

UNIVERSITY OF CAMBRIDGE

# Derivation of next generation research models of Parkinson's disease

**Rizwan Ansari**

SUPERVISORS: Dr. Emmanouil Metzakopian and Dr. Gabriel Balmus

This dissertation is submitted for the degree of

Doctor of Philosophy

October 2020



HUGHES HALL

UNIVERSITY OF CAMBRIDGE

## Table of Contents

*DECLARATIONS* ..... *10*

*ACKNOWLEDGMENTS* ..... *11*

*ABBREVATIONS* ..... *12*

*1.SUMMARY* ..... *13*

*2.INTRODUCTION* ..... *15*

**2.1 Parkinson's Disease** ..... **15**

**2.2 History and Milestones** ..... **15**

**2.3 Symptoms** ..... **17**

**2.4 Treatments** ..... **17**

**2.5 Neuropathology** ..... **18**

**2.6 Etiology** ..... **22**

**2.7 Alpha-synuclein** ..... **24**

**2.8 Pathogenesis** ..... **26**

**2.9 Endoplasmic Reticulum Stress and Unfolded Protein Response (UPR)** ..... **27**

**2.10 Mitochondria Dysfunction and Oxidative Stress** ..... **29**

**2.11 Prion-Like Spreading of Alpha-synuclein** ..... **32**

**2.11 Non-Cell-Autonomous Neurodegeneration in PD** ..... **33**

**2.12 Midbrain Dopaminergic Neuron Development *In Vivo*** ..... **34**

**2.13 Dopamine Metabolism** ..... **39**

**2.14 Models of PD** ..... **40**

**2.17. Single-Cell RNA Sequencing Technologies and Its Applications** ..... **49**

**2.18 PhD Objectives** ..... **53**

*3. HUMAN MIDBRAIN ORGANOID MODEL* ..... *55*

**3.1 Background** ..... **55**

**3.2 Results** ..... **56**

3.2.1 Generation of human midbrain organoids using a modified Jo et al. 2016 protocol ..... 56
3.2.2 Optimizing the Jo et al. 2016 Protocol to Enhance LMX1A and FOXA2 Expression ..... 58
3.2.3 Addressing Low LMX1A Expression in HMOs: CHIR Titration Experiment ..... 61
3.2.4 Refining CHIR Concentration for Enhanced LMX1A Expression in HMOs ..... 63
3.2.5 Exploration of Additional Optimization Conditions ..... 65
3.2.6 Generation and Characterization of Human Midbrain Organoids Using a Combined Approach of Jo et al. 2016 and Kriks et al. 2011 Protocols (Ansari protocol) ..... 67
3.2.7 Generation of Midbrain Organoids from Healthy and Patient iPSCs with an SNCA Triplication and It's Isogenic Control ..... 69
3.2.8 Cellular heterogeneity in HMOs following differentiation ..... 70
3.2.9 Cell Composition and Robustness of Differentiation in Human Midbrain Organoids ..... 74
3.2.10 HMOs integration with in-vitro datasets ..... 75
3.2.11 Comparing HMO Neuronal Cell Types to Human SNpc *In-Vivo* Neuronal Nuclei ..... 78
3.2.12 Assessment of Dopaminergic Marker Genes Across In-Vivo and In-Vitro Datasets ..... 81
3.2.8 Neuromelanin accumulation in Human Midbrain Organoids ..... 82
3.2.13 Evaluating Ansari Protocol Reproducibility ..... 84
3.2.14 Comparing Cellular Heterogeneity and Gene Expression in HMOs Cultured with Matrigel and Geltrex ..... 85
3.2.15 Comparison of Ansari and Unmodified Jo et al. 2016 Protocols ..... 86
**3.3 Discussion and Future Directions ..... 88**
3.3.1 Neurospheres Generation Using Kriks et al. Protocol ..... 89
3.3.2 Absence of Glial Cells (Astrocytes & Oligodendrocytes) In Midbrain Organoids 90
3.3.3 Absence of Microglia in Midbrain Organoids ..... 90
3.3.4 Necrotic Core in Midbrain Organoids ..... 91
3.3.5 Vascularization of Midbrain Organoids ..... 91
3.3.6 Midbrain Organoids on Air-Liquid Interface (ALI) ..... 92
3.3.7 Identification and Removal Of Stressed Cell Populations In Brain Organoids ..... 92

3.3.8 oDAn3 Dopaminergic Neuronal Population in Midbrain Organoid Model ..... 93
3.3.9 Mechanical Cutting of Brain Organoids ..... 93
3.3.10 Maturity of Midbrain Organoid Model ..... 94

***4. PD DISEASE MODELLING ..... 95***

**4.1 Background ..... 95**

**4.2 Results ..... 97**

4.2.1 Elevated SNCA Expression and α-Synuclein Protein Observed in SNCA-3x Organoids ..... 97
4.2.2 SNCA-3x impacts Endoplasmic Reticulum Protein Folding, Translation, and Oxidative Phosphorylation in oDAn1 Cells ..... 99
4.2.3 Effect of SNCA Triplication on oDAn2 and oDAn3 ..... 101
4.2.4 Validation of SNCA-3x Mutation Signatures Using qPCR ..... 103
4.2.5 Validation of SNCA-3x Mutation Signatures Using scRNA-seq ..... 105
4.2.6 Variations in Alpha-synuclein Protein Levels Over Time and Across Batches ..... 107
4.2.7 Effect of L-MOSES on Alpha-synuclein Protein Levels and SNCA-3x Associated Transcriptomic Profile ..... 108
4.2.8 Effect Of L-MOSES On SNCA-3x and SNCA-Corr oDAn1 Cells ..... 110
4.2.9 Identification of Rotenone-Sensitive and Resilient oDAn1 Cells through In-Silico Analysis ..... 112
4.2.10 Derivation and Characterization of Human Chimera Midbrain Organoids ..... 118
4.2.11 Investigation of non-cell autonomous effect on healthy and disease cells ..... 119

**4.3 Discussion and Future Directions ..... 123**

4.3.1 Dysregulation in Protein Translation in SNCA-3x Dopaminergic Neurons ..... 124
4.3.2 ER Stress and UPR in SNCA-3x SNCA-3x Dopaminergic Neurons ..... 125
4.3.3 Mitochondrial Dysfunction and Oxidative Phosphorylation in SNCA-3x dopaminergic Neurons ..... 126
4.3.4 oDAn1 vs oDAn2 vs oDAn3 ..... 127
4.3.5 Validation Experiments ..... 128
4.3.6 The Impact Of L-MOSES On Alpha-Synuclein Levels and Transcriptomic Characteristics of The SNCA-3x Mutation ..... 129
4.3.7 Impact of Oxidative Stress on Dopaminergic Neurons ..... 130
4.3.8 Chimera Midbrain Organoids ..... 132

*5. Induced Dopaminergic Neurons* ..... *133*

**5.1 Background** ..... **133**

**5.2 Results** ..... **135**

5.2.1 Induced Dopaminergic Neurons Differentiation Using Standard PiggyBac Vectors ..... 135

5.2.2 Induced Dopaminergic Neurons Differentiation Using NGN2 Polycistronic Constructs via Transient Expression ..... 136

5.2.3 Induced Dopaminergic Neurons Differentiation Using Subcloned iPSCs with NGN2 +3TFs ..... 137

5.2.4 Induced Dopaminergic Neurons Differentiation Using Subcloned BON NGN2 iPSCs Utilizing NGN2+3TFs ..... 138

5.2.5 A Screen Aims to Identify Optimal Combination Of TF For Induced Dopaminergic Neurons Differentiation ..... 139

5.2.6 Induced Dopaminergic Neurons Differentiation Using ATOH1 Polycistronic Vectors via Inducible Expression ..... 141

5.2.7 Induced Dopaminergic Neurons Differentiation Using ATOH1 Monocycistronic Vector Via Inducible Expression ..... 143

5.2.8 Induced Dopaminergic Neurons Differentiation Using Subcloned iPSCs With ATOH1+6TFs ..... 145

5.2.9 A Screen To Identify Optimal Combination Of TF For Induced DA Neurons Differentiation ..... 146

5.2.10 Single Cell RNA Sequencing Analysis Of Induced DA Neurons Derived Using NGN2+6TFs And ATOH1 + 6TFs Combination ..... 148

5.2.11 Characterization of Induced Dopaminergic Neurons Derived By Transient Overexpression Of NGN2 Alone, NGN2+6TFs, ATOH1 Alone, ATOH1+6TFS ..... 151

5.2.12 Characterization of Induced Dopaminergic Neurons Derived Using Controlled Ectopic Expression Of ATOH1 Alone And ATOH1+ Shh_FGF8a ..... 154

**5.3 Discussion and Future Directions** ..... **156**

5.3.1 PiggyBac vs Lentivirus Vectors ..... 157

5.3.2 ATOH1 vs ASCL1 ..... 158

5.3.3 ATOH1 in Combination with SHH and FGF8b ..... 159

5.3.4 The Importance of Proneural TF Ectopic Expression in Induced Neurons ..... 159

5.3.5 Inducible vs. Constitutive Expression Vectors ..... 160

5.3.6 Single Cell RNA Sequencing Analysis of ATOH1/NGN2 + 6TFs ..... 160
5.3.7 ATOH1 + 3TFs ..... 161
5.3.8 TFs Combination To Induce Dopaminergic Neurons Of Central Nervous System Identity ..... 161
5.3.9 Integration of TFs in Safe Harbour Sites Using OPTI X System ..... 161
5.3.10 iPSCs Derived vs Fibroblast Derived induced Neurons ..... 162

***6. Unravelling Cell Type Specific Response to Parkinson's Disease ..... 163***

**6.1 Background ..... 163**

**6.2 Results ..... 164**

6.2.1 Single Nucleus RNA-Seq Reveals Cell Type Heterogeneity in Human SNpc ..... 164
6.2.2 Robust Gene Expression of PD GWAS Associated Genes In Substantia Nigra Pars Compacta Cell Types ..... 169
6.2.3 Neuronal Subpopulations in Aged Human Snpc And Their Response In Sporadic PD ..... 171
6.2.4 Astrocytes Subpopulations in Aged Human SNpc and Their Response In Sporadic PD ..... 176
6.2.5 Microglia Subpopulations in Aged Human SNpc and Their Response In Sporadic PD ..... 182
6.2.6 Oligodendrocyte Subpopulations in Aged Human SNpc and Their Response In Sporadic PD ..... 187
6.2.7 TH Among Several Common Genes Expressed in Depleted Cell Types In Sporadic PD Samples ..... 192

**6.3 Discussion and Future Directions ..... 196**

6.3.1 Cell Types Composition in SNpc ..... 196
6.3.2 Expression Of PD GWAS Associated Genes In SNpc Cell Types ..... 198
6.3.3 Dopaminergic Neurons0 Degeneration in PD Condition ..... 198
6.3.3 GABA Neurons3 Degeneration in PD Condition ..... 199
6.3.4 UPR And Oxidative Stress in Astrocytes ..... 201
6.3.5 Microglia States in Substantia Nigra Pars Compacta ..... 201
6.3.6 "Dysfunctional" Oligodendrocyte Subpopulations In PD ..... 202
6.3.7 UPR And Oxidative Stress in Oligodendrocyte Subpopulation ..... 203
6.3.8 Shared Vulnerability of *TH*-Enriched Glial Subpopulations in Substantia Nigra 203

6.3.9 Limitation of The Study.....205

*7. Methods and Material*.....*206*

**7.1 iPSC Culture**.....**206**

**7.2 Plate Coating**.....**206**

**7.3 Antibiotic Selection**.....**207**

**7.4 iPSC Subcloning**.....**207**

**7.5 HEK293T culture**.....**207**

**7.6 DNA Extraction**.....**207**

**7.7 PCR**.....**207**

**7.7 Human Midbrain Organoid Derivation- Jo et al. 2016 Protocol**.....**208**

**7.8 Human Midbrain Organoid Derivation- Ansari Protocol**.....**209**

**7.9 Oxidative Stress Treatment**.....**209**

**7.10 L DOPA Treatment**.....**209**

**7.11Mitomycin C Treatment**.....**209**

**7.12 L-MOSES Treatment**.....**210**

**7.13 Cortical Induced Neurons Differentiation**.....**210**

**7.14 Immunohistochemistry**.....**210**

**7.15 RNA Extraction**.....**211**

**7.16 RNA Quantification**.....**211**

**7.17 RNA Quality**.....**211**

**7.18 cDNA Synthesis**.....**212**

**7.19 Quantitative qPCR**.....**212**

**7.20 Protein extraction**.....**212**

**7.21 Protein Quantification**.....**212**

**7.22 Western Blotting**.....**213**

**7.23 Sample Preparation For scRNA-Seq**.....**213**

**7.24 SNCA Triplication Screening ..... 214**

**7.25 Plasmids Designs ..... 215**

**7.26 Plasmid Amplification ..... 215**

**7.27 Bacterial Transformations ..... 215**

**7.28 Plasmid Extraction ..... 215**

**7.29 Induced Dopaminergic Neurons Differentiation ..... 216**

**7.30 Induced Dopaminergic Neurons Differentiation- Sagal et al. 2014 protocol ..... 216**

**7.31 Induced Dopaminergic Neurons scRNA-seq sample preparation ..... 217**

**7.32 Induced Dopaminergic Neurons Immunocytochemistry ..... 217**

**7.34 Cloning ..... 217**

**7.35 Human Midbrain Organoids Single Cell RNA sequencing Data Analysis ..... 218**

7.35.1 Processing and Filtering of Single Cell RNA-seq ..... 218

7.35.2 Dimensionality Reduction ..... 218

7.35.3 Cell Type Annotation ..... 218

7.35.4 2D And *In Vivo* Single-Cell Gene Expression Integration ..... 219

7.35.5 Differential Expression and Gene-Set Enrichment Analysis ..... 219

7.35.6 Rotenone Treatment Analysis ..... 219

7.35.7 Chimera Genotype Calling ..... 220

7.35.8 Visualization ..... 220

7.35.9 Wang et al. SNpc Neuronal Nuclei Data Set ..... 220

7.35.10 Fiorenzano et al. Ventral Midbrain Organoids ..... 220

**7.36 Post-Mortem Substantia Nigra Single Nucleus RNA Sequencing Data Analysis ..... 221**

7.36.1 Selection of Donors ..... 221

7.36.2 Isolation Of Nuclei From Frozen Post-Mortem Brain Tissue And Single Nuclei RNA-Seq ..... 221

7.36.3 Pre-Processing Single Nuclei RNA-Seq Data ..... 221

7.36.4 Quality Control ..... 221

7.36.5 Integration of The Datasets And Clustering ..... 222

7.36.6 Marker Identification ..... 223

7.36.7 Datasets of PD GWAS Associated Genes ..... 223
7.36.8 Differential Gene-Expression Analysis ..... 223
7.36.9 Pathway Analysis ..... 224
7.36.10 Spatial Transcriptomics - Sample Preparation ..... 224
7.36.11 Spatial Transcriptomics - Molecular Cartography ..... 224
7.36.12 Spatial Transcriptomics – Object Segmentation ..... 225
7.36.13 Spatial Transcriptomics - Quality Control ..... 225
7.36.14 Spatial Transcriptomics - Cell Type Classification and Clustering ..... 225

# DECLARATIONS

This thesis is the result of my own work and includes nothing that is the outcome of work done in collaboration except where specifically indicated in the text. This work has not been previously submitted to any university or institution for any degree or diploma or other qualification. In accordance with the Clinical and Veterinary Medicine Degree Committee, this thesis does not exceed the prescribed word limit (60,000 words).

The work presented in sections 3 and 4 of the thesis was conducted in collaboration with Nikolaos Patikas, resulting a publication in the Neuroscience Research journal (details provided below). Nikolaos Patikas and I share the co-first authorship of this paper.

**Publication Details:**

Patikas N, Ansari R, Metzakopian E. Single-cell transcriptomics identifies perturbed molecular pathways in midbrain organoids using α-synuclein triplication Parkinson's disease patient-derived iPSCs. Neurosci Res. 2023 Oct; 195:13-28. doi: 10.1016/j.neures.2023.06.001. Epub 2023 Jun 3. PMID: 37271312.

The research presented in section 6 of the thesis was conducted in collaboration with Araks Martirosyan and Francisco Pestana. The manuscript prepared from this work is currently in the second round of review for the Molecular Neurodegeneration journal. Araks Martirosyan, Francisco Pestana, and I share the co-first authorship of this paper. An earlier version of the manuscript is already available online on the bioRxiv preprint server. I was not initially mentioned as an author in that manuscript because I joined the project shortly before it was made available on bioRxiv server. Due to my significant contributions in improving the manuscript, which included highlighting a potential novel mechanism associated with PD, I have been included as a co-first author in the revised version under consideration in the Molecular Neurodegeneration journal.

**Publication Details (preprint):**

Martirosyan A, Ansari R, Pestana F. Unravelling cell type-specific responses to Parkinson's Disease at single cell resolution.bioRxiv. 2023 Jan. https://doi.org/10.1101/2023.01.04.522691.

# ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor, Dr. Emmanouil Metzakopian, for his support and guidance throughout my entire PhD journey. Without his mentorship, this PhD journey would not have been possible. I extend my heartfelt thanks to Nikolaos Patikas, Araks Martirosyan, and Francisco for their invaluable assistance in single cell RNA sequencing data analysis and for providing plots for my thesis. I am deeply indebted to my wife, Maviya Noor Mohammadi, for her motivation and support at various stages of my PhD, especially during challenging times. Her role in my academic journey has been pivotal. I would like to express my appreciation to my parents for their support in all my life decisions, including my pursuit of a PhD. I am grateful to Allah for His blessings, which have enabled me to achieve so much during my PhD. A special thanks goes out to all my lab mates, with gratitude to Sandeep Rajan, Sofia Pavlou, William Plumbly, Hugo Fernandes, Emma Wilson, Silvia Hnatova, Kimberly Cheam, Jose Vicente, and Stefanie Foskolou for their invaluable assistance and support in the lab. I would like to acknowledge the Cancer Research UK CI Genomics department for their assistance in performing single cell RNA sequencing on my midbrain organoid samples and induced dopaminergic neuron samples. I also want to acknowledge BioRender for the schematics used in this thesis. Once again, thank you to everyone who has played a part in my academic journey. Your support has been instrumental in my success.

# ABBREVATIONS

# 1.SUMMARY

Parkinson's disease (PD) is the second most common neurodegenerative disorder following Alzheimer's disease. PD is characterized by the degeneration of dopaminergic neurons in the substantia nigra region of the midbrain and the presence of intracellular inclusions, notably Lewy bodies composed of the alpha-synuclein protein (encoded by the SNCA gene), in surviving neurons. The precise molecular and cellular mechanisms governing PD's onset, progression, and pathology remain incompletely understood. The development of advanced research models is pivotal in unravelling the mysteries surrounding PD's pathogenesis and progression.

Although traditional monolayer cell culture models have significantly enhanced our understanding of the disease, their results have not consistently translated into clinical success. These models lack the desired cellular diversity and fail to mimic *in vivo* physiological conditions. To address these limitations, our goal was to establish a human midbrain organoid model that closely replicates *in vivo* cellular diversity, tissue cytoarchitecture, and physiology. Within this thesis, we introduce a robust novel method for deriving human midbrain organoids from both healthy individuals and PD patients induced pluripotent cells (hiPSCs). We characterized the cellular diversity within these midbrain organoids using single-cell RNA sequencing (scRNA-seq), revealing the presence of dopaminergic neuron populations resembling *in vivo* substantia nigra neurons. We demonstrated the utility of the midbrain organoid model by capturing a comprehensive transcriptomic profile of dopaminergic neurons in response to *SNCA* triplication mutation and oxidative stress. The single-cell transcriptomics highlighted perturbations in oxidative phosphorylation and protein translation in response to *SNCA*-3X mutation. The vulnerability of dopaminergic neurons to oxidative stress was associated with the enrichment of cholesterol biosynthesis and synaptic signalling processes.

Recent studies employing single-cell RNA sequencing have shown that neuronal models derived from human iPSCs contain undesired cellular heterogeneity and noticeable variability across different batches. Our aim is to develop a monolayer neuronal model that can offer highly homogeneous substantia nigra neurons on a large scale within a short timeframe through the ectopic expression of transcription factors (TFs). In this study, we validated a significant finding: ectopic expression of *ATOH1* alone is sufficient to generate

Tyrosine hydroxylase (*TH*)-positive neurons. Additionally, we identified several TF combinations that also induced *TH*-positive neurons. Single-cell RNA sequencing analysis of neurons generated through ectopic expression of combination of (*ATOH1* or *NGN2*), *FOXA2, LMX1A, NURR1, PITX3, SOX6,* and *MSX1* revealed three distinct *TH*-positive neuronal clusters. However, these neurons were also found to be positive for *PRPH*, a peripheral nervous system marker. The characterization of *TH*-positive neurons induced by *ATOH1* alone, with or without midbrain patterning factors such as SHH and FGF8b, also showed positivity for *PRPH*. Further refinement of the TF combination is needed to correct the regional identity of the *TH*-positive neurons. Additionally, the generated neurons require characterization beyond scRNA-seq/qPCR techniques.

Human post-mortem material remains the gold standard for formulating hypotheses and subsequently testing them in both *in vivo* and *in vitro* models of PD. It also serves as an invaluable tool to validate hypotheses derived from experimental models of PD, confirming their relevance to human disease. In this study, we generated a comprehensive single-nucleus RNA sequencing dataset of a prominent PD-affected region of the midbrain, the substantia nigra pars compacta, from a large cohort of 15 sporadic PD patients and 14 control individuals. This study involved transcriptomic-level characterization of all major cell types, including their subpopulations, present in the brain. We demonstrated specificity of GWAS PD associated genes to different cell types. Importantly, we identified glial subpopulations enriched in the *TH* gene, which was found to be depleted in PD samples. This dataset provides a valuable resource for future hypothesis-driven experiment and represents a significant step forward in understanding PD's disease mechanisms.

# 2.INTRODUCTION

## 2.1 Parkinson's Disease

Neurological disorders represent a significant source of disability worldwide, with the incidence of these disorders, including Parkinson's disease (PD), on the rise in our current aging population (Jankovic et al., 2020). PD is a progressive neurological disorder, ranking as the second most common after Alzheimer's disease (Poewe et al., 2017). According to the global burden of disease study, the global PD population reached 6.1 million individuals in 2016, with projections indicating an anticipated increase to approximately 13 million by 2040(Dorsey et al., 2018).

## 2.2 History and Milestones

PD was first described by James Parkinson in an essay titled "An Essay on the Shaking Palsy" in 1817 (Parkinson, 1817). Within this essay, Parkinson outlined the clinical symptoms observed in six individuals. Parkinson observed the degenerative nature of the disease and described classic parkinsonian features, including tremor, gait abnormalities, and distinctive postures (McDonald et al., 2018). Nearly five decades later, in 1872, Jean-Martin Charcot officially bestowed the name "Parkinson's disease" upon this medical condition in honour of James Parkinson's pioneering contribution. Jean-Martin Charcot distinguished PD from conditions such as multiple sclerosis (MS) and other disorders marked by tremors (Goetz et al., 2011).

\

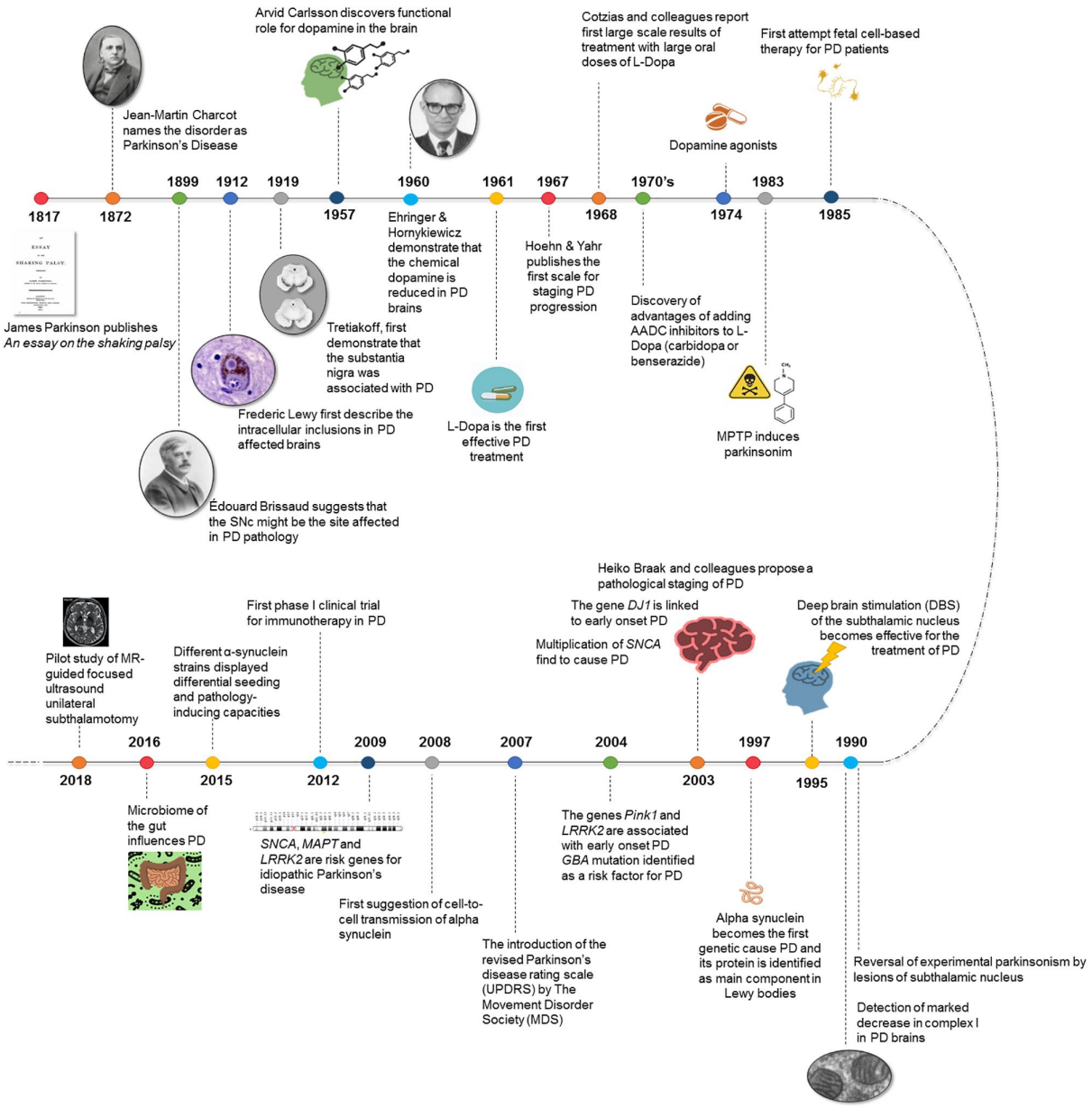



**Figure 1.1**. A timeline of major milestones in Parkinson's disease. (Image reference: Del rey et al., 2018)

In 1899, Edouard Brissaud proposed the substantia nigra as a potential locus for PD pathology. Subsequently, in the years 1912 to 1919, Frederick Lewy and Konstantin Tretiakoff identified inclusions situated within the substantia nigra as indicative of PD. In 1960, Oleh Hornykiewicz showed a reduction in dopamine levels within the striatum of PD-affected brains. Based on this discovery, in 1961, Andre Barbeau introduced a groundbreaking milestone by utilizing L-DOPA, a dopamine precursor, as the first orally administered and effective anti-PD medication. The year 1970 saw the introduction of Deep Brain Stimulation

(DBS) for alleviating PD-associated tremors. Meanwhile, the revelation of MPTP-induced parkinsonism in 1983 led research into environmental factors role in PD, supported by confirmation in animal models. The work of George Cotzias and Donald Brian Calne in testing the antiparkinsonian effects of apomorphine and bromocriptine, a dopamine receptor agonist, marked a significant advancement. The identification of alpha-synuclein's pivotal role in Lewy bodies, and its linkage to the first genetic cause of PD was accomplished by Mihael Polymeropoulos in 1997. In 2003, John Nutt conducted the pioneering double-blind controlled trial of cell-based therapy in PD. In the same year, Heiko Braak proposed a groundbreaking pathological staging framework for PD. The period spanning 2003 to 2007 witnessed the identification of crucial genes such as *DJ1, PINK1, NR4A2, LRRK2, GBA*, and *ATP13A2* in the context of PD. The year 2014 marked another milestone with the discovery of cell-to-cell transmission of alpha-synuclein in PD, adding to our understanding of the disease's progression and mechanisms.

## 2.3 Symptoms

The symptoms of PD include both motor and non-motor aspects. Among the motor symptoms, the most frequently recognized in PD are the loss of postural reflexes, bradykinesia, rest tremor, and rigidity. Additional motor symptoms linked to PD comprise respiratory disruptions, bulbar dysfunction, and neuro-ophthalmological anomalies. Non-motor manifestations encompass cognitive impairments, mood fluctuations, sleep disturbances, and autonomic dysfunctions. These symptoms collectively put a profound impact on the quality of life for individuals with PD, highlighting the essential need for comprehensive healthcare support to facilitate effective management (Ciga et al., 2020; Pang et al., 2019; Varadi 2020).

## 2.4 Treatments

Due to our limited understanding of the disease, a cure for PD remains elusive at present. Current treatments for PD primarily focus on symptom management; none possess the capability to slow down or halt the progression of the disease. A pivotal turning point occurred with Arvid Carlsson's discovery of dopamine in the mammalian brain in 1958, led to two profound breakthroughs: firstly, the finding that the loss of neurons in the substantia nigra leads to dopamine deficiency; secondly, the oral administration of the dopamine precursor levo-dopa (L-DOPA) effectively alleviates most PD symptoms (Reference 7). Nonetheless, a significant proportion of patients undergoing long-term L-DOPA treatment develop a challenging

involuntary movement disorder referred to as "dyskinesias," which poses difficulties in management and significantly diminishes their quality of life (Pandey et al.,2019).

Various other drugs, including Catechol-O-methyltransferase inhibitors, monoamine oxidase type B inhibitors, and dopamine agonists, have been employed to compensate for the loss of striatal dopamine (Poewe et al., 2017; Iarkov et al., 2020)). In cases where drug treatments prove ineffective, DBS emerges as an alternative to manage PD symptoms (Del Rey et al., 2018). Efforts have also been directed towards dopamine replacement through cell therapy-based approaches, though the outcomes have yielded disappointing results (Del Rey et al., 2018; Fan et al., 2020). These setbacks highlight the formidable challenges posed by PD and underscore the imperative for increased scientific attention.

## 2.5 Neuropathology

The distinctive features of PD are the loss of neurons in the substantia nigra par compacta (SNpc) and the presence of intracellular alpha-synuclein aggregates in the form of Lewy bodies and Lewy neurites in the surviving neurons (Spillantini et al., 2019).

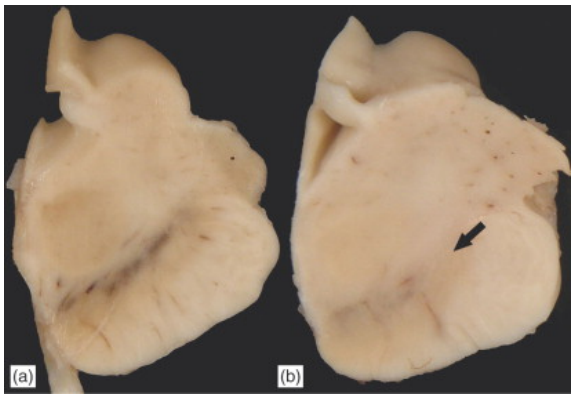



**Figure 1.2:** Transverse section of the midbrain from (a) aged matched control (b) PD patient. Black arrow indicating loss of neuromelanin pigmentation in PD patient. Image reference: Petrucelli et al., 2008.

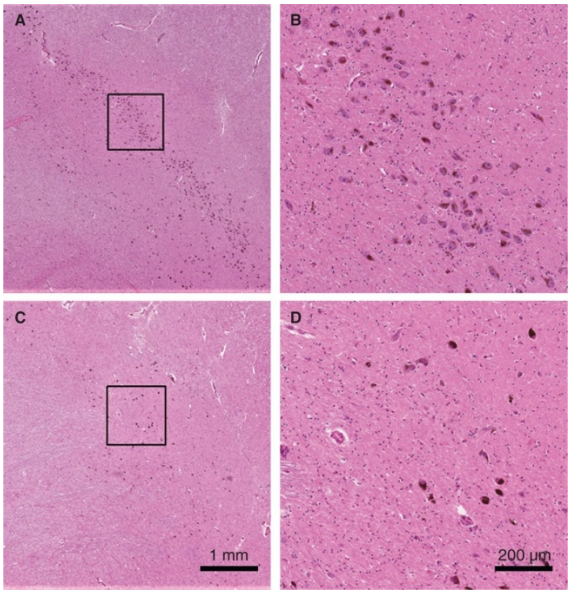



**Figure 1.3**: Coronal sections of the Substantia Nigra pars compacta (SNpc) in a Control (A & B) and a Parkinson's disease (PD) brain (C & D) section. The dark brown cells represent neuromelanin-containing dopaminergic neurons, which are remarkably lost in PD. Image courtesy: Kouli et al., 2018.

At a macroscopic level, the brains of individuals with PD may outwardly appear unremarkable, albeit demonstrating mild frontal cortex atrophy, and ventricular dilation, in most cases. However, upon analysing transverse sections of the brainstem in PD cases, a distinct absence of the densely pigmented regions within SNpc and the locus coeruleus becomes evident. This loss of pigmentation has been observed to correspond directly with the depletion of noradrenergic neurons in the locus coeruleus and neuromelanin-containing neurons in the SNpc (Dickson et al., 2012). Strikingly, a particular subset of neuromelanin-containing dopaminergic neurons, specifically the A9 neurons, experiences primary loss within the SNpc, while other neuronal and glial cell types tend to remain relatively unaffected. Studies have indicated that up to 30% of SNpc dopaminergic neurons may be lost by the time PD symptoms manifest, and this loss can escalate to 60% or more over time (Cheng et al., 2012). Such substantial depletion of SNpc dopaminergic neurons leads to decreased dopamine levels in the striatum, ultimately contributing to the cardinal motor symptoms of PD (Kouli et al., 2018).

The scope of cell loss in PD extends beyond the SNpc and locus coeruleus. Areas such as the olfactory bulb, subcortical nuclei, the dorsal motor nucleus of the vagus nerve, and the nucleus basalis of Meynert also exhibit widespread cellular loss (Giguere et al., 2018). Additionally, it's noteworthy that the impact of PD isn't confined solely to the dopaminergic

neurotransmitter system. Research highlights the influence on other neurotransmitter systems, including glutamatergic, GABAergic, serotoninergic, and cholinergic systems (Kalia et al., 2012). This diverse neurotransmitter involvement may contribute to the varying responses among PD patients to dopamine therapy.

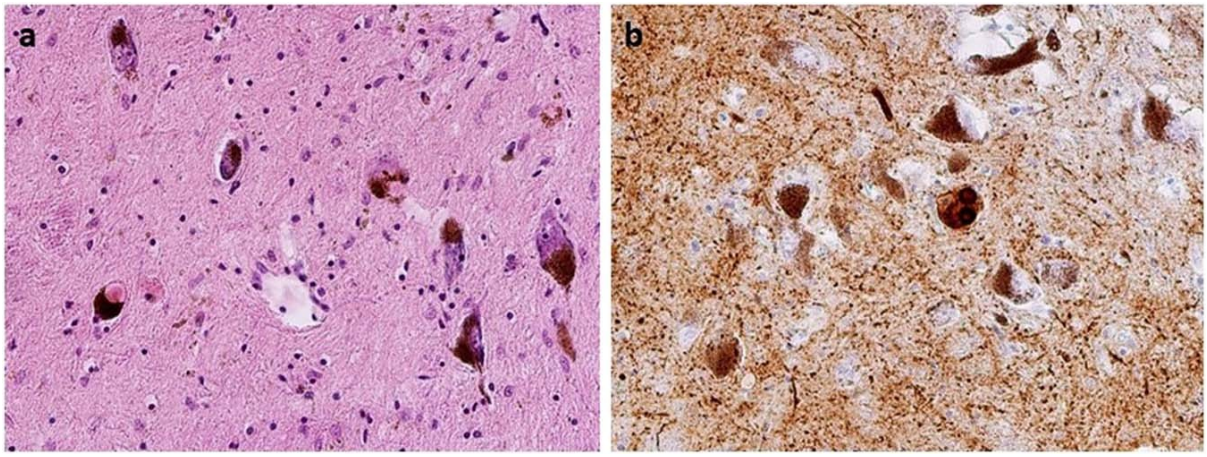



**Figure 1.4:** (a) A typical morphology of Lewy bodies in the neuromelanin pigmented dopaminergic neurons of SNpc, H&E staining. (b) Lewy bodies and dystrophic neurites stained with alpha-synuclein antibody. Image courtesy: Mensikova et al., 2022.

At the microscopic level, surviving neurons within PD-afflicted brains exhibit the presence of Lewy bodies within their cell bodies, often accompanied by Lewy neurites (dystrophic neurites). A characteristic Lewy body in the SNpc of PD cases displays a central granular and fibrillar core encased by a surrounding halo, varying in size from 5 to 25μm (Spillantani et al, 1991; Kouli et al, 2018). Lewy bodies primarily consist of aggregated, insoluble forms of alpha-synuclein. In addition to alpha-synuclein, Lewy bodies include a range of other proteins, including tau, ubiquitin, cytoskeletal proteins, heat shock proteins (HSPs), as well as mitochondria, lipid membrane, proteasomal and lysosomal components (Gomez-Benito et al., 2020; Moussaud et al., 2014; Kouli et al., 2018).

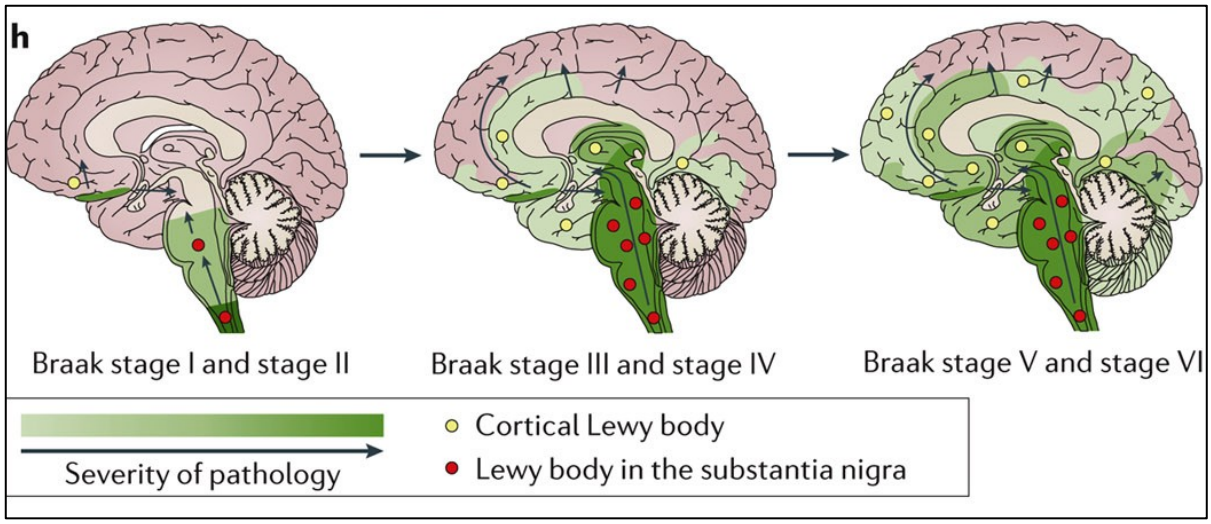



**Figure 1.5:** Schematic illustration of Braak staging model for PD disease progression. Poewe et al 2017.

The Braak staging model, first introduced in 2003 by Braak and colleagues, is the most widely cited model for describing the neuropathological progression of PD (Braak et al., 2003). This model suggests that the initial pathological changes manifest in the medulla and olfactory bulb (Braak stages 1 and 2). During these early stages (1 and 2), patients do not exhibit any motor symptoms; however, they may display non-motor symptoms such as autonomic issues (e.g., constipation) and sleep-related disorders. As the disease advances to stages 3 and 4, the pathology extends to include the SNpc as well as other midbrain and basal forebrain structures. At this juncture, patients start displaying the cardinal clinical motor symptoms closely linked to PD, and it is at this point that the disease is typically diagnosed (Armstrong et al., 2020). Progressing further to stages 5 and 6, the disease enters its most advanced phase, characterized by the spread of pathology to the cerebral cortex (Kouli et al., 2018). This advanced stage is generally accompanied by the onset of cognitive impairment and hallucinations. It is important to acknowledge that the Braak staging model is based on the distribution of Lewy bodies rather than neuronal cell loss. Furthermore, despite its widespread acceptance, the model has yet to be definitively validated through comprehensive studies.

## 2.6 Etiology

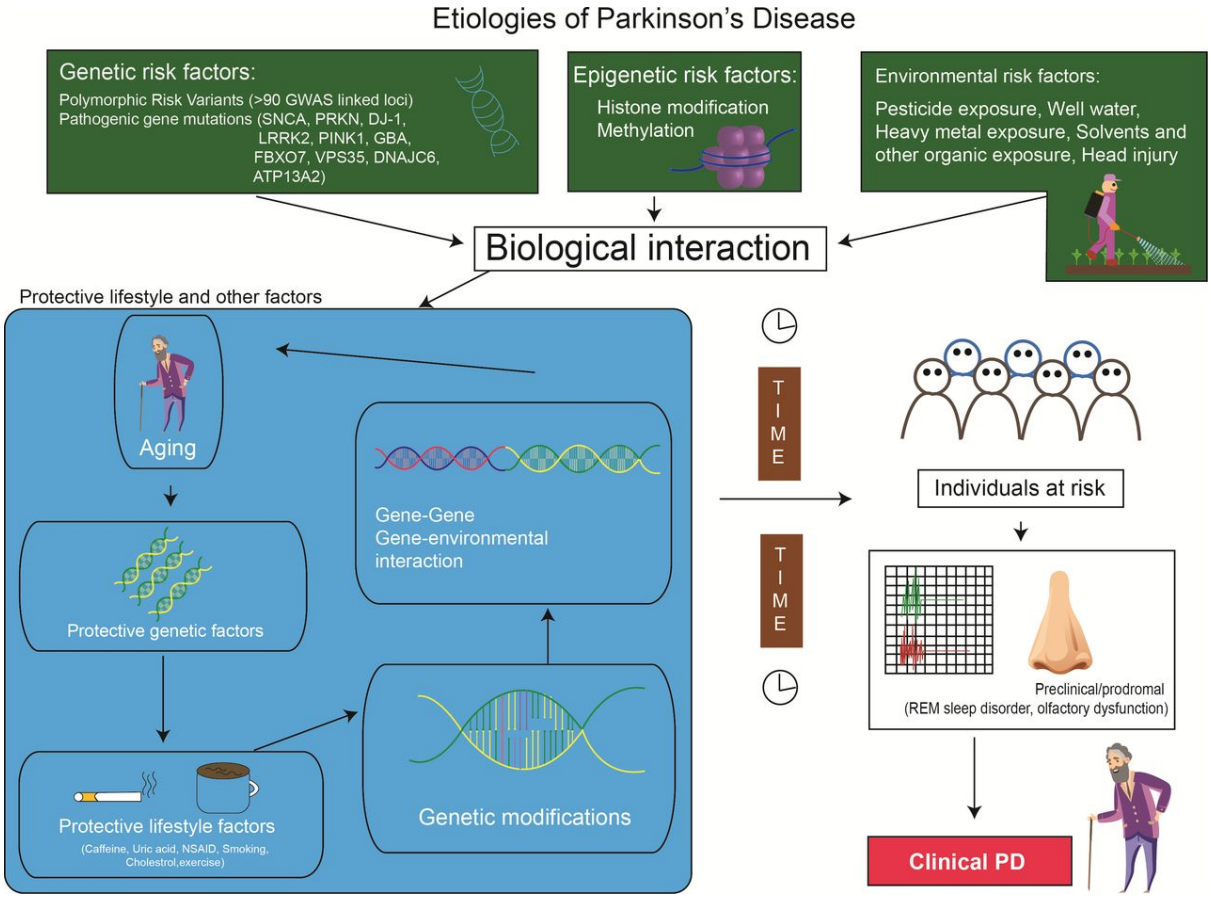



**Figure 1.6:** Schematic illustration depicting biologic interaction between genetic, epigenetic, and environmental factors contributing to the PD developmental. Image courtersy: Jankovic et al 2020.

PD is a multifactorial disease, including sex, aging, genetic, epigenetic and environmental elements that contribute to its development. Aging is the biggest risk factor, with the onset of PD typically occurring around the age of 60. The frequency of PD appears higher in men compared to women (Jankovic et al., 2020).

Among the acknowledged environmental risk factors, several factors carry significant weight, including head injuries, consumption of dairy products, pesticide exposure, heavy metal exposure, and others (Ball et al., 2019). Of notable importance, individuals exposed to 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) have demonstrated the development of nigrostriatal degeneration, accompanied by cardinal features of PD (Balestrino et al., 2017).

Within the body, the lipophilic MPTP is capable of crossing the blood-brain barrier. Once inside, it undergoes metabolic transformation catalysed by the enzyme monoamine

oxidase B (MAO-B), primarily found in glial cells, especially astrocytes (Ransom et al., 1987). This process leads to the generation of MPP+ (1-methyl-4-phenylpyridinium), a potent neurotoxin. Operating as a selective inhibitor of mitochondrial complex-I, MPP+ damage specifically dopaminergic cells situated in the substantia nigra. Notably, MPP+ exhibits a strong affinity for dopamine receptors, resulting in its preferential uptake and accumulation within dopamine neurons, ultimately culminating in their demise (Langston et al., 2017).

Curiously, the dopamine neurons in the ventral tegmental area (VTA) appear to be largely resilient to the toxic effects of MPP+. This phenomenon could potentially be attributed to the presence of calbindin (Lavoie et al., 1991). Following uptake by nigral dopamine neurons, MPP+ becomes markedly concentrated within mitochondria, profoundly influencing their functionality through complex I inhibition. This inhibition consequently leads to diminished ATP production and subsequently generation of toxic reactive oxygen species (Ramsay et al., 1986). Likewise, rotenone and paraquat (compounds structurally similar to MPTP) have exhibited the capacity to selectively affect mitochondrial complex I and induce targeted damage to nigral dopamine neurons in animal models (Di Monte et al., 1986; Betarbet et al., 2000).

PD presents itself in two forms: familial and sporadic, each with distinct origins (Reference 12). Familial PD is inherited genetically, either through an autosomal recessive or dominant pattern. Conversely, sporadic, also known as idiopathic PD, arises due to unknown gene-gene and/or gene-environment interactions. A mere 10-15% of PD instances are attributed to be familial, while the majority stem from sporadic origins (Ball et al., 2019).

The identification of multiple families displaying Mendelian inheritance patterns (both dominant and recessive) supported the genetic influence on the PD development. This led to the discovery the first PD-associated gene, *SNCA*, in 1997. Through the progression of genetic methodologies and comprehensive population analyses, including genome-wide association studies (GWAS), researchers have identified more than 20 monogenic variations of PD. Additionally, the identification of over 100 loci as potential contributors to PD risk has emerged (Jankovic et al.,2018; Ciga et al., 2020).

The literature includes numerous autosomal dominant, autosomal recessive, and susceptibility genes associated with PD. Among the autosomal dominant genes are *SNCA,*

*LRRK2, VPS35,* and *DNAJC13*. Autosomal recessive genes include *Parkin, PINK1, DJ1, FBXO7*, and *DNAJC6*. Additionally, susceptibility genes like *GBA, MAPT, NR4A2,* and *HLA* have been reported (Emamzadeh et al., 2018).

Mutations in *LRRK2* give rise to the most prevalent form of autosomal dominant monogenic PD. Within the *LRRK2* gene, six distinct pathogenic mutations have been verified, with p.G2019S standing out as the most prevalent. Globally, p.G2019S accounts for 4% of familial PD cases and 1% of sporadic PD cases. Mutations in *PRKN, PINK1,* and *DJ1* are implicated mitochondria homeostasis dysregulation. Both *PRKN* and *PINK1* play integral roles in the process of mitophagy. Mutations in *PRKN* are the predominant cause of autosomal familial PD, contributing to as much as 50% of early onset instances. *DJ1*, on the other hand, contributes to the regulation of calcium flux within mitochondria and offers protection against oxidative stress. The mutation in *GBA*, which encodes the enzyme glucocerebrosidase, correlates with lysosomal storage disorders. Mutations in *GBA* have demonstrated a fivefold increase in PD risk, highlighting the significance of the lysosomal pathway in the pathogenesis of PD (Billingsley et al., 2018).

## 2.7 Alpha-synuclein

Undoubtedly, the pivotal role of alpha-synuclein, which is encoded by the *SNCA* gene, in the pathogenesis of PD enjoys widespread acceptance within the PD research community. The first-ever genetic link to PD was unearthed in the *SNCA* gene by Polymeropoulos et al. in 1997(Polymeropoulos et al., 1997). This groundbreaking discovery revealed a G209A substitution mutation, leading to an A53T amino acid alteration. Notably, this mutation exhibited autosomal dominant inheritance and resulted in the early onset of PD. Subsequently, two more point mutations, A30P and E46K, were identified within the *SNCA* gene, both exhibiting autosomal dominant inheritance patterns. Intriguingly, additional mutations in the *SNCA* gene came to light, including triplication and duplication mutations. Notably, these mutations exhibit a dosage effect. Triplication gives rise to early onset PD and manifests as a more severe form compared to cases of duplication. Research has unveiled that *SNCA* overexpression can induce nigral degeneration and motor impairments (Yamada et al., 2004; Koprich et al., 2017). The multiplication mutation of *SNCA* holds intriguing implications, suggesting an excess of wild-type *SNCA* could be a causal factor for PD (Pang et al., 2019). Despite the strides made in PD research, genetics can only elucidate a fraction of cases

(Billingsley et al., 2018). Extensive research is essential to unravel the intricate genetic makeup of PD, which holds promise for identifying therapeutic targets.

Alpha-synuclein is a relatively small and inherently disordered protein composed of 140 amino acids. Alpha-synuclein lacks a well-defined structure when in an aqueous solution. However, it can adopt α-helical configurations when it interacts with negatively charged lipids, like the phospholipids found in cellular membranes (Stefanis et al., 2012). The lack of a defined structure in native state is thought to be significant for alpha-synuclein function. It may allow alpha-synuclein to adopt different conformations when interacting with various cellular components, such as membranes and other proteins. This structural property also contributes to its propensity to aggregate, as it can undergo conformational changes that lead to the formation of beta-sheet-rich aggregates (Bisi et al., 2021; Meade et al., 2019).

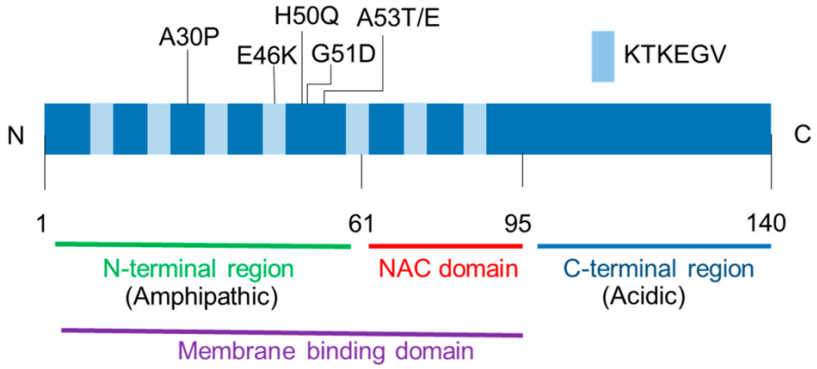



**Figure 1.7:** Alpha-synuclein protein structure. Image reference: Mori et al., 2020

Alpha-synuclein consists of three distinct domains: 1) N terminal region ranging from residues 1 to 60, 2) a central hydrophobic region spanning residues 61 to 95, which is known as the NAC (non-Aβ component) domain implicated in the protein's aggregation process, and 3) a C-terminal region (carboxyl terminus). The N terminal region serves as an amphipathic region, support interactions with lipid membranes. Intriguingly, this region is associated with all known mutations linked to synucleinopathies (neurodegenerative diseases characterized by alpha-synuclein aggregation), including A53T, A53E, A30P, E46K, H50Q, and G51D, the majority of which are found within an eight-residue span. Alpha-synuclein is part of synuclein family which also include beta synuclein and gamma synuclein. NAC domain largely separate alpha-synuclein from beta and gamma synuclein (Stefanis et al., 2012). Conversely, the C-terminal region of alpha-synuclein is marked by its acidic, highly negatively charged nature, and it is largely unstructured, which contributes to prevention of alpha-synuclein aggregation by electrostatic repulsion (Bisi et al., 2021; Burre et al., 2015). Phosphorylated alpha-synuclein

at residue S129 (pS129) has been reported in the PD brain, as well as in transgenic animal models and *in vitro* cellular models of PD. However, the effect of phosphorylation on residue 129, located in the C-terminal region of the alpha-synuclein protein, in promoting its aggregation and toxicity *in vivo*, remains unclear (Oueslati et al., 2016).

Alpha-synuclein is predominantly and ubiquitously expressed in the central nervous system, comprising 1% of total cytosolic protein. Alpha-synuclein is abundantly expressed in neurons, it is found to be expressed in glial and blood cells (Barba et al., 2022). Alpha-synuclein is physiologically localized in presynaptic terminals and is involved in essential neurological functions, such as vesicle priming, docking, fusion, vesicle trafficking, and neurotransmitter release (Jankovic et al., 2020). Within presynaptic terminals, neurotransmitter release depends on the soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) complex. Alpha-synuclein promotes SNARE complex assembly by acting as a chaperone for SNARE proteins (Yoo et al., 2023). Presynaptic terminals undergo neurotransmitter release thousands of times per minute, requiring the constant assembly and disassembly of the SNARE complex. However, this process generates highly reactive unfolded SNARE protein intermediates, making the synaptic terminal a highly vulnerable site for the initiation of neurodegenerative disorders (Kramer et al., 2007; Stafanis et al., 2010).

Lewy bodies predominantly consist of aggregated, insoluble forms of alpha-synuclein. *SNCA* mutations or multiplications can initiate a process wherein initially soluble alpha-synuclein monomers form oligomers. Subsequent stages involve the assembly of small protofibrils and the eventual emergence of large insoluble fibrils, constituting the core of Lewy bodies (Lashuel et al., 2013). The complete comprehension of the processes governing the structural conversion of individual alpha-synuclein protein from a monomeric state to oligomers, and ultimately into fibrils, remains a goal that has not yet been fully achieved.

## 2.8 Pathogenesis

There are several mechanisms have been implicated in pathogenesis of PD. Apart from aggregation of alpha-synuclein, number of other processes have been associated with PD which includes perturbed alpha-synuclein homeostasis contributed by dysregulated ubiquitin-proteasome system (UPS) and lysosomal autophagy system (LAS), endoplasmic reticulum

stress (ER), unfolded protein response (UPR), mitochondria dysfunction, oxidative stress, and neuroinflammation.

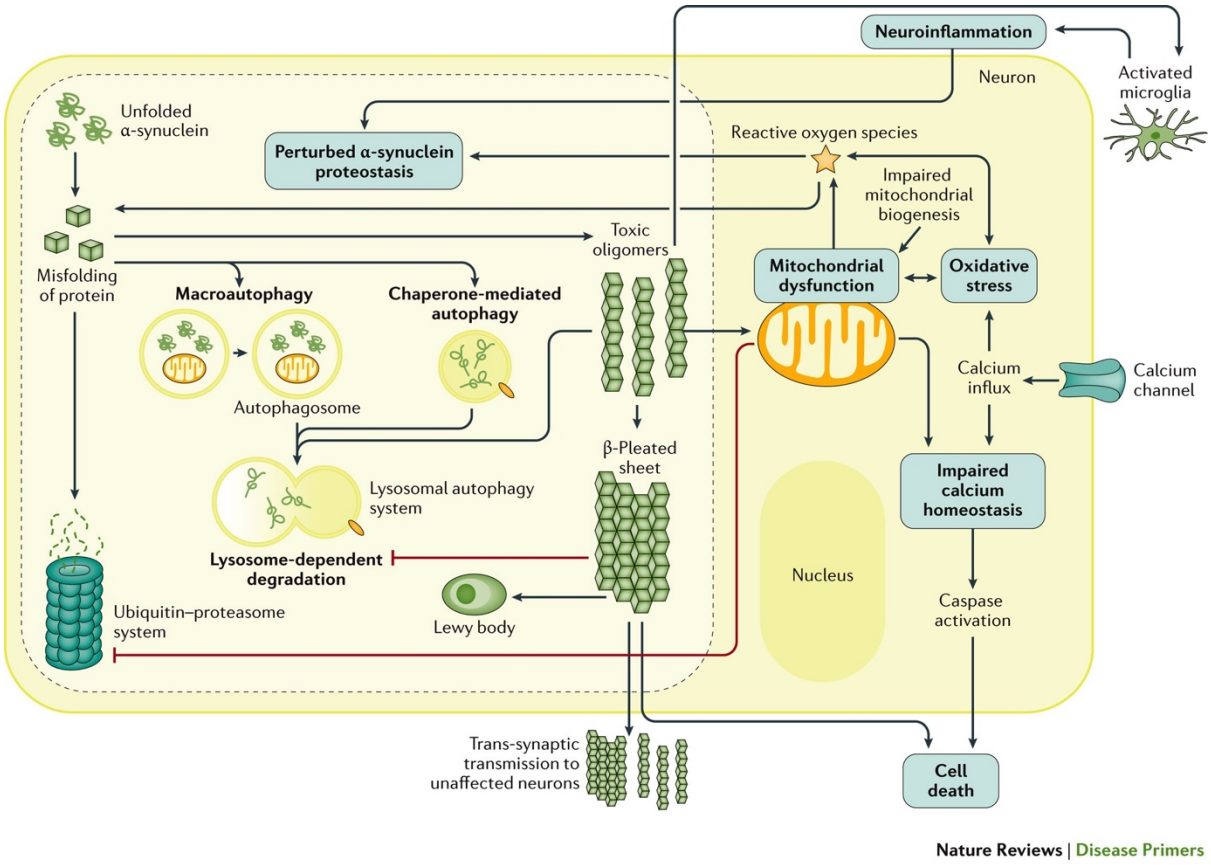



**Figure 1.8:** Schematic illustration of interplay between major cellular and molecular pathways associated with PD pathogenesis. Image reference: Poewe et al., 2017.

## 2.9 Endoplasmic Reticulum Stress and Unfolded Protein Response (UPR)

ER contributes to several physiological functions, such as protein synthesis, folding, and export; lipid synthesis and storage; glucose metabolism; and calcium homeostasis. It has been suggested that disruptions in the endoplasmic reticulum may play a significant role in the pathological impact of alpha-synuclein. Despite having sophisticated protein-folding machinery, protein misfolding still occurs in ER. The misfolded proteins are primarily cleared through either the endoplasmic reticulum-associated degradation pathway (ERAD) or the unfolded protein response (UPR). ERAD identifies misfolded proteins, returning them to the cytosol for subsequent ubiquitylation and degradation by the 26S proteasome, thereby preventing the accumulation of misfolded or unfolded proteins (Hetz et al., 2018).

Perturbations in ER functions, caused by factors like oxidative stress or dysregulation of calcium homeostasis, can surpass the ER's protein folding capacity, leading to a high load of accumulated misfolded or unfolded proteins within the ER. This triggers ER stress. In response, cells activate the UPR mechanism to address the abnormal accumulation of misfolded proteins. The primary goal of the UPR is to restore ER homeostasis. The UPR response has three main objectives: 1) reducing general protein translation, 2) enhancing the synthesis of protein-folding chaperones, and 3) increasing the translation of ERAD proteins. However, when these objectives are not achieved or in cases of chronic ER stress, the UPR can lead to cell apoptosis.

The UPR response involves three key transmembrane proteins or sensors located in the ER: protein kinase RNA-like ER kinase (PERK), inositol-requiring transmembrane kinase/endonuclease (IRE1) α/β, and activating transcription factor 6 (ATF6) α/β.

IRE1 possesses kinase and endoribonuclease activities that play a pivotal role in the UPR response. Upon activation, IRE1 dimerizes and undergoes auto-phosphorylation, activating its endoribonuclease function. This activity splices an unconventional mRNA called XBP1 (X-box binding protein 1). The translated product of spliced XBP1 is an active transcription factor that upregulates genes involved in ERAD, lipid synthesis, protein folding, and protein quality control (Costa et al., 2020; Martinez, 2019).

PERK is a type I transmembrane kinase that phosphorylates eIF2α (eukaryotic translation initiation factor 2α) at serine 51. This phosphorylation impedes protein translation into the ER (Walter et al., 2011), reducing the influx of nascent proteins and alleviating ER workload. However, eIF2α phosphorylation can lead to selective translation of specific genes, including ATF4 (activating transcription factor 4), which aids in restoring ER homeostasis by regulating the expression of chaperones, foldases, and autophagy components (Hetz et al., 2018). In prolonged ER stress, ATF4 induces the expression of proapoptotic elements like reactive oxygen species (ROS) production, upregulation of the CHOP transcription factor, and the BCL-2 family (Tabas and Ron, 2011; Pihan et al., 2017).

ATF6, a transcription factor normally situated in the ER membrane, undergoes translocation to the Golgi apparatus during ER stress. Here, it is cleaved by endopeptidases SP1 and SP2, releasing cytosolic fragments (ATF6f) containing a transcription factor domain.

ATF6f activates XBP1 and genes involved in ERAD, ER protein folding and quality control, and lipid synthesis (Bailey et al., 2007).

In 2007, Hoozemans and colleagues provided evidence of the modulation of unfolded protein response (UPR) mediators in the human brain affected by PD. They observed elevated levels of phosphorylated PERK and phosphorylated EIF2α proteins within the substantia pars compacta of PD brain samples, comparing them to individuals of similar age without the condition (Hoozemans et al., 2007). Other studies have also reported changes in the expression of BiP/GRP78, ATF4, and CHOP proteins in post-mortem human PD brain tissues from the substantia nigra pars compacta (Esteves et al., 2020; Selvaraj et al., 2012).

Both animal models and neuronal models derived from induced pluripotent stem cells (iPSCs) have further supported the association between PD and UPR response activation. Upregulation of phosphorylated IRE1 has been reported in iPSC-derived neurons from PD patients with *SNCA* triplication (Heman-Ackah et al., 2017). Additionally, iPSC-derived dopaminergic neurons carrying *GBA* mutations exhibited increased endoplasmic reticulum (ER) stress markers and impaired ER-associated degradation (Kovaleva et al., 2021).

The induction of ER stress and UPR response has been effectively replicated in both *in vivo* and *in vitro* models using pharmacological agents such as rotenone, MPTP, 6-hydroxydopamine (6-OHDA), and tunicamycin (Costa et al., 2020). Despite this growing body of evidence, the precise cause-and-effect relationship between ER stress/UPR and PD remains to be fully established.

## 2.10 Mitochondria Dysfunction and Oxidative Stress

Dopaminergic neurons are known for their high energy demands, and mitochondria play a crucial role in meeting this energy requirement through ATP production (Zainul Ali 2022). The majority of adenosine triphosphate (ATP) within mitochondria is generated via the oxidative phosphorylation pathway (OXPHOS), which is coupled to the electron transport chain (ETC). OXPHOS consists of five protein complexes: Complex I (NADH ubiquinone oxidoreductase), Complex II (Succinate dehydrogenase), Complex III (Ubiquinol: Cytochrome c reductase), Complex IV (Cytochrome c oxidase), and Complex V (ATP synthase) (Ahmed et al., 2020). Complex I establishes the mitochondrial membrane potential through proton

translocation. Subsequently, ATP is produced in Complex V with the assistance of Complex II, III, and IV. Complex V differs functionally from the other complexes as it is involved in ATP generation, while the others are responsible for proton transfer into the intermembrane space (Ojas et al., 2023). OXPHOS relies on both nuclear DNA and mitochondrial DNA-encoded proteins (Exner et al., 2012). Studies have demonstrated an acute shortage of energy (ATP) in the substantia nigra pars compacta (SNpc) region of PD patients, indicating the involvement of mitochondria in PD pathogenesis (Wellstead et al., 2011).

The first link between mitochondria and PD became apparent with the discovery that exposure to a selective mitochondrion Complex I inhibitor, MPTP, leads to the cardinal features of PD in humans (Langston et al., 1983). Inhibition of Complex I reduces ATP production and increases the generation of reactive oxygen species (ROS), which can damage mitochondrial DNA and components of the respiratory chain. Consequently, the detrimental effects of Complex I inhibition result in mitochondrial impairment and oxidative stress (Ramsay et al 1986).

Apart from MPTP, other environmental toxins such as trichloroethylene, rotenone, and paraquat also induce selective degeneration of substantia nigra dopaminergic neurons in animal models, implicating mitochondrial dysfunction in PD pathogenesis as these toxins inhibit mitochondrial Complex I of the respiratory chain (Jiang et al 2018).

Both nuclear and mitochondrial genomes have been implicated in the pathogenesis of PD. Several mitochondrial DNA (mtDNA) deletions/rearrangements have been observed in the substantia nigra of PD patients compared to age-matched controls (Vastegani et al., 2023; Gu et al., 2002). Numerous nuclear-encoded factors are essential for mtDNA replication, transcription, and translation. Modulation of these factors can result in mitochondrial dysfunction. For example, mutations in polymerase gamma (*POLG*), a gene involved in mtDNA synthesis, are associated with PD pathogenesis due to mitochondrial dysfunction (Luoma et al., 2004).

So far, several genes including *PRKN, PINK1, LRRK2, SNCA,* and *ATP13A2* have been identified as monogenic causes of familial PD, with pathogenic mutations in these genes linked to mitochondrial dysfunction (Moon et al., 2015; Vastegani et al., 2023; Park et al., 2018). Mitochondrial dysfunction can result from oxidative stress, impaired mitochondrial biogenesis,

calcium imbalance, altered mitochondrial dynamics, dysfunctional electron transport chain, impaired mitochondrial trafficking, and defective mitophagy (Park et al., 2018).

The *PRKN* (PARK2) gene is considered the most frequent cause of early-onset autosomal recessive PD. The Parkin protein encoded by the *PRKN* gene is a multifunctional E3 ubiquitin ligase involved in mitochondrial quality control (QC), mitophagy, and mitochondrial homeostasis (Ge et al., 2022). Pathogenic mutations in PRKN genes lead to dysregulation of mitochondrial QC, resulting in increased ROS and subsequent cellular damage (Johnson et al., 2012). *PINK1* is the second most frequent cause of autosomal recessive early-onset PD and encodes a mitochondrial serine/threonine kinase. It is involved in mitochondrial QC along with Parkin. Mutations in the *PINK1* gene can disrupt mitochondrial complex I and mitochondrial morphology, leading to the accumulation of damaged mitochondria and dysregulation of mitochondrial functions (Morais et al., 2009). The *PARK7* gene encodes the DJ-1 protein and is also a cause of autosomal recessive early-onset PD. DJ-1 has been shown to be neuroprotective by preventing oxidative stress-induced mitochondrial dysfunction. Pathogenic mutation in PARK7 has been associated with oxidative stress cell death cascade in PD (Pantcheva et al., 2014). Overexpression of *SNCA* has also been associated with impairment of mitochondrial function through the accumulation of alpha-synuclein aggregates in mitochondria (Parihar et al., 2009). A study has demonstrated that alpha-synuclein can influence the master regulator of mitochondrial biogenesis, peroxisome proliferator-activated receptor gamma coactivator 1-alpha (*PGC-1α*) (Ryan et al., 2013). The molecular mechanism associated with the aetiology of mitochondrial dysfunction and subsequent oxidative stress in PD pathogenesis remain unclear.

## 2.11 Prion-Like Spreading of Alpha-synuclein

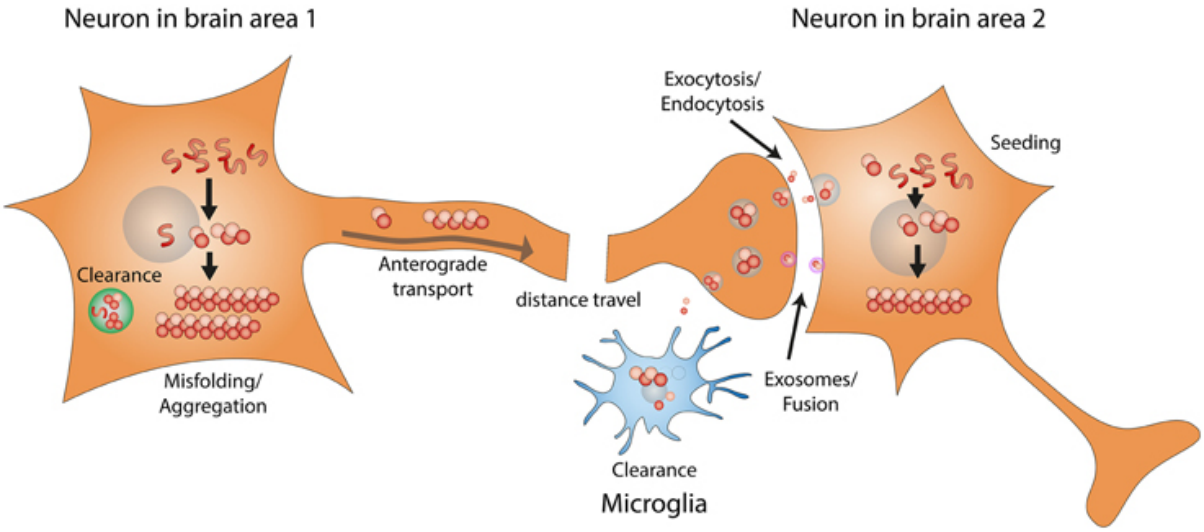



**Figure 1.9:** Prion like spreading of alpha-synuclein aggregates. Image reference: Zhang et al., 2015.

The prion-like spreading hypothesis for alpha-synuclein suggest that once pathogenic alpha-synuclein aggregates developed within neurons, the aggregates can travel intra-axonally and be released via exosomes to other brain regions (Zhang et al., 2015). Uptake of these pathogenic aggregates by neighbouring neurons triggers endocytosis and further aggregation of endogenous alpha-synuclein, potentially spreading pathology throughout the brain over time. Recent findings support the notion that initial alpha-synuclein aggregation occurs in the olfactory bulb and peripheral nervous system, later extending to the brainstem and midbrain as suggested by the Braak hypothesis (Braak et al., 2003).

The notion of prion-like transmission of alpha-synuclein was initially suggested through the observation of alpha-synuclein aggregation in transplanted healthy fetal mesencephalic dopaminergic neurons, occurring years after the transplantation itself. Subsequently, a noteworthy percentage, up to 5% and 30%, of grafted healthy dopamine neurons displayed Lewy body pathology at the 10-year and 16-year marks post-transplantation, respectively. This observation hinted at the spread of Lewy pathology from the host to the graft. Such transmission might be attributed to the direct transfer of aggregated alpha-synuclein seeds to healthy neurons or influenced by the unique microenvironment within the PD brain (Kordower et al., 2008; Li et al., 2008). Since these initial findings, several *in vitro* and *in vivo* studies have reported the cell-to-cell transfer and prion-like propagation of alpha-synuclein (Guan et al., 2020; Gustafsson et al., 2018; Domert et al., 2016)).

Employing purified alpha-synuclein seeds, both *in vitro* cellular models and animal models have effectively shown that externally assembled alpha-synuclein fibrils (referred to as preformed fibrils) can prompt the aggregation of endogenous soluble alpha-synuclein upon uptake by healthy cells (Vargas et al., 2019). Recently, a study highlighted the influence of pathological mutations in alpha-synuclein on the transmission of pathological effects (Guan et al., 2020). Specifically, this research unveiled that mutations like H50Q and A53T in alpha-synuclein lead to a significant increase in the secretion of alpha-synuclein from cells. Additionally, A53T, G51D, and H50Q PFF fibrils exhibited a seeding effect in the rat substantia nigra pars compacta. These discoveries present promising avenues for PD therapeutic research. However, a crucial prerequisite involves unravelling the mechanisms that underlie the prion-like spreading of alpha-synuclein before harnessing them for targeted therapeutic interventions.

## 2.11 Non-Cell-Autonomous Neurodegeneration in PD

The majority of PD research focuses on "cell-autonomous" mechanisms, overshadowing the investigation of "non-cell-autonomous" effects. While the involvement of non-cell-autonomous effects on dopaminergic neurons in PD is recognized, the precise mechanisms remain elusive. In PD, the loss of dopaminergic neurons from SNpc and alpha-synuclein accumulation is usually accompanied by a glial cell reactivity and neuroinflammation. The role of glial reactivity and neuroinflammation is unclear whether it is protective or detrimental for dopaminergic neurons (Pons-Espinal et al., 2020). It is possible that astrocytes and microglia are activated in early stages of disease to protect SNpc dopaminergic neurons, however, as the disease progresses activated glia becomes toxic. The involvement of microglia in PD was first highlighted by a study published in 1988, which reported enrichment of human major histocompatibility class II(MHC2) positive amoeboid shaped microglia in PD brains tissues compared to the aged-matched Controls (McGeer et al., 1988). MHC2 protein is involved in antigen presentation and inflammatory signalling, whereas microglia amoeboid morphology suggests reactive microglia. This finding was later confirmed by other studies which showed growing number of MHC2 positive microglia over time in substantia nigra of PD compared to healthy controls (Imamura et al., 2003).

PD associated mutations and or aging may contribute to altered reactive glia cell response (Kam et al., 2020). Animal study have shown that selective expression of A53T alpha-

synuclein in astrocytes causes neurodegeneration, indicating non-cell autonomous effect on neurons (Gu et al., 2010). *In vitro* co-culture studies have also indicated that PD-associated mutant astrocytes can contribute to non-cell-autonomous neurotoxicity in neighbouring healthy dopaminergic neurons (Chai et al., 2019; di Domenico et al., 2019). Notably, abnormal alpha-synuclein levels in astrocytes trigger the release of proinflammatory chemokines and cytokines, inducing microglia activation. Activated microglia have been shown to induce neurotoxic astrocytes (Liddelow et al., 2017). A precise role of activated microglia in PD's pathogenesis remains unclear. Understanding and targeting non-cell autonomous mechanisms are crucial for developing effective therapies for PD.

## 2.12 Midbrain Dopaminergic Neuron Development *In Vivo*

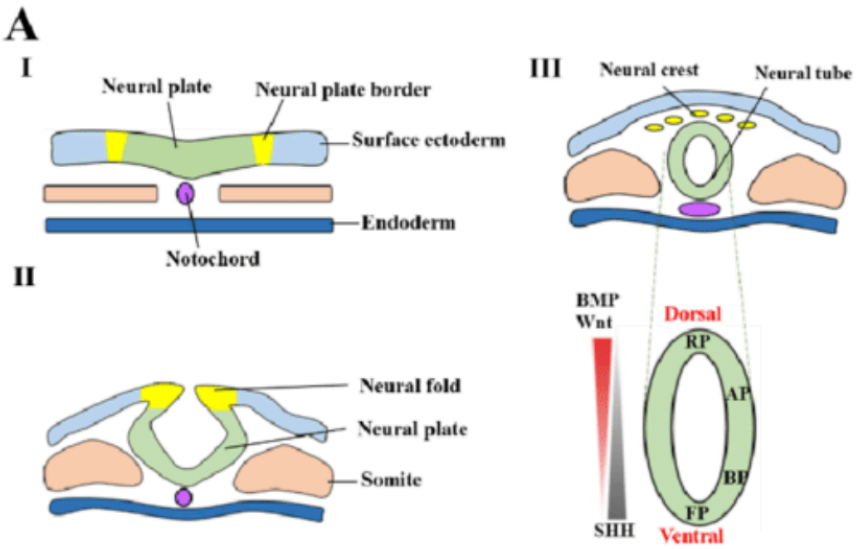



**Figure 1.10:** Schematic of neural tube development and signaling governing neural tube patterning. Image reference: Wang et al., 2022.

The development of midbrain dopaminergic neurons takes place in a complex milieu of genetic and environmental cues. Midbrain development itself is a complex process that rely on the interplay between these cues. The entire nervous system is developed through neurulation, a process where the neural tube and neural crest emerge from the neural plate (Chhetri et al., 2023). This neural plate arises from neuroepithelial cells of the neuroectodermal layer via induction of notochord, which then fold and fuse to form the neural tube.

The neural tube is comprised of two main domains: the ventral midline floor plate and the dorsal midline roof plate, adjacent to the crest cells that form the peripheral nervous system.

Additionally, the neural tube contains lateral alar and basal plates. The roof plate secretes bone morphogenic protein (BMP), countering the effect of sonic hedgehog protein (SHH) secreted by the notochord and floor plate. This crucial interplay between SHH and BMP creates morphogen gradients, which in turn define distinct neuronal cell types along the dorsoventral axis of the neural tube (Wang et al., 2020). Progenitor cells in the basal plate develop into interneurons and lower motor neurons, while those in the alar plate give rise to sensory neurons (Ishikawa et al., 2022).

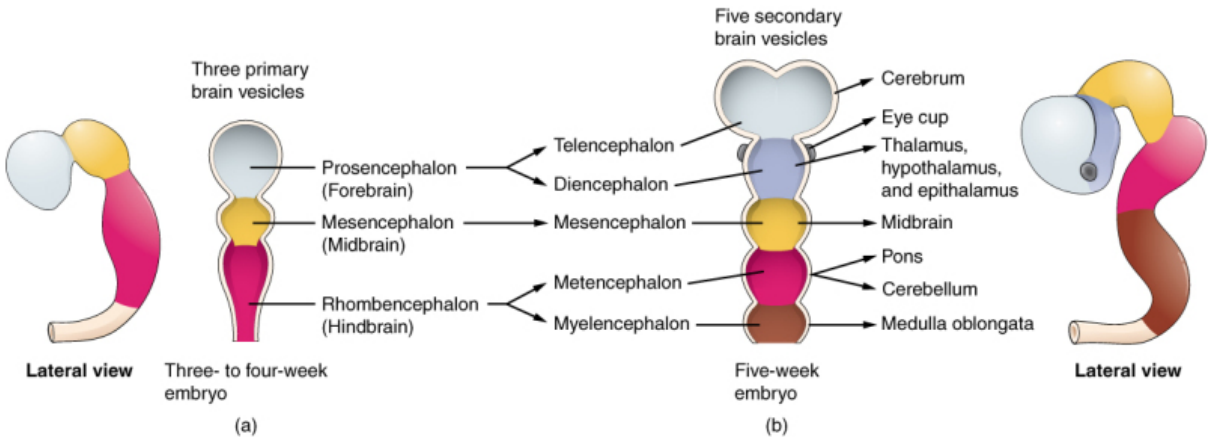



**Figure 1.11:** Primary and secondary vesicles development from Neural tube. Image reference: Chhetri et al., 2023.

In the early stages of brain development, the neural tube differentiates into three primary vesicles: the forebrain or prosencephalon, the midbrain/mesencephalon, and the hindbrain/rhombencephalon. Subsequently, the prosencephalon gives rise to the diencephalon and telencephalon, while the mesencephalon remains undivided. Within the rhombencephalon, the myelencephalon and metencephalon emerge (Chhetri et al., 2023).

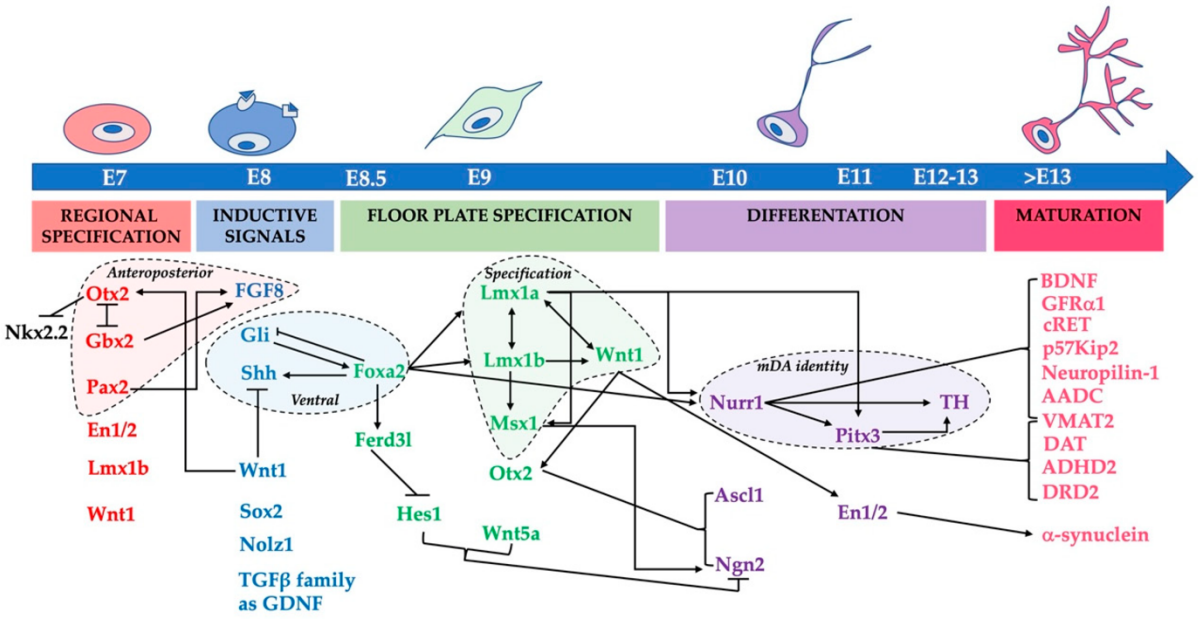



**Figure 1.13:** Molecular players involved in the development of mDA neurons in the mouse brain. Image reference: Volpicelli et al., 2020.

The organization of the ventral midbrain rely on two critical signaling centers: the midbrain floor plate (mFP) and the midbrain-hindbrain boundary, known as the isthmic organizer (Asgrimsdottir et al., 2020). The mFP secretes SHH, inducing the expression of *FOXA2*, a well-established marker of the floor plate (Hegarty et al., 2013). The development of midbrain dopaminergic neurons (mDA) is governed by the *SHH-FOXA2* signaling pathway. The isthmic organizer develops at the boundary between the midbrain and hindbrain, crucial for the development of both regions (Mesman et al., 2020). The accurate positioning of the isthmic organizer depends on the coordinated expression and mutual repression of two vital transcription factors: *OTX2* and *GBX2*. GBX2 is expressed in the hindbrain, whereas *OTX2* is confined to the midbrain. The regional expression of *OTX2* and *GBX2* establishes the midbrain-hindbrain boundary (Arenas 2015). Several genes, including *PAX2, LMX1B, EN1/2*, and *Wnt1*, play pivotal roles in determining the correct positioning of the isthmic organizer and the secretion of FGF8(Volpicelli et al., 2020). The low level of FGF8 in the midbrain induces midbrain identity, while elevated FGF8 in the hindbrain results in hindbrain identity. Additionally, Wnt1, secreted by the isthmic organizer, regulates the expression of midbrain transcription factors such as *EN1/2, OTX2, NURR1,* and *LMX1A* (Wang et al., 2020).

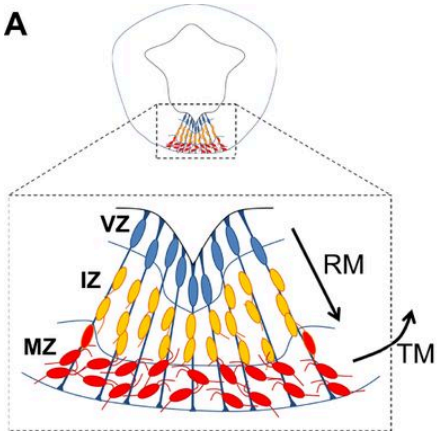



**Figure 1.12:** Schematic representation of a section through the mFP at E11.5. VZ = Ventricular zone, which contains radial glial cells (Blue); IZ = intermediate zone, which contains post-mitotic neuroblasts (Yellow); MZ = Marginal zone, which contains fully differentiated mDA neurons (Red). RM = radial migration of radial glial cell neuroblasts; TM = tangential migration of mDA neurons towards the substantia nigra pars compacta (SNpc). Image reference: Arenas et al., 2015.

As the ventral neural tube progresses, a cascade of transcription factors is sequentially activated, crucially initiating the differentiation of midbrain floor plate (mFP) cells into mature midbrain dopamine neurons (mDA). This process is essential for establishing the molecular features of immature mDA neural progenitors and the unique identity of fully developed mDA neurons. After mFP specification, mDA progenitors in the ventricular zone (VZ) begin to express two crucial transcription factors necessary for mDA development: *FOXA2* and *LMX1A* (Andersson et al., 2006; Ferri et al., 2007). These progenitor cells proliferate and subsequently undergo neurogenesis, prompted by the bHLH proneural gene Neurog2 (*NGN2*) (Kele et al; 2006). This neurogenesis leads to the formation of post-mitotic mDA neuroblasts expressing *NURR1*. The expression of *NGN2* is induced by *MSX1* (Msh homeobox 1), which, in turn, is activated by *LMX1A* expression (Volpicelli et al., 2020). While *MSX1* is primarily expressed in mDA neuroblasts, LMX1A expression persists in adult post-mitotic mDA neurons. The *OTX2* indirectly regulates *NGN2* expression by activating *LMX1A* through Wnt1 (Volpicelli et al., 2020). As neurogenesis is initiated, the neuroblasts migrate to their destination in the mantle zone. During this migration, the neuroblasts acquire *PITX3* and *PBX1* expression, necessary for the further development of mDA neurons, as well as markers of mDA identity, including TH, dopamine transporter (*SLC6A3*/DAT), and vesicular monoamine transporter (*SLC18A2*/VMAT2) (Asgrimsdottir et al., 2020). *NURR1* has been found to regulate the expression of VMAT2 and DAT, as well as neurotrophic factors BDNF (brain-derived

neurotrophic factor) and the response to GDNF (glial-derived neurotrophic factor) via Ret receptor, both critical for the maturation and survival of mDA neurons (Hyman et al., 1991; Galleguillos et al., 2010; Volpicelli 2007).

Midbrain dopaminergic neurons can be categorized into three primary groups: 1) the retrorubral field (RRF), also known as A8, 2) the substantia nigra pars compacta (SN), denoted as A9, and 3) the ventral tegmental area (VTA), referred to as A10. SNpc A9 neurons project to the dorsal striatum via the dorsal striatum pathway, essential for voluntary movement control. On the other hand, VTA A10 and RRF A8 neurons project to the ventral striatum and prefrontal cortex through the mesolimbic and mesocortical pathways (Arenas et al., 2015). Notably, VTA and SN mDA neural progenitor cells express distinct markers. The transcription factor *SOX6* (SRY-box 6) is predominantly expressed in the SNpc region, while *OTX2* and Nolz1 are primarily observed in VTA precursor cells (Panman et al., 2014).

Numerous studies highlight the significance of these transcription factors in mDA neuron development and survival. *FOXA2* is expressed in all ventral DA progenitors and post-mitotic mDA neurons, regulating mDA specification and neurogenesis (Wang et al., 2020). In *FOXA1/2* knockout mice, there is reduced expression of *TH* and *AADC*, leading to diminished dopamine release into the striatum (Pristera et al., 2015). The absence of SNpc mDA neurons and subsequent impairment of motor functions have been observed in heterozygous *FOXA1/2* models (Arenas et al., 2008). Knockout of *LMX1A/B* results in substantial mDA neuron loss in animal models (Yan et al 2011). *FOXA1/2* knockout mutants showing a lack of *LMX1A/B* expression. Conversely, gain-of-function experiments have demonstrated that *FOXA1/2* induces *LMX1A/B* expression (Lin et al., 2009). NURR1-deficient mice lack TH expression entirely, highlighting the importance of *NURR1* in inducing *TH* expression in mDA neurons (Zetterstrom et al., 1997; saucedo-cardenas et al., 1998). *PITX3* knockout mice display defects in SNpc mDA neuronal development. Early *PITX3* expression is crucial for specifying SNpc mDA neurons; otherwise, mDA neurons adopt a VTA fate (Maxwell et al., 2005). Wnt1-null mice exhibit malformation of the mDA neurons and midbrain. Although VM DA neural progenitors do develop in Wnt1-null mice, they fail to proliferate and differentiate effectively (Prakash et al., 2006; McMahon et al., 1990). Wnt1 critical role in *EN1/2* expression has been demonstrated, as *EN1/2* knockout mice exhibit a phenotype similar to Wnt1-null mice, suggesting that the loss of Wnt1 might lead to a loss of EN1/2 expression (Hagarty et al., 2013).

These evidence outline a complex interplay of various transcription factors and signaling molecules in the development of ventral dopamine neurons.

## 2.13 Dopamine Metabolism

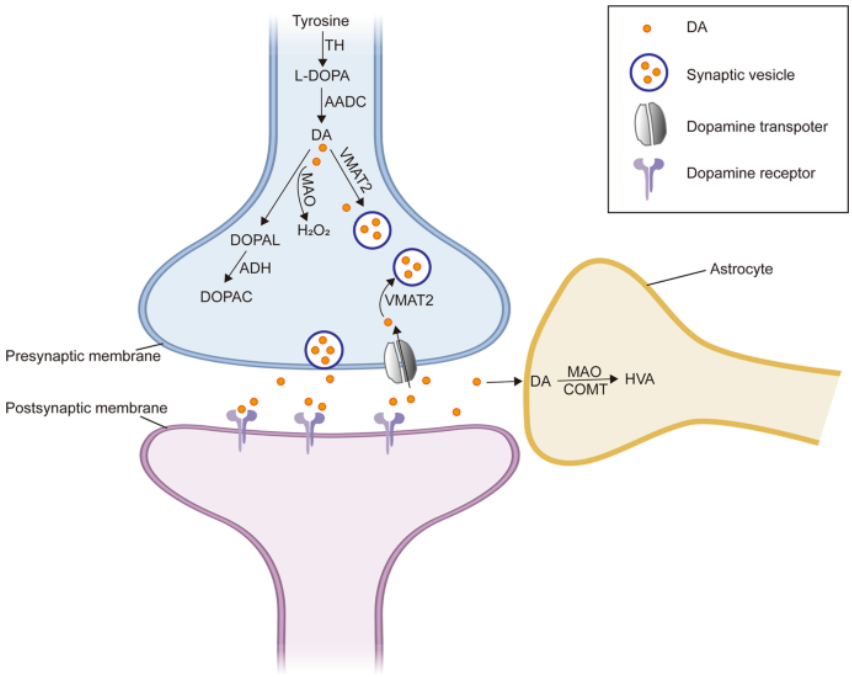



**Figure 1: Dopamine metabolism in dopaminergic neurons. Image reference: Xu et al., 2022.**

Dopamine is a neurotransmitter that plays a crucial role in regulating reward and movement processes (Olguin et al., 2016). Dopaminergic neurons are responsible for its synthesis, starting with the conversion of tyrosine to L-3,4-dihydroxyphenylalanine (L-DOPA) by tyrosine hydroxylase (TH). Subsequently, L-DOPA is further transformed into cytosolic dopamine by the enzyme aromatic amino acid decarboxylase (AADC) (Best et al., 2009). The resulting dopamine is then packaged into synaptic vesicles by the vesicular monoamine transporter 2 (VMAT2) and stored within these vesicles until it is released (Xu et al., 2022). Upon an action potential in the dopaminergic neuron, dopamine-filled synaptic vesicles are released into the synaptic cleft (Xu et al., 2022). Here, dopamine binds to dopamine receptors on both presynaptic and postsynaptic neurons (Speranza et al., 2021). Following synaptic transmission, unbound dopamine can be reuptaken into the presynaptic neurons' cytosol through the dopamine transporter (DAT) and subsequently repackaged into synaptic vesicles (Egana et al., 2009). Any excess dopamine left in the synaptic cleft is metabolized by glial cells into various metabolites, including homovanillic acid (HVA), through the action of enzymes

such as monoamine oxidase (MAO) and catechol-o-methyltransferase (COMT) (Xu et al., 2022). Additionally, cytosolic dopamine within neurons can also undergo breakdown by MAO, yielding inactive metabolites such as 3,4-dihydroxyphenylacetaldehyde (DOPAL), 3,4-dihydroxyphenylacetic acid (DOPAC), or hydrogen peroxide ($H_2O_2$).

## 2.14 Models of PD

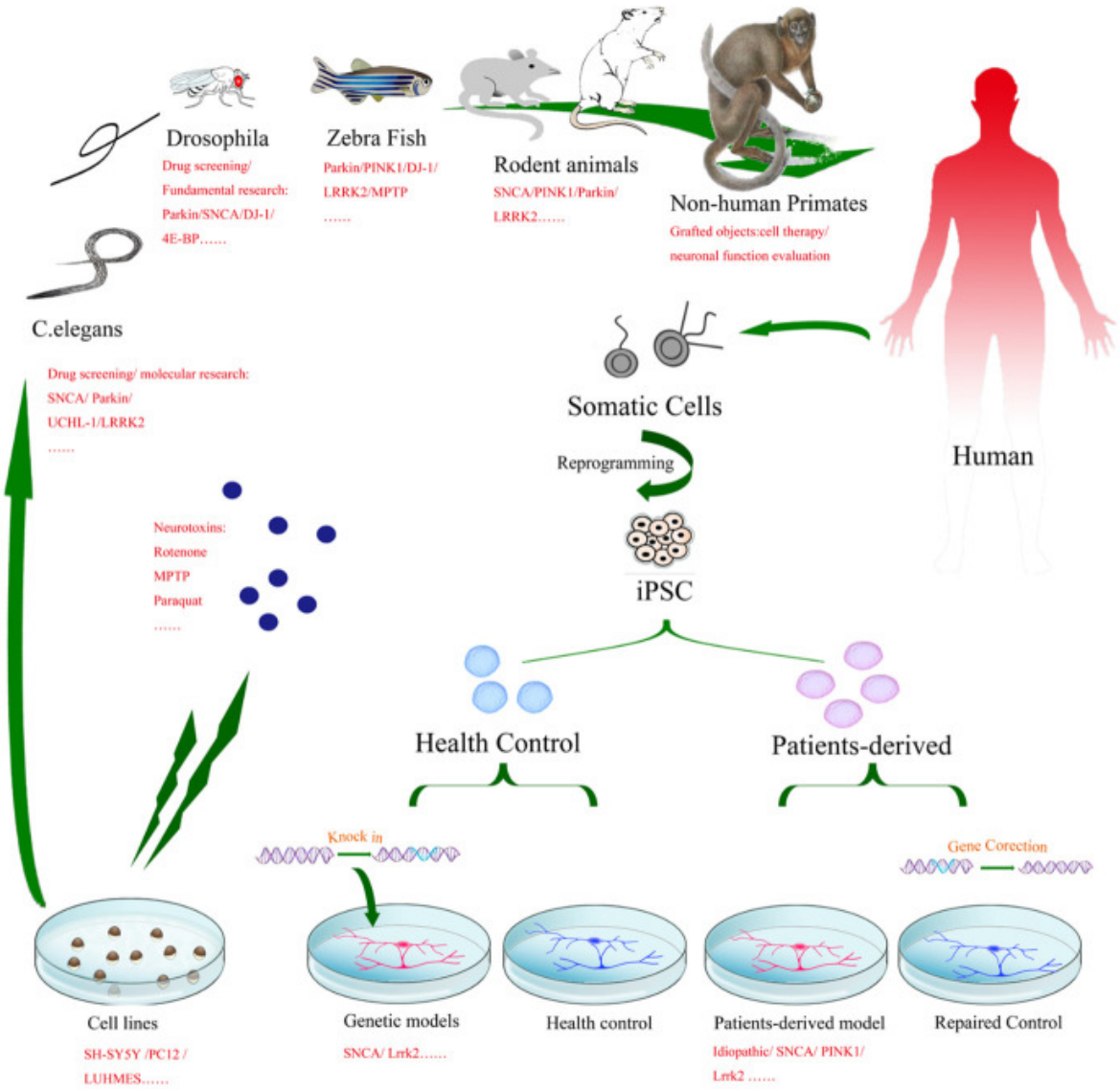



**Figure 1.15:** Schematic illustration depicting a variety of model being used to study PD. Image reference: Ke et al., 2021.

The cellular and molecular mechanisms underlying key pathological features of PD, such as dopaminergic neuronal degeneration and Lewy body formation, have yet to be fully elucidated. This lack of knowledge is primarily due to the inaccessibility of human brain tissue from different stages of PD. While postmortem tissues have been employed to study PD, they

typically represent the end stage of the disease, making them unable to reveal important mechanisms involved in the early stages and disease progression. To bridge this gap, several animal and cellular models have been developed, significantly advancing our understanding of the disease. However, none of these models fully replicate the disease, primarily for two reasons: 1) The etiopathogenesis of sporadic PD remains elusive, making it challenging to reproduce something we do not fully understand. 2) Most models aim to induce pathology rapidly, in contrast to the gradual development and progression of PD in humans.

### 2.14.1 Animal Models

Animal models are instrumental for studying the progression of PD, as they offer the capacity for behavioral experiments. These models include both mammalian (rodents and non-human primates) and non-mammalian models (C. elegans, Drosophila, zebrafish) (Konnova et al., 2021). In animal experiments, researchers mainly employ two approaches: exposure to neurotoxins and the introduction of pathogenic mutations (Konnova et al., 2021). While these animal models successfully replicate cardinal PD features such as loss of dopaminergic neurons in SNpc and Lewy body formation, they have significant limitations. For instance, neurotoxin-based models induce rapid dopaminergic neuronal loss without forming Lewy bodies and involve a faster disease progression, unlike the human PD development. Furthermore, animal models carrying pathogenic mutations in genes such as *DJ-1, LRRK2, SNCA, and ATP13A2* do not reliably produce SNpc degeneration and associated motor symptoms (Blesa et al., 2014).

### 2.14.2 *In Vitro* Models

Traditionally, a two-dimensional (2D) *in vitro* culture system has been used to study PD. These systems have been highly valuable and significantly improved our understanding of the disease. In comparison to animal models, *in vitro* models induce PD pathology more quickly, are easy and reliable for genetic and pharmacological manipulations, cost-effective, and do not involve ethical constraints (Ferrari et al., 2020; Ke et al., 2021). Although *in vitro* cell culture systems do not offer behavioral assessment and pharmacokinetic features, they provide a platform to investigate complex pathological mechanisms and drug discovery in a controlled environment. Several cellular models have been developed over the years for studying PD, including established cell lines, primary cultures, and pluripotent stem cell-derived neuronal cultures.

#### 2.14.2.1 Immortalized Cell Lines

Established cell lines are immortalized cells being used in PD research. Examples include SH-SY5Y, PC12, LUHMES (Lund Human Mesencephalic cells), MEs23.5, MN9D, N27 line, and Neuro 2a (Ferrari et al., 2020; Ke et al., 2021). These cell lines offer several advantages, such as a high proliferation rate, reproducibility, ease of genetic manipulation, and some exhibiting dopamine metabolism and synaptogenesis (Lopes et al., 2017). However, they also have disadvantages, such as their oncogenic origin and characteristics, and the laborious differentiation steps needed to convert them into dopaminergic-like cells (Lopes et al., 2017; Herrup and Yang., 2007). For example, SH-SY5Y cell lines, a subclone of the SK-N-SH neuroblastoma cells, can be efficiently differentiated into dopaminergic neurons expressing TH, DAT, D2, and D3 receptors using specific protocols (Biedler et al., 1978; Xicoy et al., 2007). These cells are a useful tool for studying cellular mechanisms such as mitochondrial dysfunction, lysosomal pathways, autophagy, and the neuroprotection of dopamine agonists against dopamine and neurotoxin-induced toxicity. The LUHMES cell line is a subclone of MESC 2.10 derived from the human mesencephalic brain region of an 8-week-old fetus and immortalized by inserting a tetracycline-controlled v-myc gene (TET-off) (Lotharius et al., 2005). Like SH-SY5Y, LUHMES cells exhibit immortalized characteristics, can be permanently subcultured and maintained, and can be relatively easily differentiated into dopaminergic neurons that resemble primary neurons in morphology. They express key dopaminergic markers such as TH, DAT, VMAT, and D2 receptors and release dopamine upon treatment with DOX, BDNF, and GDNF. These cells have been extensively used to investigate neurotoxin-induced dopaminergic cell death and alpha-synuclein aggregation through the introduction of pathogenic mutations in the *SNCA* gene (Ferrari et al., 2020; Ke et al., 2021).

#### 2.14.2.2 Primary Dopaminergic Cultures

Primary dopaminergic neuronal cultures are derived from postnatal substantia nigra brain tissue of rats and mice (Lopes et al., 2017). Studies have shown that dissociated SNpc primary cultures consist of a high percentage of dopaminergic neurons, and the physiological properties of these neurons resemble *in vivo* dopaminergic neurons. These primary neurons also display the same receptors and transmitters as those *in vivo* (Mauko et al., 1992). These neuronal cultures are sensitive to neurotoxins such as MPTP, rotenone, paraquat, and 6-OHDA (Blum et al., 2001; Radad et al., 2006). Interestingly, dopaminergic neurons in primary neuronal cultures derived from transgenic mice lacking *NDUFS4* (mitochondrial complex I

essential subunit) do not show compromised survival despite lacking dysfunctional complex I (Choi et al., 2008). This result suggests that other intrinsic factors could play a role in neurotoxin-mediated toxicity, such as rotenone, MPTP, and paraquat. Despite these advantages, primary cultures are heterogeneous, lack reproducibility, contain a high percentage of glial cells, and originate from a non-human genetic background (Ferrari et al., 2020).

#### 2.14.2.3 Pluripotent Stem Cell-Derived Dopaminergic Neuronal Cultures

Pluripotent stem cells, including human embryonic stem cells (hESCs) and induced pluripotent stem cells (hiPSCs), are now being used to develop PD models. The discovery of hiPSCs has significantly accelerated neuroscience research, as they can be generated by reprogramming fully differentiated cells from healthy individuals and PD patients (Ke et al., 2021). hiPSCs offer several advantages, including species-specific cellular pathways compared to animal models and no ethical issues associated with ESCs (Ferrari et al., 2020). hiPSCs provide an opportunity to generate dopaminergic neurons with specific genetic backgrounds, allowing for the study of the role of PD-associated mutations. These neurons can also be used for drug screening and to study the effects of environmental risk factors such as neurotoxins on PD pathogenesis. Furthermore, hiPSCs generated from PD patients can be useful in investigating the underlying cellular and molecular mechanisms of disease onset and progression. The combination of iPSC technology and genome editing, such as CRISPR-Cas9, has proven invaluable for dissecting PD pathogenic mechanisms. For example, Byers et al. (2011) generated *SNCA* triplication PD patient-derived iPSCs and isogenic controls, revealing alpha-synuclein accumulation and oxidative stress in PD neurons. Similarly, Nguyen et al. (2011) generated iPSCs harboring *LRRK2* G2019S mutations and isogenic controls, which exhibited alpha-synuclein accumulation, increased sensitivity to neurotoxins, and oxidative stress upon differentiation. A recent study identified histone deacetylase 4 (HDAC4) as a potential therapeutic target using dopaminergic neurons derived from PD-GBA patients through high-resolution single-cell transcriptomics (Lang et al., 2019). However, iPSC models have their limitations, including the maturity of differentiated iPSCs compared to fetal counterparts and an altered epigenetic profile as observed in parental cells (Kim et al., 2011).

### 2.14.2.4 Monolayer Dopaminergic Neuronal Differentiation

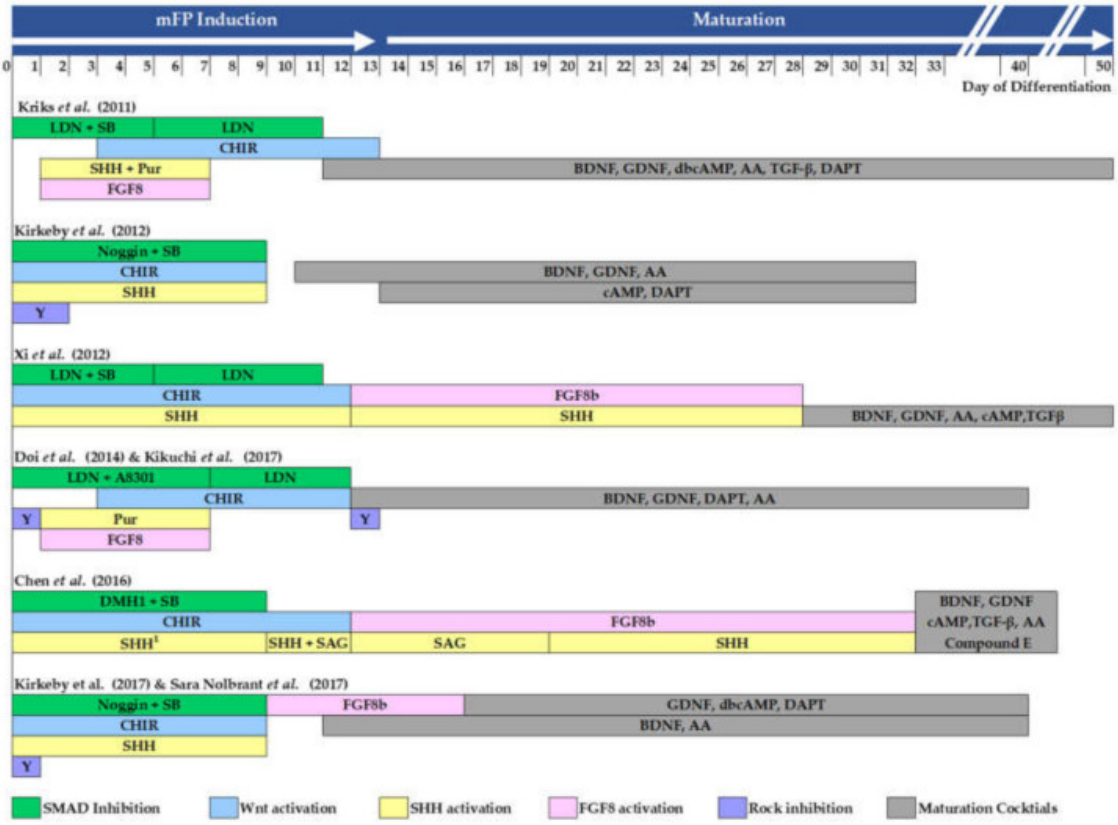



**Figure 1.16:** Schematic illustration strategy and timeline of midbrain DA neuron differentiation.

Dopaminergic neurons comprise less than 1% of neurons in the brain; however, they play crucial roles in movement, reward processing, stress regulation, mood control, and addiction (Fiorenzano et al., 2021). Various strategies have been employed to differentiate dopaminergic (DA) neurons, including: 1) Co-culture of embryonic stem cells (ESCs) with neural-inducing stromal feeder cells, such as the MS5 or PA6 method, 2) Generation of embryoid bodies (EB), 3) Direct fate conversion through the ectopic expression of transcription factors (TFs), 4) Directed differentiation of pluripotent stem cells into DA neurons using small molecules, growth factors, and morphogens (Wang et al., 2020). Several protocols have been developed to differentiate pluripotent stem cells into DA neurons that express genes similar to *in vivo* dopaminergic cells (Asgrimdottir et al., 2020). The differentiation of PSC-derived DA neurons involves three main steps: 1) Neural induction, 2) Midbrain floor plate (mFP) specification, 3) Neural maturation (Kreutz et al., 2023). Various small molecules, growth factors, and morphogens have been used to induce these changes in the cultures. Neural induction is achieved by inhibiting dual SMAD signaling using combinations of SB 431542,

A8301, Noggin, LDN-193189, and dorsomorphin. Inhibiting dual SMAD prevents the development of endodermal and mesodermal lineages, resulting in neural induction (Chang et al., 2007). The SMAD signaling pathway involves BMP and nodal signaling, with SB 431542 and A8301 used to inhibit nodal signaling and Noggin, LDN-193189, and dorsomorphin for inhibiting BMP signaling (Wang et al., 2020). mFP patterning is achieved by activating SHH, Wnt, and FGF8 signaling. SHH signaling can be activated using SHH protein or small molecules like purmorphamine and SAG, which induce *FOXA2* expression. Wnt signaling is activated using the potent GSK3β inhibitor CHIR99021 (Law et al., 2022). CHIR99021 is applied either at a high concentration (3uM) post neural induction or at a low concentration (0.7-0.8uM) from day 0 to induce midbrain fate (Krikeby et al., 2012). FGF8 signaling has been shown to promote midbrain DA neural progenitor differentiation (Xi et al., 2012). In DA differentiation protocols, either of the two FGF8a isoforms is used to activate FGF8 signaling. Post-mFP specification, the correctly patterned cells express *LMX1A* and *FOXA2*, which are unique to the ventral midbrain (Struder et al., 2012). For neural maturation and maintenance, various neurotrophic factors such as BDNF, GDNF, ascorbic acid (AA), dbcAMP, DAPT, and TGF-beta are commonly used in most DA differentiation protocols. The protocol proposed by Kriks et al. in 2011 is the most cited and widely used method to generate DA neurons from PSCs. Nevertheless, extensive research is still being conducted to optimize the protocols with the aims of: 1) Improving the yield and efficiency of DA neuron production, 2) Reducing undesired heterogeneity within the culture 3) Minimizing variability between batches 4) Decreasing the time needed to produce DA neurons, 5) Achieving Good Manufacturing Practice (GMP) compliant conditions, 6) Optimizing for transplantation, 7) Adapting to various starting stem cell types (Marton et al., 2019).

#### 2.14.2.5 Dopaminergic Neurons Differentiation Through Ectopic Expression of Transcription Factors (Tfs)

Monolayer neuronal models derived from iPSCs have exhibited challenges, including undesired cellular heterogeneity and batch-to-batch variability (Volpato et al., 2020; Schwartzentruber et al., 2018). Single cell RNA sequencing analysis of dopaminergic neuronal cultures, based on the widely used Kriks et al. 2011 protocol, has revealed significant heterogeneity among the cells (Fernandes et al., 2020). Furthermore, distinct transcriptional profiles and sensitivities to genetic and toxin-induced stress were observed among different subtypes of neurons within the culture. The efficiency of dopamine neuron differentiation using

current protocols varies widely, typically falling within the 10-70% range (Marton et al., 2019). These protocols are cost-intensive due to their reliance on small molecules and morphogens and often require substantial time investments to yield highly pure dopaminergic neurons. Interestingly, cultures generated using these differentiation methods do not exclusively contain substantia nigra A9 neurons; instead, they encompass various dopaminergic neuronal subtypes, including ventral tegmental A10 neurons (Fernandes et al., 2020). Despite advances in the field, efficiently generating specific subtypes of midbrain dopamine neurons remains a formidable challenge. Researchers have explored direct fate conversion and cell reprogramming as potential solutions but have encountered limited success.

In the direct cell reprogramming approach, a set of transcription factors is ectopically overexpressed to drive the differentiation of somatic cells or pluripotent stem cells into the desired cell type, with or without the use of patterning small molecules or growth factors (Wang et al., 2021). The feasibility of this approach was initially demonstrated by Vierbuchen et al. in 2010 when they ectopically overexpressed *Ascl1, BRN2,* and *MYT1L* to differentiate mouse embryonic and post-natal fibroblasts into neural progenitor cells (Vierbuchen et al., 2010). Various transcription factors associated with DA neuron development, such as *LMX1A, LMX1B, FOXA2, NURR1, PITX3, ASCL1, OTX2,* and *EN1*, have been ectopically overexpressed in different cell types to generate induced dopaminergic (DA) neurons. For instance, ectopic overexpression of *FOXA2* and *NURR1, LMX1A* and *FOXA2,* or *LMX1A, LMX1B, FOXA2*, and *OTX2* in BAM iNPCs has resulted in functional DA neurons (Lim et al., 2015; Pereira et al., 2014; Pfisterer et al., 2011). Combining *ASCL1, LMX1A,* and *NURR1* or *FOXA2, PITX3, and EN1* has also been shown to produce DA neurons from both mouse and human fibroblasts. These DA neurons have demonstrated dopamine release and electrophysiological function after transplantation into mouse PD brains (Caiazzo et al., 2011; Kim et al., 2011). Additionally, DA neurons have been generated from human iPSCs through ectopic expression of *ASCL1, LMX1B, and NURR1* or *ASCL1, NURR1, EN1, FOXA2, LMX1A, and PITX3* (Powell et al., 2021; Ng et al., 2021). Furthermore, ectopic overexpression of *NGN2* and *ATOH1*, either alone or in combination with SHH, FGF8, DAPT, or *ATOH1* alone, has been demonstrated to produce DA neurons (Xu et al., 2019; Sagal et al., 2014). Despite these promising results, most of these protocols face limitations, such as inefficiency for large-scale experiments, insufficient characterization of generated neurons, or the need for costly small molecules and morphogens in addition to TF overexpression. To overcome these challenges, there is a pressing need to develop an efficient protocol capable of generating highly

homogeneous and fully characterized A9 substantia nigra dopaminergic neurons on a large scale within a relatively short timeframe.

### 2.14.2.5 Midbrain Organoids

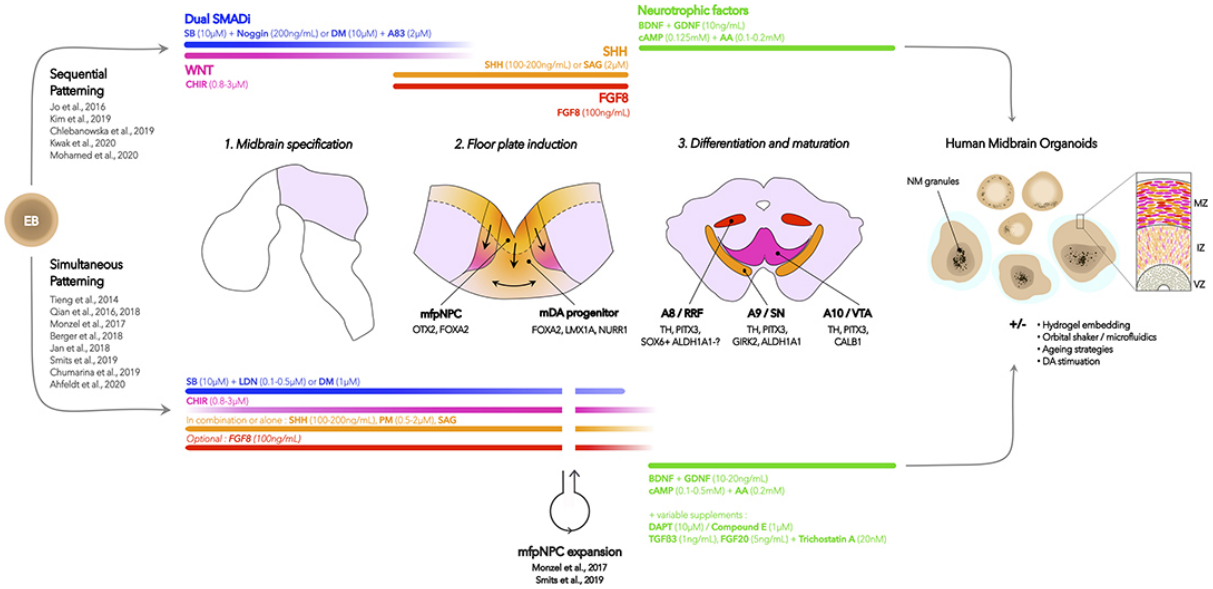



**Figure 1.17:** Schematic illustration depicting midbrain organoids differentiation strategy. Image reference: Galat et al., 2020.

Traditionally, the study of PD has relied on a two-dimensional (2D) *in vitro* culture system, which has indeed enhanced our understanding of the disease. However, this 2D approach falls short in capturing the cellular diversity and intricate architecture found in the human neural tissue (Jensen et al., 2020). Consequently, the insights gained from these models have struggled to translate effectively into clinical applications (Sun et al., 2018). Recent advances in 3D culture systems have addressed some of these limitations by enabling the generation of midbrain-like organoids from human iPSCs. These organoids closely mimic human brain development and replicate key aspects of *in vivo* architecture and physiology (Jo et al., 2016; Fiorenzano et al., 2021). Obtaining and characterizing early-stage PD brain tissue is often infeasible for various reasons, thus limiting research to end-stage disease scenarios. Therefore, there is a pressing need to establish advanced and reliable *in vitro* models that can explore the mechanisms underlying disease onset and progression.

The derivation of midbrain organoids is built upon our understanding of midbrain development, with several published protocols available (Galat et al., 2020). These protocols

primarily differ in the use of specific small molecules, their timing, and concentrations. In essence, pluripotent stem cells are aggregated to form embryoid bodies (EBs) and are initially directed toward a neuroectoderm fate through dual SMAD inhibition (targeting the BMP and TGFB/Activin/nodal pathways). Subsequently, activation of sonic hedgehog (Shh), Wnt, and FGF8 signalling pathways leads to the development of midbrain floor progenitor cells. The EBs are then embedded in a hydrogel to encourage tissue growth and organization. Further differentiation and maturation are achieved through the application of neurotrophic factors such as Glia-derived neurotrophic factor (GDNF) and Brain-derived neurotrophic factor (BDNF). Most protocols employ orbital shakers to facilitate nutrient and gas exchange within the midbrain organoids. The transcriptomic profile of human midbrain organoids closely resembles that of the human midbrain, in contrast to 2D cultures (Jo et al., 2016). Consequently, results obtained using midbrain organoids are anticipated to be more accurate and have better clinical relevance (Galat et al., 2020).

The increasing utilization of human brain organoids for studying brain biology and modelling neurological disorders is underpinned by their shared cellular, molecular, and physiological characteristics with the human brain (Smit et al., 2020). Notably, a study by Lancaster et al. (2013) has effectively demonstrated the potential of cerebral organoids as a model for identifying altered neurodevelopment. Using RNA interference and patient-specific induced pluripotent stem cells (iPSCs), they successfully modelled microcephaly, a condition challenging to replicate in animal models (Lancaster et al., 2013). Research using brain organoids has also revealed a link between Zika virus (ZIKV) infection and impaired neurogenesis. These discoveries became possible through the utilization of a 3D brain organoid model with advanced organizational features that faithfully mimic the specific embryonic development of the human brain. Furthermore, two distinct approaches employing 3D human neural cell culture systems have been reported for replicating Alzheimer's disease (AD) pathological features *in vitro* (Choi et al., 2014; Raja et al., 2016). These 3D cultures provide an environment conducive to the formation of amyloid-β (Aβ) plaques and neurofibrillary tangles (NFTs), pathological phenomena that were previously challenging to link sequentially when using 2D cultured human neurons (Raja et al., 2016; Choi et al., 2014).

Addressing the challenge of fully recapitulating alpha-synuclein pathologies has been pivotal in PD research. Recent studies have shown promising results, generating Lewy body-like inclusions in human midbrain organoids carrying glucocerebrosidase (*GBA*) and alpha-

synuclein multiplication mutations (Jo et al., 2021; Becerra-Calixto et al., 2023). Midbrain organoids harbouring the G2019S mutation in the *LRRK2* gene exhibit PD-like phenotypes, such as a reduced number of dopamine neurons (Smits et al., 2019). Additionally, in another study using midbrain organoids with the G2019S mutation in *LRRK2* or *SNCA* triplication mutation, the presence of hyperphosphorylated alpha-synuclein aggregates, a characteristic feature of pathogenic alpha-synuclein, was observed (Kim et al., 2019; Mohamed et al., 2021). These findings highlight the significant potential of 3D human midbrain organoids and their associated midbrain dopaminergic neurons as robust tools for *in vitro* disease modelling.

## 2.17. Single-Cell RNA Sequencing Technologies and Its Applications

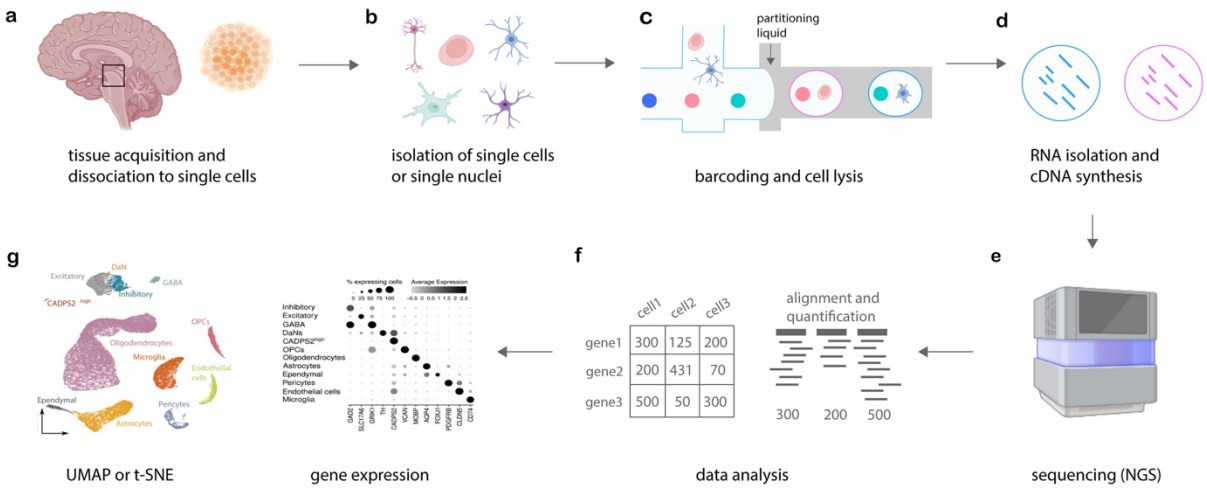



**Figure 1.18:** An overview of the single-cell RNA-sequencing procedures. Image reference: Pozojevic et al., 2023.

Single-cell RNA sequencing (scRNA-seq) technology has emerged as a powerful tool for deciphering the cellular diversity and complexity of RNA transcripts in individual cells, providing valuable insights into various cell types composition and their roles within highly organized tissues, organs, or organisms (Jovic et al., 2022). Unlike conventional sequencing methods, which identify the mean transcript expression level in populations of cells, scRNA-seq offers the ability to explore individual cells (Wang et al., 2023). A significant milestone in scRNA-seq history was the work of Tang et al. in 2009 when they successfully sequenced the transcriptomes of individual blastomeres and oocytes (Tang et al., 2009). Since then, the scRNA-seq protocol has seen substantial technical improvements, while its core principles have remained unchanged.

Over the years, scRNA-seq throughput has increased from analyzing single cells to thousands of cells per experiment. Simultaneously, the cost of scRNA-seq has decreased, and its publication rate has significantly risen. The scRNA-seq workflow primarily involves the isolation of single cells, single cell transcript capture, molecular barcoding, cDNA amplification, library preparation, deep sequencing, and data analysis (Pozojevic et al., 2023; Jovic et al., 2022). Several single-cell transcriptomic platforms, including VASA-seq, Smart-seq 2, Smart-seq 3, Drop-seq, and 10X Genomic, have been developed (Wang et al., 2023). Among these, the 10X Genomic platform has become the preferred choice due to its high throughput and cost-effectiveness. These platforms differ mainly in their single-cell separation and transcript amplification methods (e.g., all transcripts, 5' end, or 3' end). Common single-cell separation methods include micromanipulation, laser capture microdissection (LCM), fluorescence-assisted cell sorting (FACS), and microfluidics (Valihrach et al., 2018; Gross et al., 2016). Currently, microfluidics is the most preferred method and can be divided into microporous, integrated fluidic circuits (IFCs), and droplet methods (Wang et al., 2023).

The 10X Genomic platform utilizes the droplet method to capture single cells and prepare cDNA libraries for next-generation sequencing (NGS) (Zhang et al., 2017). In this approach, single cells, reverse transcription (RT) reagents, gel beads containing barcoded oligonucleotides, and oil are combined on a microfluidic chip to create reaction vesicles called Gel Beads in emulsion (GEMs) (Zheng et al., 2017). Multiple GEMs are generated in parallel through the microfluidic channels in a single run. Within each GEM, an individual cell undergoes lysis, causing the Gel Bead to dissolve, releasing identically barcoded RT oligonucleotides into the solution. This enables the reverse transcription of polyadenylated mRNA. As a result, all cDNAs from one specific cell share the same barcode, simplifying the mapping of sequencing reads back to their respective single cells. Subsequently, NGS libraries are efficiently created from these barcoded cDNAs in bulk.

NGS reads are then aligned to the reference genome, transcriptome, or epigenome, followed by quantification (Pozojevic et al., 2023; Sreenivasan et al., 2022). After quantification, the data is structured into a feature-barcode matrix. A normalization step is applied to facilitate intercellular comparisons. Principal Component Analysis (PCA) is employed to reduce dimensionality, and clustering based on shared gene expression profiles is carried out. These clusters represent groups of cells exhibiting similar molecular characteristics. To visualize these clusters, further dimensionality reduction techniques like

UMAP (Uniform Manifold Approximation and Projection) or t-SNE (t-distributed Stochastic Neighbor Embedding) are applied. In these representations, each point on the scatter plot signifies an individual cell, with neighboring cells showing similar gene expression profiles. Within these clusters, cells can be further subdivided into subtypes, often identified based on the high expression of well-known genes associated with specific cell types within distinct clusters (Pozojevic et al., 2023; Sreenivasan et al., 2022).

Despite its advantages, scRNA-seq technology has several limitations. Due to the limited starting material, scRNA-seq exhibits low transcript capture efficiency (estimated to be around 30% of the total transcriptome) and high dropout rates (Chen et al., 2019; Zhang et al., 2017). In comparison to traditional bulk RNAseq technology, scRNA-seq yields variable and noisier data (Chen et al., 2019). Additionally, most scRNA-seq platforms rely on the hybridization of barcoded oligonucleotide primers with poly(A) sequences in transcripts, which prevents the detection of non-coding RNAs and selective promoters, limiting its applicability (Wang et al., 2023). Furthermore, sample processing methodologies for scRNA-seq can induce false transcriptional stress responses due to the harsh cell/tissue dissociation process (Brink et al., 2017). Single-nucleus RNA sequencing (snRNA-seq) has been introduced as an alternative to scRNA-seq. The primary difference lies in the fact that snRNA-seq captures only mRNA present in the nucleus of the cell, unable to detect mRNA in the cytoplasm (Walter et al., 2023). This incomplete mRNA capturing remains a limitation of snRNA-seq. However, snRNA-seq offers the advantage of using samples from fresh or frozen tissues and tissues that are challenging to dissociate (Wiegleb et al., 2022).

scRNA-seq/snRNA-seq technology has revolutionized the investigation of selective cellular vulnerability in neurodegenerative diseases at the single-cell level. In PD, the mechanisms underlying the selective vulnerability of dopamine neurons have remained elusive. *In vitro* cellular models of PD often exhibit undesired heterogeneity and batch-to-batch variability (Volpato et al., 2020; Schwartzentruber et al., 2018). With the advancements in scRNA-seq technology, researchers can address these limitations and focus on relevant cell types within heterogeneous cultures. For instance, a recent study by Lang et al. (2019) used scRNA-seq analysis of iPSC-derived dopamine neurons from control and PD GBA-370S patients to identify HDAC4 as a potential therapeutic target in PD. They found that HDAC4 represses genes related to ER stress and autophagy/lysosomal pathways, shedding light on PD pathogenesis.

In a previous scRNA-seq study conducted by our group, we highlighted the cellular heterogeneity in monolayer DA neuronal cultures generated using the widely used Kriks et al. (2011) protocol. This study identified multiple DA neuronal and progenitor subpopulations, with a specific vulnerable DA neuronal population revealed after rotenone treatment. These vulnerable DA neurons exhibited gene expression profiles associated with PD, cholesterol biosynthesis, reticulons, synaptic signaling, and vesicle dynamics. They also showed downregulation of oxidative phosphorylation and cholesterol biosynthesis and upregulation of oxidative stress and heat shock genes in response to rotenone. Tunicamycin-treated cells exhibited similar downregulation of cholesterol biosynthesis, and upregulation of ER stress and apoptosis signaling. Additionally, analysis of DA neurons carrying the A53T mutation demonstrated downregulation in cholesterol metabolism, oxidative phosphorylation, glycolysis, synaptic signaling, DNA damage, and ubiquitin-proteasomal degradation, along with upregulation in chromatin remodeling and stress response genes (Fernandes et al., 2020).

While much emphasis in PD research has been placed on dopamine-producing neurons, it is essential to recognize the significance of other cell types, such as microglia, in the context of PD due to their involvement in neuroinflammation. Smajic et al. conducted a single nucleus RNA-seq study on frozen human post-mortem midbrain tissue from control and sporadic PD patients (Smajic et al., 2020). This study identified a subpopulation of neurons expressing high levels of *CADPS2*, specifically in sporadic PD samples. These neurons shared gene expression profiles with dopaminergic neurons but had low *TH* levels. Additionally, the study revealed an increase in microglia and astrocytes and a decrease in oligodendrocytes in PD samples. PD microglia showed upregulation of genes related to inflammation, while oligodendrocytes displayed upregulation of *S100B*. These findings highlight the importance of utilizing scRNA-seq/snRNA-seq technology in PD research.

## 2.18 PhD Objectives

The overarching objective of this thesis is to advance the field of Parkinson's disease research by developing next-generation cellular models.

**Objective 1:** Develop a robust 3D human midbrain organoid model and perform comprehensive characterization using single-cell RNA sequencing.

Rationale: Traditional 2D cell culture models lack the necessary cellular heterogeneity and fail to replicate *in vivo* physiological conditions. Human midbrain organoids (HMOs) offer a solution by closely emulating cellular heterogeneity, tissue cytoarchitecture, and physiological processes found *in vivo*. Utilizing midbrain organoids is expected to yield more accurate results with greater clinical relevance.

**Objective 2:** Evaluate dopaminergic neurons in the human midbrain model in response to oxidative stress inducer (rotenone) and the PD pathogenic SNCA triplication mutation and identify pathways associated with these stressors.

Rationale: The molecular and cellular mechanisms underlying the onset and progression of PD remain poorly understood. Investigating the impact of the *SNCA* triplication mutation and oxidative stress on dopaminergic neurons may provide significant insights into their roles in PD pathogenesis.

**Objective 3:** Develop induced dopaminergic neuron model using ectopic expression of transcription factors (TFs).

Rationale: Traditional induced pluripotent stem cell (iPSC)-derived monolayer neuronal models present challenges due to undesired cellular heterogeneity and batch-to-batch variability. Additionally, existing differentiation protocols for dopaminergic neurons are inefficient and resource intensive. Developing an efficient and well-characterized induced dopaminergic neuronal model through TF-based induction may address these limitations.

**Objective 4**: Investigate the cellular heterogeneity of the substantia nigra pars compacta (SNpc) and uncover cell-type-specific responses in PD using single nucleus RNA sequencing.

Rationale: Despite extensive research, the molecular and cellular mechanisms underlying PD pathogenesis remain elusive. While much attention has been given to dopamine neurons, it is essential to recognize the significance of other cell types, such as microglia, astrocytes, and oligodendrocytes. High-throughput single-cell/nucleus RNA sequencing technologies provide a unique opportunity to systematically explore cell-type-specific responses to the disease, particularly in the substantia nigra, a region of the brain affected by PD.

# 3. HUMAN MIDBRAIN ORGANOID MODEL

## 3.1 Background

Developing advanced research models is essential for understanding the cellular and molecular mechanisms underlying Parkinson's disease. Although traditional cell culture models have enhanced our disease understanding, their translation to clinical trials has been limited. These models lack the desired cellular heterogeneity and fail to mimic *in vivo* physiological environments. Human midbrain organoids (HMOs) offer a solution by closely emulating cellular heterogeneity, tissue cytoarchitecture, and physiology found *in vivo*.

Numerous protocols exist in literature for deriving midbrain organoids from pluripotent stem cells like embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs). Initially, we investigated the widely adopted protocol by Jo et al. (2016), which outlined the derivation of HMOs from ESCs. We adapted this protocol with slight modifications (substituting Noggin with LDN and Shh protein with SAG) to generate HMOs from human iPSCs. The tissue cytoarchitecture and dopaminergic identity of the resulting cells were assessed using immunohistochemistry and quantitative PCR (qPCR). Initial experiments revealed low expression of dopaminergic specification markers such as *FOXA2* and *LMX1A*. To address this misspecification, titration experiments with small molecules CHIR (a Wnt signalling activator) and SAG (a Shh signalling agonist) were conducted. Improved conditions were identified that enhanced *FOXA2* and *LMX1A* expression, but tissue cytoarchitecture remained lacking.

To address these limitations, a novel protocol named the "Ansari protocol" was devised. This protocol combined elements of Kriks et al. (2011) for patterning cells toward midbrain fate and Jo et al. (2016) for tissue induction and organoid development. HMOs generated using the Ansari protocol exhibited high *FOXA2* and *LMX1A* expression, along with *in vivo*-like tissue cytoarchitecture. Additionally, dissection of HMOs using single cell RNA sequencing revealed multiple cell types, including a dopaminergic population resembling *in vivo* human substantia nigra dopaminergic neurons. HMOs also displayed the production of neuromelanin-like pigmentation upon exposure to the dopamine precursor L-DOPA. The Ansari protocol consistently produced similar cell types across HMOs generated from various iPSC lines including patient iPSCs, embedding matrices (such as Geltrex and Matrigel), and batches. In

summary, the Ansari protocol proves to be robust, reproducible, and efficient in generating high-quality human midbrain organoids.

## 3.2 Results

### 3.2.1 Generation of human midbrain organoids using a modified Jo et al. 2016 protocol

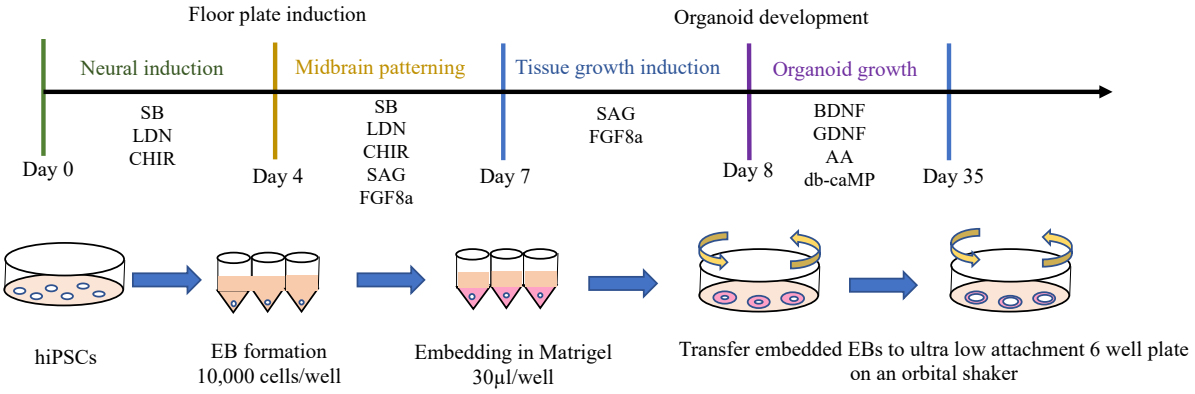



**Figure 3.1**: (A) Schematic illustration of the strategy used for HMO generation via the modified Jo et al. (2016) protocol. Components: SB = SB431542, LDN = LDN193189, SAG = Sonic Hedgehog Agonist, Pur = Purmorphamine, CHIR = CHIR99021, BDNF = Brain-Derived Neurotrophic Factor, GDNF = Glia-Derived Neurotrophic Factor, AA = Ascorbic Acid, db-cAMP = Dibutyryl Cyclic-AMP.

To generate human midbrain organoids from hiPSCs, we initially assessed a protocol published by Jo et al. in 2016, making slight modifications for our purposes (substituting Noggin with LDN and Shh protein with SAG). Originally demonstrated for generating midbrain organoids from human embryonic stem cells (hESCs), the Jo et al. 2016 protocol served as a foundation. In this study, we adapted the protocol to generate midbrain organoids from hiPSCs.

The process involved dissociation of hiPSCs into single cells, which were then seeded into a V-bottom 96-well plate to form embryoid bodies (EBs). The resulting EBs were treatment with dual SMAD inhibitors, SB431542 and LDN193189, along with a Wnt pathway activator (CHIR99021), resulting in the induction of neuroectoderm differentiation. Further neural patterning was achieved by supplementing the culture media with SAG (Sonic hedgehog signaling agonist) and FGF8a, directing the cells toward a mesencephalic fate. Subsequently, the patterned EBs were embedded in Matrigel, facilitating 3D tissue growth and organization.

To advance organoid development, the Matrigel-embedded EBs were transferred to an ultra-low attachment plate, where the presence of neurotropic factors like BDNF and GDNF promoted growth. To enhance optimal nutrient and gas exchange throughout the organoids, the plates were placed on an orbital shaker.

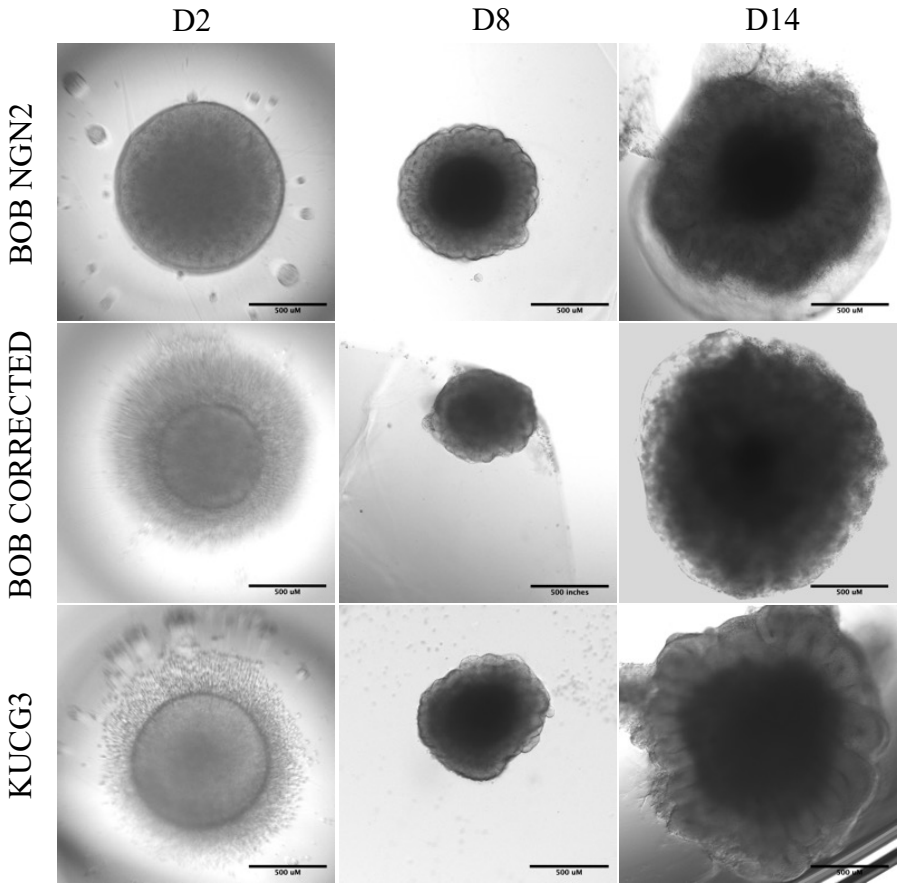



**Figure 3.2**: Brightfield images depicting the morphology of HMOs at different developmental stages generated via the modified Jo et al., 2016 protocol. Scale bar = 500µm.

The morphology of the human midbrain organoids (HMOs) was observed at various stages of their development. By day 2, the embryoid bodies (EBs) exhibited smooth, rounded, and translucent edges, indicative of successful EB formation (Figure 3.2). Following the embedding of the EBs in Matrigel on day 8, distinct neuroepithelium buds became evident (Figure 3.2). Throughout the developmental phase of the organoids, substantial growth was observed, resulting in a noticeable increase in size and the formation of a dense core (Figure 3.2).

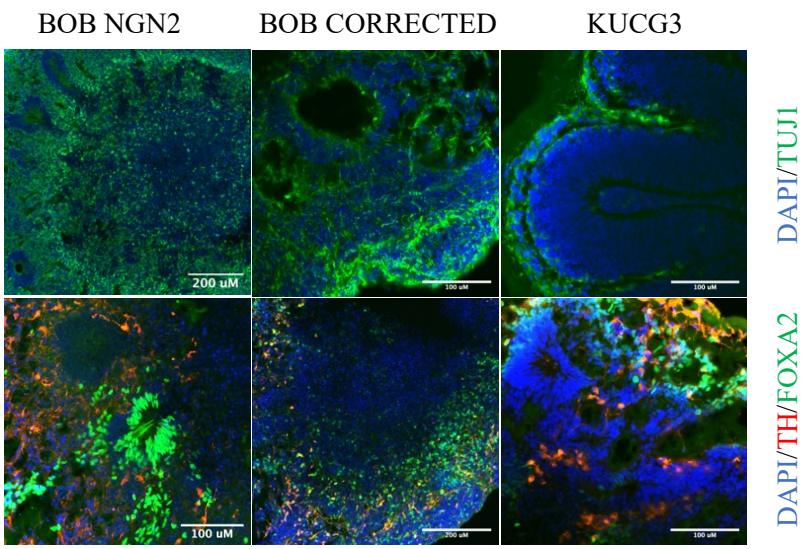



**Figure 3.3**: Immunohistochemical analysis illustrating the characterization of 4-week-old Human Midbrain Organoids (HMOs) generated using the modified Jo et al. (2016) protocol. The organoids originate from BOB NGN2, BOB-WT, and KUCG3 iPSCs. The analysis includes dopaminergic markers (TH, LMX1A, FOXA2) and pan-neuronal Markers (TUJ1). LMX1A Staining images are excluded due to non-detection in HMOs.

For the initial characterization of the HMOs, 35-day-old HMOs were fixed and cryosectioned to produce 12µm-thick sections. These sections were subsequently subjected to immunostaining to enable the investigation of tissue cytoarchitecture and the identification of midbrain dopaminergic neurons. To accurately pinpoint and confirm midbrain dopaminergic neurons, the sections were immunostained for FOXA2, LMX1A, and TH—well-established markers within the field. Additionally, TUJ1 was utilized as a pan-neuronal marker. TUJ1-positive neurons were found to be situated on the periphery of the ventricle-like structure, indicative of the midbrain organoids displaying a tissue organization (Figure 3.3). Notably, the organoids featured TH-positive cells, a subset of which exhibited dual positivity for FOXA2. However, cells co-expressing FOXA2 and TH appeared to be relatively sparse in numbers (Figure 3.3). An evaluation of LMX1A, another pivotal midbrain progenitor marker, yielded unexpectedly low counts of positive cells. This finding suggested that a significant portion of midbrain progenitor cells within our organoids were potentially mis-patterned.

### 3.2.2 Optimizing the Jo et al. 2016 Protocol to Enhance LMX1A and FOXA2 Expression

Within the context of addressing the challenge of low LMX1A and FOXA2 expression, an optimization effort was undertaken based on the Jo et al. (2016) protocol.

During midbrain development, the regulation of *LMXA1* expression is orchestrated through a feedback loop mediated by Wnt signalling. Achieving co-expression of *LMX1A* and *FOXA2* within midbrain progenitor cells requires robust activation of the sonic hedgehog signalling pathway. Consequently, the optimization process included the manipulation of various concentrations of CHIR99021 (a Wnt pathway activator) and SAG (a sonic hedgehog signalling agonist), ranging from low to high, during distinct stages of midbrain floor specification.

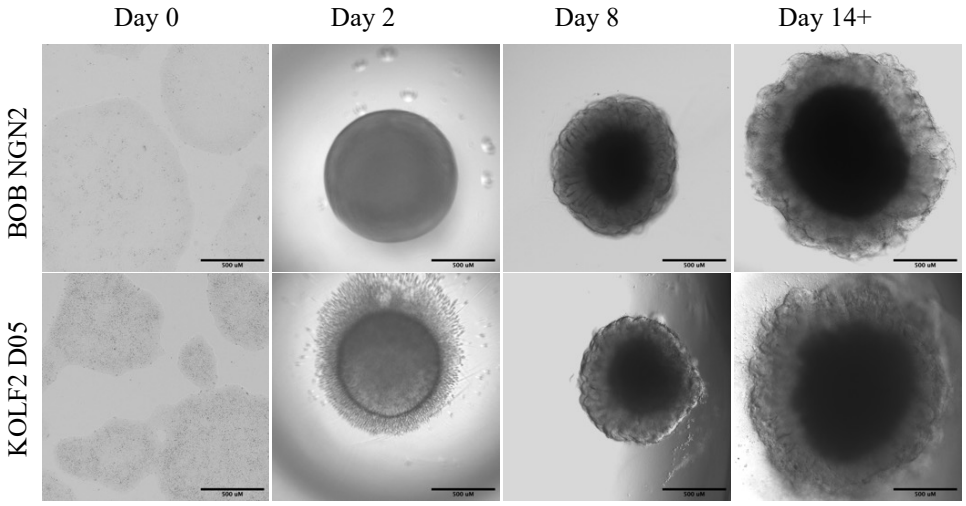



**Figure 3.4**: Brightfield images depicting the morphology of HMOs generated using a modified Jo et al. (2016) protocol with CHIR99021(3uM) and SAG(0.5uM) at different developmental stages. Scale bar = 500µm.

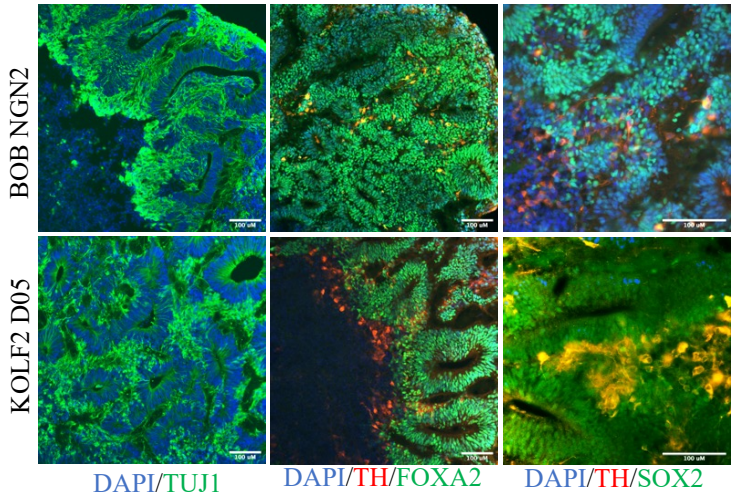



**Figure 3.5**: Immunohistochemical analysis illustrating the characterization of 4-week-old Human Midbrain Organoids (HMOs) generated using a modified Jo et al. (2016) protocol with

CHIR99021(3uM) and SAG(0.5uM). The organoids originate from BOB NGN2, and KOLF2 D05 iPSCs. The analysis includes dopaminergic markers (TH, LMX1A, FOXA2) and pan-neuronal Markers (TUJ1). LMX1A Staining images are excluded due to non-detection in HMOs.

Initial attempts involved the generation of human midbrain organoids (HMOs) by employing a media supplemented with high levels of CHIR99021 (3µM) and SAG (0.5µM). Notably, the morphology of both embryoid bodies (EBs) and HMOs exhibited a normal progression across different developmental stages (Figure 3.4). After the generation of these HMOs, an immunohistochemical assessment was conducted, targeting LMX1A, FOXA2, TH, SOX2, and TUJ1.

The immunohistochemical analysis revealed the presence of ventricle-like structures having cells positive for SOX2 and FOXA2(Figure 3.5). Cells expressing TUJ1 were predominantly situated on the periphery of these ventricle-like structures (Figure 3.5). These findings strongly suggested that the cellular organization mimicked the organization observed *in vivo* tissue. Intriguingly, a significant population of FOXA2-expressing cells was evident in both KOLF2 D05 and BOB NGN2-derived HMOs (Figure 3.5). The number of TH positive cells were found to be in low number for both KOLF2 D05 and BOB NGN2-derived HMOs (Figure 3.5). However, it's noteworthy that HMOs generated using the 3µM CHIR and 0.5µM SAG concentrations failed to exhibit a significant increase in LMX1A expression.

### 3.2.3 Addressing Low LMX1A Expression in HMOs: CHIR Titration Experiment

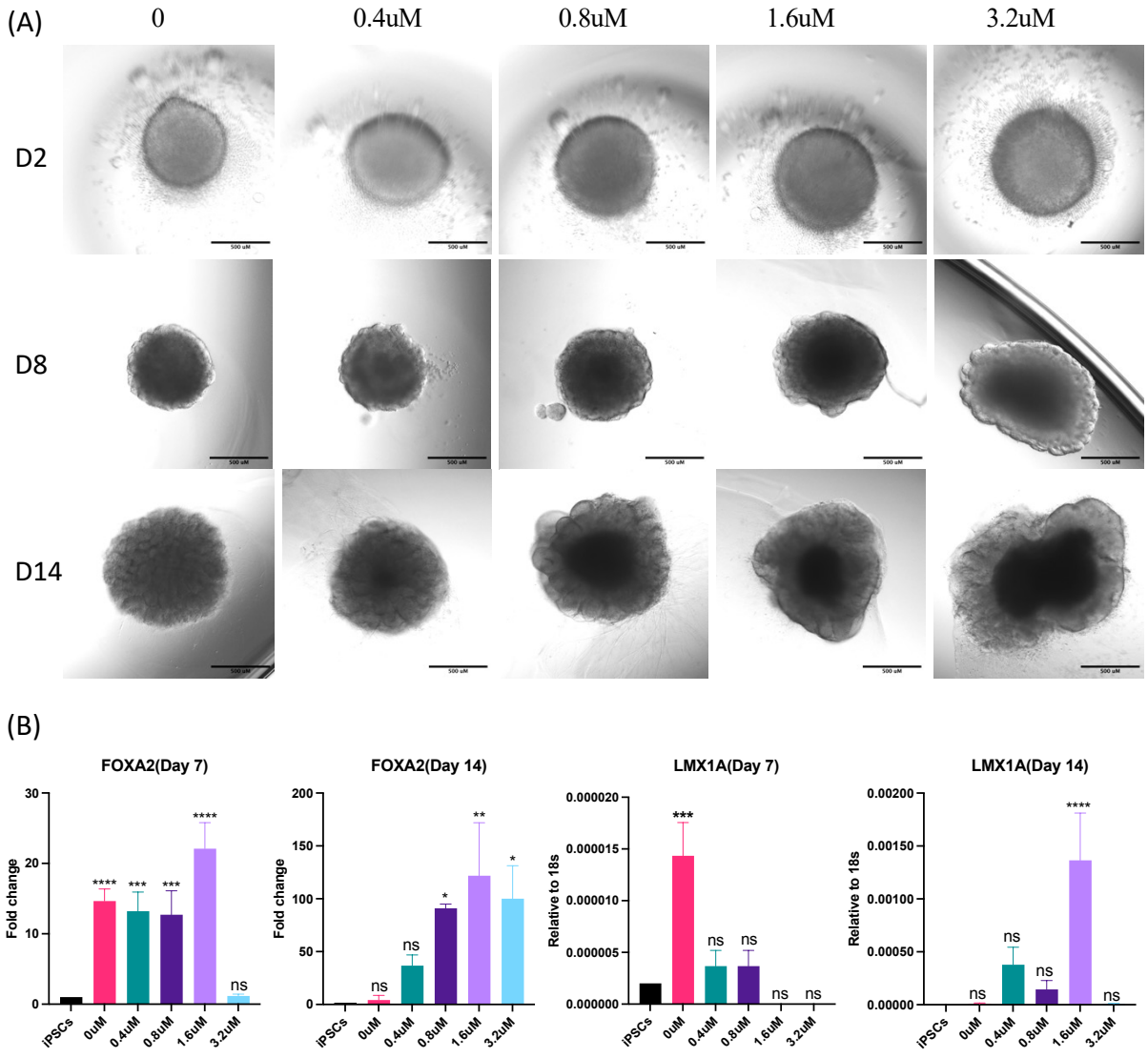



**Figure 3.6**: (A): Brightfield images depicting the morphology of HMOs generated using a modified Jo et al. (2016) protocol with various concentrations of CHIR99021 at different developmental stages. Scale bar = 500 µm. (B) Bar plots presenting qPCR analysis results for the expression of *LMX1A* and *FOXA2* in HMOs generated using the Jo et al. (2016) protocol with various CHIR99021 concentrations on day 7 and day 14. Data are represented as ± SEM, n = 3. Statistical analysis was performed using one-way ANOVA with Dunnett's multiple comparisons test. Significance levels: *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001, in comparison to the control (iPSCs).

To overcome the challenge of insufficient LMX1A expression within HMOs, a comprehensive CHIR titration experiment was conducted. This approach included generating

HMOs across a range of CHIR concentrations, ranging from low to high. Specifically, CHIR titration concentrations included 0.4µM, 0.8µM, 1.6µM, and 3.2µM.

The primary objective of this endeavour was to identify optimal midbrain patterning of the cells. To this end, developing HMOs were collected at two timepoints, day 7 and day 14, for subsequent qPCR analysis. The qPCR analysis showed a noteworthy finding: the highest *LMX1A* expression was evident in HMOs generated with 1.6µM CHIR concentration on day 14(Figure 3.6B). Intriguingly, negligible *LMX1A* expression was observed in the HMOs utilizing 1.6µM CHIR on day 7. Concurrently, HMOs produced with 1.6µM CHIR displayed the most robust *FOXA2* expression levels on both day 7 and day 14(Figure 3.6B).

Turning our attention to the morphological aspect of HMOs at various timepoints, a distinct pattern emerged. HMOs generated with 1.6µM CHIR or lower concentrations exhibited a smaller overall size compared to their counterparts generated using 3.2µM CHIR (Figure 3.6A). Furthermore, it was noted that HMOs cultured with lower concentrations of CHIR displayed a correspondingly diminished necrotic core, as expected due to their smaller size(Figure 3.6A).

### 3.2.4 Refining CHIR Concentration for Enhanced LMX1A Expression in HMOs

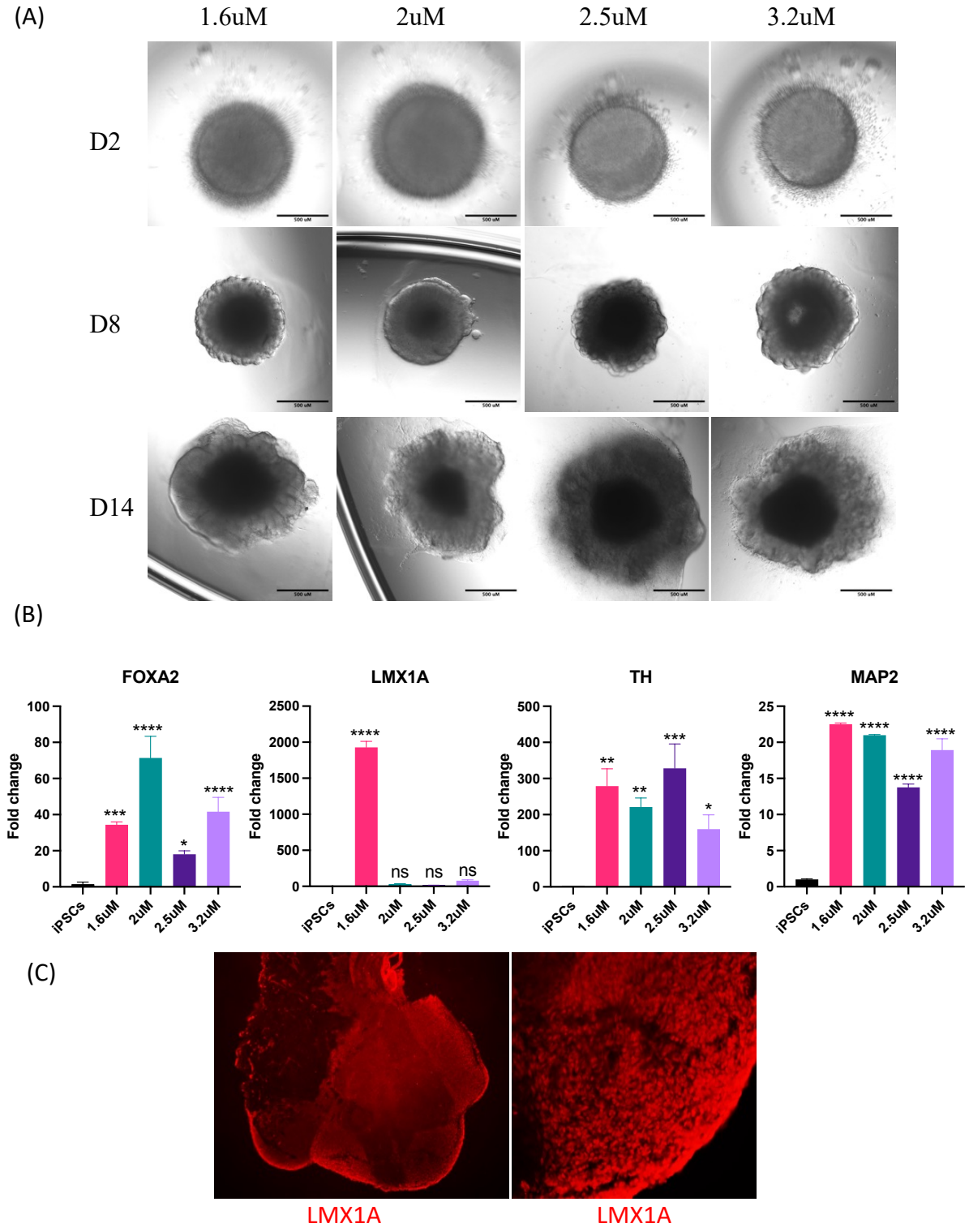



**Figure 3.7**: (A) Brightfield images depicting the morphology of HMOs generated using a modified Jo et al. (2016) protocol with various concentrations of CHIR99021 at different developmental stages. Scale bar = 500µm. (B) Bar plot presenting qPCR analysis results for the expression of *LMX1A* and *FOXA2* in HMOs generated using the Jo et al., 2016 protocol

with various CHIR99021 concentrations on day 7 and day 14. Data are represented as ± SEM, n = 3. Statistical analysis was performed using one-way ANOVA with Dunnett's multiple comparisons test. Significance levels: *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001, in comparison to the control (iPSCs). (C) Immunohistochemical images depicting expression of LMX1A in 1.6 µM HMOs at 4x(left) and 10x(right) magnification.

Taking a step further in the pursuit of optimizing CHIR concentration to elevate LMX1A expression within HMOs, an investigation including CHIR concentrations ranging from 1.6µM to 3.2µM was conducted. This experiment aimed to uncover the most optimal CHIR concentration for enhanced *LMX1A* expression in HMOs. The assessment of HMOs was conducted via qPCR analysis specifically focused on the 28th day of organoid development.

The qPCR analysis yielded a striking result, with the highest *LMX1A* expression distinctly observed within HMOs generated using the 1.6µM CHIR concentration (Figure 3.7B). In contrast, other concentrations exhibited either low or negligible *LMX1A* expression (Figure 3.7B). Notably, the 1.6µM CHIR-derived HMOs also exhibited elevated expression levels of *MAP2* and *SOX2*, indicating a higher proportion of neuronal and stem cells within these organoids. *TH* and *FOXA2*, important dopaminergic markers, were likewise found to be upregulated in the 1.6µM CHIR-generated HMOs (Figure 3.7B).

To corroborate the LMX1A expression in the 1.6µM CHIR-generated HMOs, immunohistochemistry was performed, revealing a significant number of cells expressing LMX1A. However, it's worth noting that ventricle-like structures were not observed within the 1.6µM CHIR-derived HMOs (Figure 3.7C).

### 3.2.5 Exploration of Additional Optimization Conditions

(A)

| | Day 1- Day 8 | Day 0 | Day 2 | Day 3 | Day 4 | Day 6 | Day 7 | Day 8 | Day 9 |
|---|---|---|---|---|---|---|---|---|---|
| Condition | SAG | CHIR99021 | | | | | | | |
| 1 | 0.5 | 1.6 | 1.6 | | 1.6 | | Matrigel | Maturation | |
| 2 | 0.5 | 1.6 | 1.6 | | 3.2 | | Matrigel | Maturation | |
| 3 | 0.5 | 1.6 | 3.2 | | 3.2 | | Matrigel | Maturation | |
| 4 | 0.5 | 1.6 | 1.6 | 3.2 | 3.2 | | Matrigel | Maturation | |
| 5 | 2 | 1.6 | 1.6 | | 1.6 | | Matrigel | Maturation | |
| 6 | 2 | 3.2 | 3.2 | | 3.2 | | Matrigel | Maturation | |
| 7 | 0.5 | 1.6 | 1.6 | 3.2 | 3.2 | 3.2 | | Matrigel | Maturation |
| 8 | 2 | 1.6 | 1.6 | | 1.6 | 1.6 | | Matrigel | Maturation |
| 9 | 2 | 3.2 | 3.2 | | 3.2 | 3.2 | | Matrigel | Maturation |

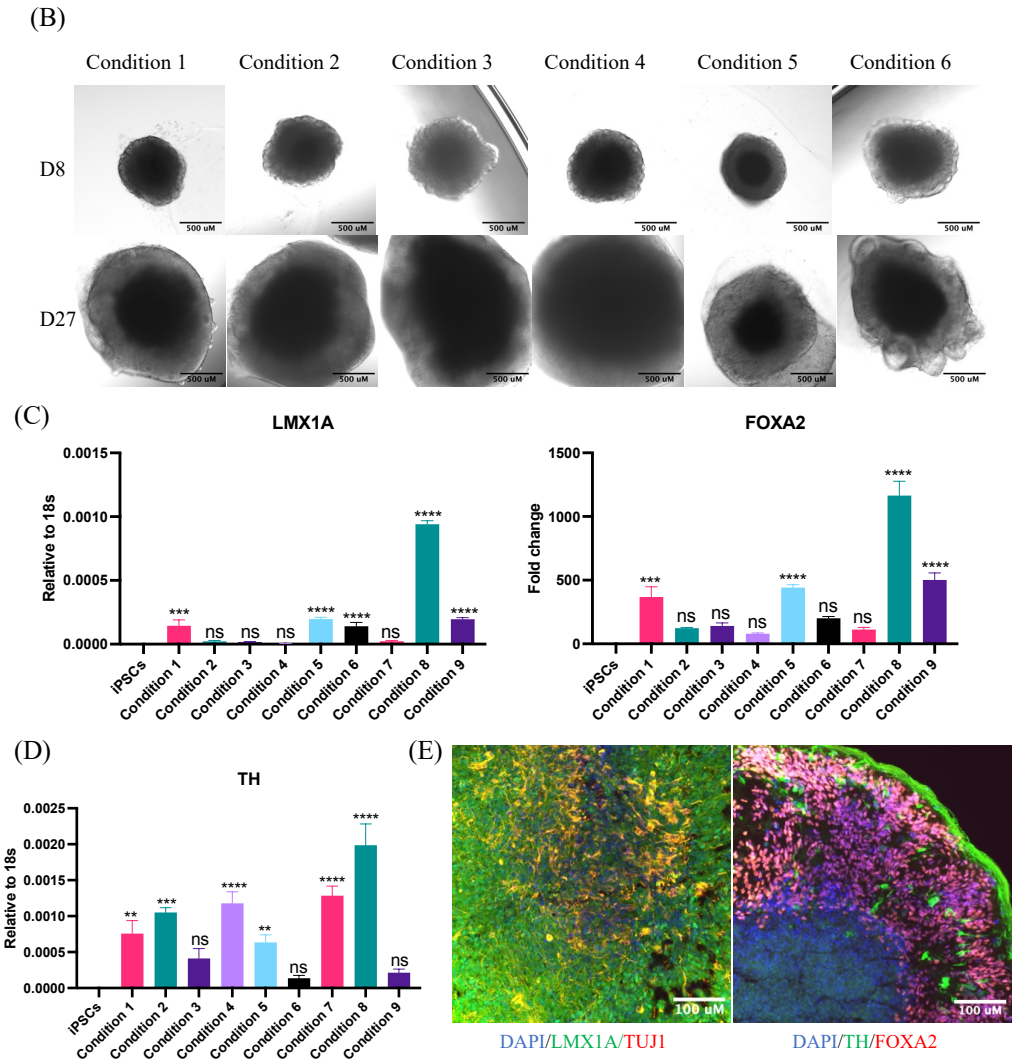



**Figure 3.8**: (A) A matrix plot illustrating the combination of SAG and CHIR concentration tested to generate HMOs. (B) Brightfield images depicting the morphology of HMOs generated using a modified Jo et al. (2016) protocol with changes highlighted in (A) at different

developmental stages. Scale bar = 500µm. (C, D) Bar plot presenting qPCR analysis results for the expression of *LMX1A, FOXA2,* and *TH* in 4-weeks old HMOs on day 28. Data are represented as ± SEM, n = 3. Statistical analysis was performed using one-way ANOVA with Dunnett's multiple comparisons test. Significance levels: *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001, in comparison to the control (iPSCs). (C) Immunohistochemical staining of 4-week-old HMOs sections for TH, TUJ1, FOXA2, and LMX1A.

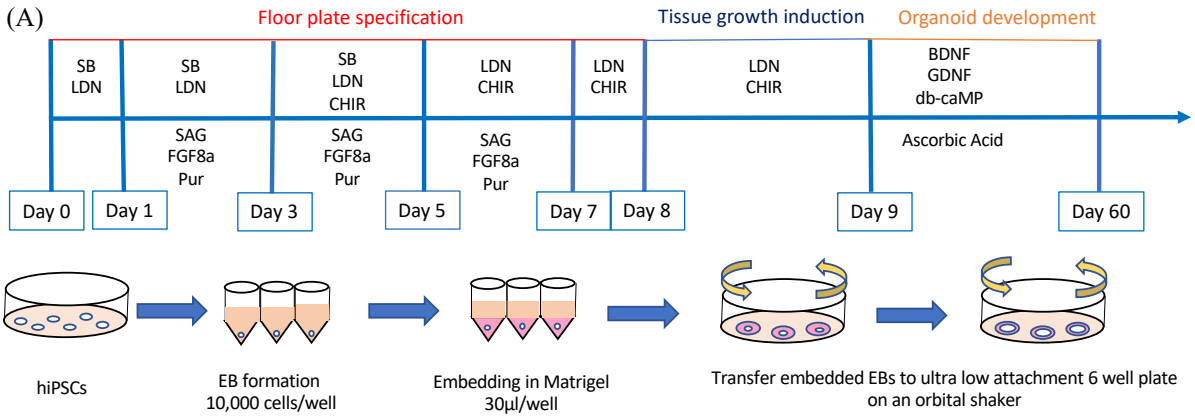
In a pursuit of further optimizing the midbrain organoid development protocol, an extended array of conditions was explored. This includes varying concentrations of CHIR99021 and SAG across a range from low to high, strategically implemented at distinct stages of midbrain floor specification. These optimization experiment also included an extension of cell patterning duration to 5 days, a day more than the previously applied 4-day period (Figure 3.8A).

Intriguingly, the subsequent qPCR analysis of these optimization conditions revealed a consistent trend of low LMX1A expression across all tested conditions (Figure 3.8C). Nonetheless, conditions 8 and 9, characterized by the longer 5-day cell patterning duration, emerged as noteworthy exceptions, showcasing the highest LMX1A expression levels (Figure 3.8C).

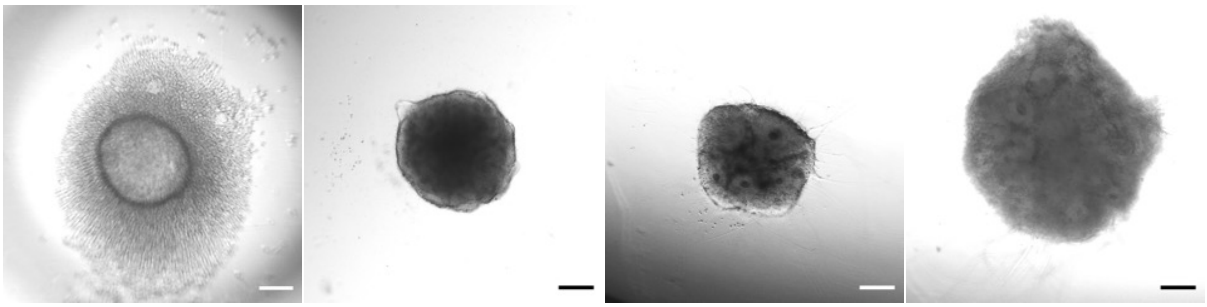
However, immunohistochemical analysis of the midbrain organoids generated using condition 8 and 9 showed that, despite the higher *LMX1A* expression detected by qPCR, the number of *LMX1A*-positive cells remained exceedingly low (Figure 3.8E). Furthermore, the characteristic ventricle-like structures, indicative of proper midbrain development, were absent within the midbrain organoids of condition 8 and 9(Figure 3.8E).

### 3.2.6 Generation and Characterization of Human Midbrain Organoids Using a Combined Approach of Jo et al. 2016 and Kriks et al. 2011 Protocols (Ansari protocol)

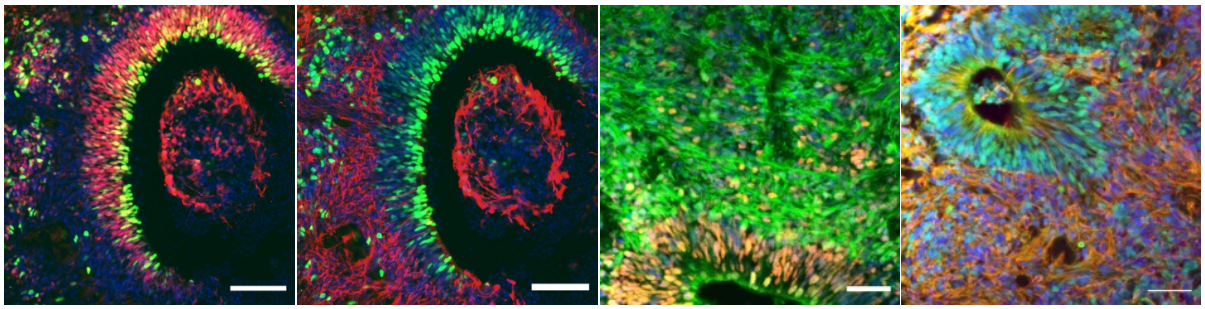
(A)




(B)




(C)




(D)

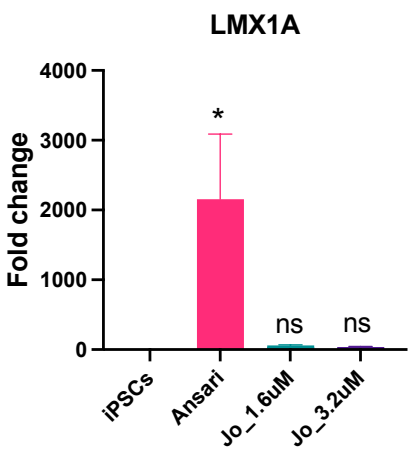


**Figure 3.9:** (A) Schematic illustration of the strategy employed for HMO generation using Ansari protocol. Components: SB = SB431542, LDN = LDN193189, SAG = Sonic Hedgehog Agonist, Pur = Purmorphamine, CHIR = CHIR99021, BDNF = Brain-Derived Neurotrophic Factor, GDNF = Glia-Derived Neurotrophic Factor, AA = Ascorbic Acid, db-cAMP = Dibutyryl Cyclic-AMP. (B) Brightfield images depicting the morphology of HMOs generated using Ansari protocol at different developmental stages. Scale bar = 500µm. (C) Immunohistochemical staining of 4-week-old HMOs sections for TH, TUJ1, SOX2, Ki67, FOXA2, and LMX1A. (D) Bar plot presenting qPCR analysis results for the expression of *LMX1A* in 4-weeks old HMOs generated using Ansari protocol and Jo et al. (2016) (SAG= 0.5µM; CHIR99021= 1.6µM & 3.2µM). Data are represented as ± SEM, n = 3. Statistical analysis was performed using one-way ANOVA with Dunnett's multiple comparisons test. Significance levels: *$p < 0.05$, in comparison to the control (iPSCs).

Considering our previous attempts, which yielded less than satisfactory results through the modified Jo et al. 2016 protocol, we employed a novel approach. This involved a strategic combination of the Jo et al. 2016 and Kriks et al. 2011 protocols. In this innovative method, midbrain patterning of embryoid bodies (EBs) was carried out by employing the Kriks et al. 2011 protocol, while the subsequent organoid development was orchestrated through the Jo et al. 2016 protocol. We named this protocol as "Ansari protocol".

Central to this protocol was the simultaneous induction of neural differentiation and midbrain patterning. Similar to the previous approach, the activation of Wnt signalling was achieved using CHIR99021, albeit on day 3(Figure 3.9A). Sonic hedgehog signalling activation involved the synergistic use of SAG and purmorphamine. Notably, the duration of the patterning stage was extended to 7 days, compared to the previous 4-day interval. Interestingly, the morphological progression of cells throughout various stages of organoid development resembled that observed in the previous protocol (Figure 3.9B).

Consequently, midbrain organoids generated through this protocol were collected on day 30 for characterization. Immunohistochemical analysis showed promising results, unveiling the presence of desired FOXA2+/LMX1A+ midbrain progenitor cells (Figure 3.9C). These organoids also featured distinct ventricle-like structures, comprising SOX2 and/or Ki67 positive cells (Figure 3.9C). TUJ1-positive neurons were predominantly situated on the exterior of these ventricle-like structures, with a noteworthy proportion exhibiting TH positivity (Figure

3.9C). These observations collectively confirmed the highly organized layers, comprising both proliferative and non-proliferative neural stem cells, as well as diverse neuron subtypes.

To quantitatively assess *LMX1A* expression, a comparative qPCR analysis was conducted on the organoids generated using this optimized protocol with those produced using earlier methodologies. Strikingly, the midbrain organoids derived using this protocol showed a significantly higher expression of *LMX1A*(Figure 3.9D).

### 3.2.7 Generation of Midbrain Organoids from Patient iPSCs with an *SNCA* Triplication and It's Isogenic Control

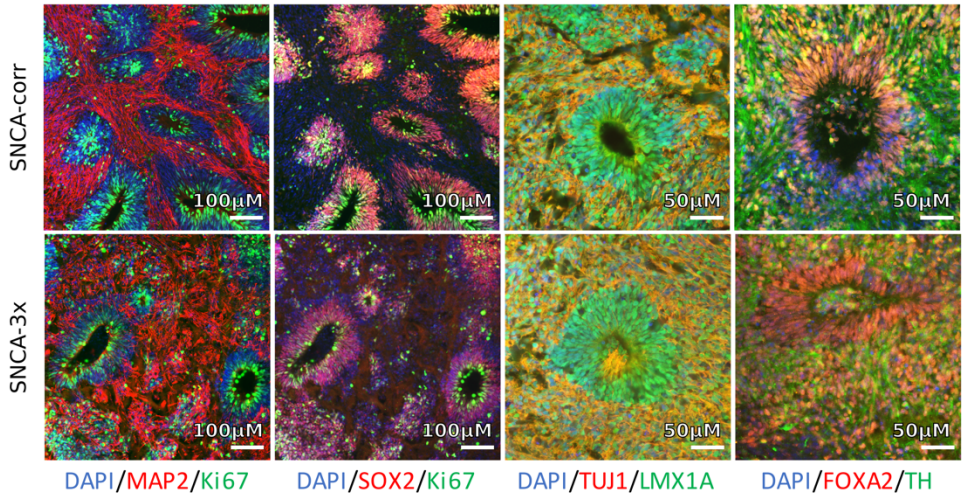



**Figure 3.10**: Immunohistochemical analysis depicting the characterization of 4-week-old human midbrain organoids (HMOs). The organoids were derived from iPSCs from PD patients with *SNCA* triplication, and isogenic control iPSCs with corrected *SNCA* triplication. The analysis includes dopaminergic markers (TH, FOXA2, LMX1A), pan-neuronal markers (TUJ1, MAP2), the proliferation marker Ki67, and the stem cell marker SOX2.

To evaluate the robustness of the Ansari protocol in generating HMOs from various human iPSC lines, including those derived from patients, we generated HMOs using patient-derived iPSCs with an *SNCA* triplication (*SNCA*-3x). These iPSCs carry two extra copies of the *SNCA* gene. For comparison purposes, we also differentiated an isogenic control iPSC line with the corrected SNCA triplication (SNCA-corr).

Immunohistochemistry was performed to demonstrate the organized tissue structure of the organoids, as well as the presence of specific cell types, including dopaminergic neurons.

Immunostaining highlighted the existence of *in vivo*-like tissue organization within the organoids, as indicated by the proliferation marker Ki67, the neural stem cell marker SOX2, and the mature neuronal marker MAP2 (Figure 3.10). The staining also revealed that ventricle-like structure, exhibited cells positive for Ki67 and SOX2. Notably, SOX2-positive cells were predominantly located in the outer layers, whereas Ki67-positive cells were mainly found in the inner layers (Figure 3.10). The MAP2-positive cells were primarily situated on the periphery of these ventricle-like structures (Figure 3.10). Collectively, these results pointed toward a *in vivo*-like tissue organization within the HMOs.

The dopaminergic identity of the cells was evaluated through the assessment of dopaminergic marker genes, including FOXA2, LMX1A, and TH. The ventricle-like structures displayed positive staining for SOX2, Ki67, FOXA2, and LMX1A, suggesting the presence of both proliferating and non-proliferating midbrain progenitors (Figure 3.10). Dopaminergic neurons within the HMOs were identified using TH expression, with a significant proportion also showing positivity for FOXA2(Figure 3.10). The co-staining of a pan-neuronal marker TUJ1 and LMX1A confirmed the existence of LMX1A-positive neurons in the organoids (Figure 3.10). These results collectively indicated that HMOs exhibited tissue organization resembling *in vivo* conditions and contained midbrain dopaminergic neurons. Remarkably, similar immunostaining results were obtained for HMOs generated from both healthy and SNCA triplication iPSCs.

### 3.2.8 Cellular heterogeneity in HMOs following differentiation

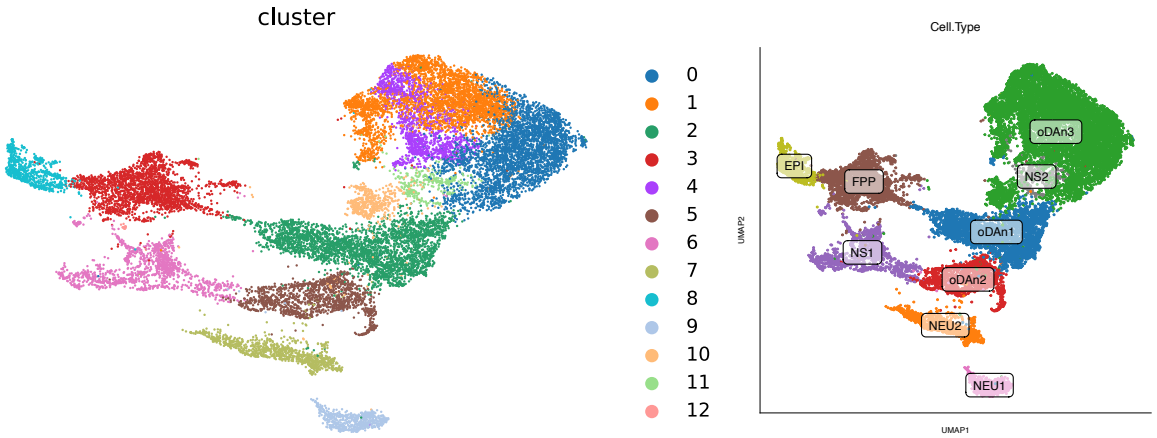
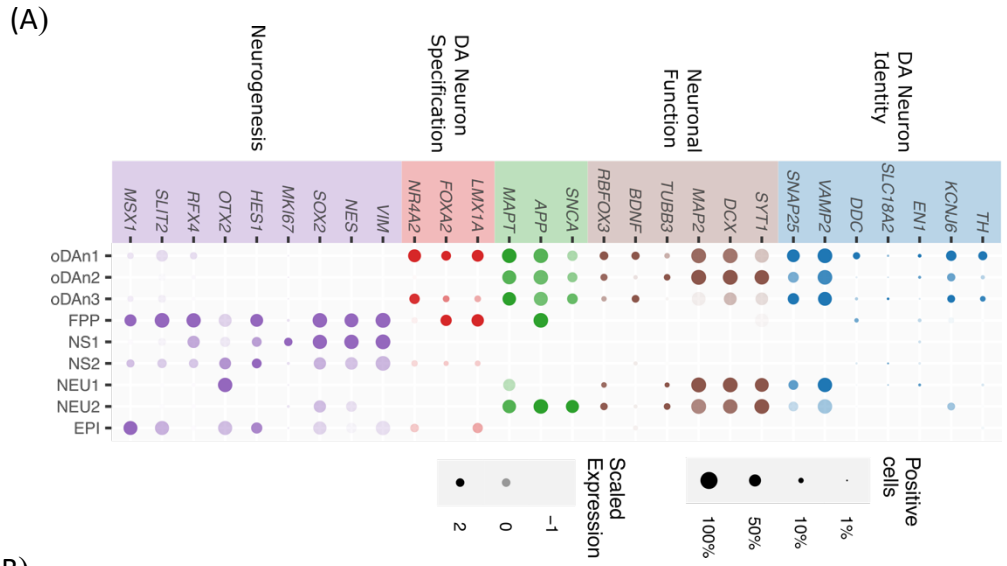


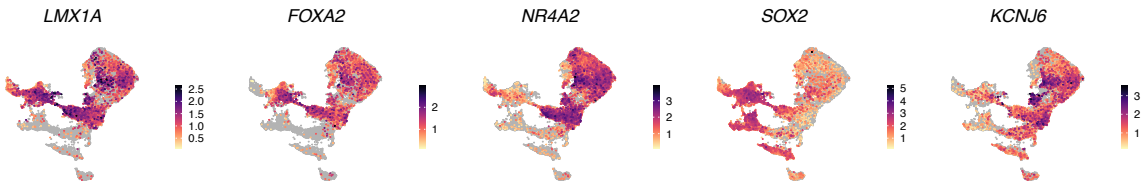

**Figure 3.11**: UMAP embedding of a unified dataset including cells from KOLF2, *SNCA*-3x, and *SNCA*-corr, totaling 19,157 cells. Through unsupervised clustering, 12 distinct clusters

were identified and further classified into 9 different cell types using a panel of marker genes. Annotated cell types are represented by varying colors. Cell type annotations include: Floor Plate Progenitors (FPP), Neural Stem Cells (NS1, NS2), Organoid Dopamine Neurons (oDAn1, oDAn2, oDAn3), and Non-Dopaminergic Neurons (NEU1, NEU2). **Note:** These plots were generated by Nikolas Patikas.

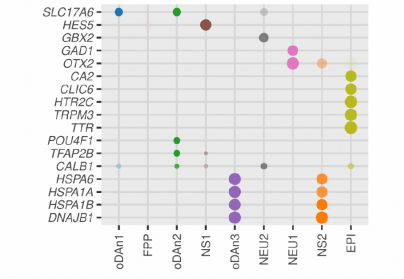
(A)



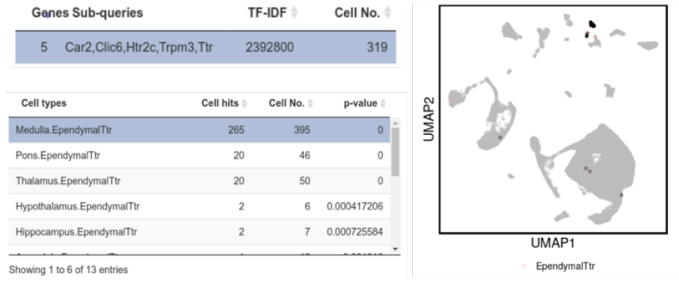

(B)




(C)




(D)



(E)

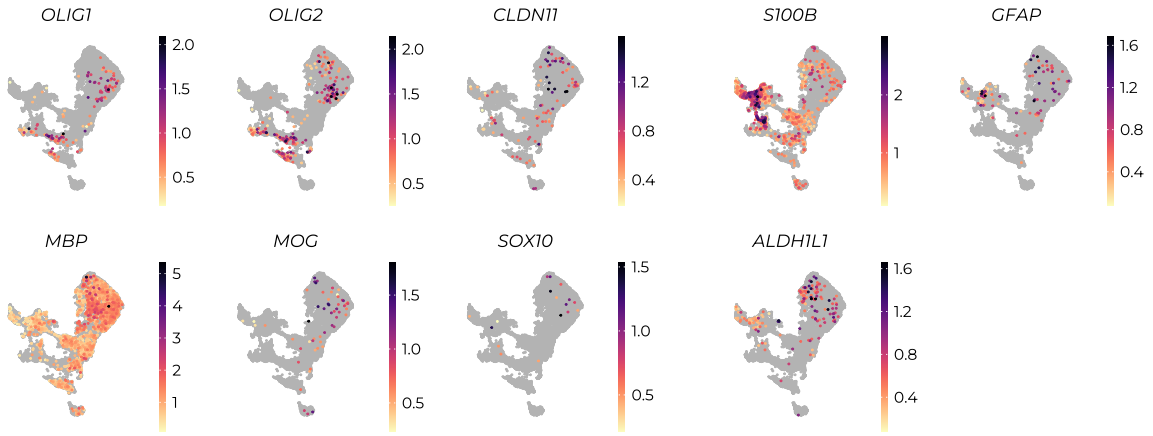



**Figure 3.12:** (A) Matrix plot illustrating the gene expression of markers associated with various stages of dopaminergic neuron development across distinct cell types within HMOs. (B) UMAP heatmaps displaying dopaminergic neuron markers and the stem cell marker *SOX2*, highlighting their specificity within the identified cell types. (C) Dot plot showing the expression of genes linked to the Unfolded Protein Response (UPR) (*HSPA6, HSPA1A, HSPA1B*), the VTA marker *OTX2*, hindbrain marker *GBX2*, and the GABAergic marker *GAD1* in distinct cell types HMOs. (D) Cell annotation as EPI was determined by the expression of *CAR2, CLIC6, HTR2C, TRPM3,* and *TTR* utilizing the scfind tool. (E) UMAP heatmaps presenting log2-normalized expression of markers for oligodendrocytes (e.g., *OLIG2, OLIG1, CLDN11*) and astrocytes (*MBP, MOG, SOX10*). **Note:** These plots were generated by Nikos Patikas.

To explore the cellular heterogeneity and gene expression profiles of the distinct cell types, single-cell RNA sequencing (scRNA-seq) was employed. For this experiment, day 49 HMOs from each line were combined and dissociated into single cells, subsequently undergoing scRNA-seq processing.

Initial processing of the raw scRNA-seq data involved filtering to eliminate doublets and cells with a low count of detected genes. Data from KOLF2, *SNCA*-3x, and *SNCA*-corr HMOs were merged to form a comprehensive dataset comprising 19,157 cells. Unsupervised clustering analysis unveiled 12 distinct clusters, which were characterized and labeled as 9 distinct cell types using a panel of marker genes (Figure 3.11).

Within the HMOs, three populations of dopaminergic neurons (labeled as oDAn1, oDAn2, and oDAn3), two non-dopaminergic neuron populations (designated as NEU1 and

NEU2), two neural stem cell populations (referred to as NS1 and NS2), a floor plate progenitors (FPP) population, and an ependymal-like cell population (EPI) were identified (Figure 3.11). EPI cells were annotated based on the expression of *CAR2, CLIC6, HTR2C, TRPM3*, and *TTR* (Figure 3.12D). Interestingly, glial cell populations were notably absent in the day 49 HMOs, aligning with findings from Fiorenzano et al., 2021, at similar timepoint. Genes associated with glial cells, such as *GFAP, OLIG1, MOG, SOX10,* and *CLDN11*, were expressed at minimal levels (Figure 3.12E).

Floor plate progenitors (FPP) and neural stem cells (NS1 and NS2) expressed *SOX2* and *VIM*, with NS1 also displaying expression of the proliferative cell marker *MKI67*, indicative of proliferating neural stem cell identity (Figure 3.12A). FPP cells expressed dopaminergic markers including *FOXA2* and *LMX1A*, while none of these populations expressed neuronal markers (Figure 3.12A). NS2 expressed low levels of *FOXA2* and *LMX1A*, but notably elevated levels of *HES1*, a gene linked to dopaminergic neuron development (Reference) (Figure 3.12A). NEU1 and NEU2, the non-dopaminergic neuronal populations, exhibited the expression of neuronal markers such as *SNAP25, SYT1, MAP2,* and *DCX* (Figure 3.12A). NEU1 also expressed *OTX2*, a ventral tegmental area (VTA) marker, and *GAD1*, a GABAergic marker (Figure 3.12C). Interestingly, neurodegenerative disease-associated genes like *SNCA, APP,* and *MAPT* were prominently expressed in NEU2(Figure 3.12A). Moreover, NEU2 exhibited high expression of *GBX2*, a hindbrain marker (Figure 3.12C).

All three dopaminergic neuronal populations exhibited expression of pan-neuronal marker genes, including *SYT1* and *SNAP25*, along with dopaminergic neuronal markers such as *TH, SNCA,* and *KCNJ6*(Figure 3.12A). Notably, oDAn1 demonstrated increased expression of dopaminergic markers *LMX1A, NURR1, and FOXA2*(Figure 3.12A*)*. In contrast, oDAn2 expressed low levels of *FOXA2, LMX1A, NURR1,* and *TH* (Figure 3.12A). As for oDAn3, robust expression of *TH, SNCA, NURR1,* and *SNAP25* was observed (Figure 3.12A). However, the expression of neuronal functional markers, including *SYT1, DCX,* and *MAP2*, was limited in oDAn3(Figure 3.12A). Intriguingly, oDAn3 exhibited significant expression of genes associated with the unfolded protein response, such as *HSPA1A, HSPA1B, DNAJB1, and HSPA6*(Figure 3.12C*)*.

In summary, scRNA-seq analysis of the HMOs revealed complex cellular heterogeneity, unraveling distinct cell types and their unique gene expression profiles, shedding light on the complex composition and dynamics within these midbrain organoids.

### 3.2.9 Cell Composition and Robustness of Differentiation in Human Midbrain Organoids

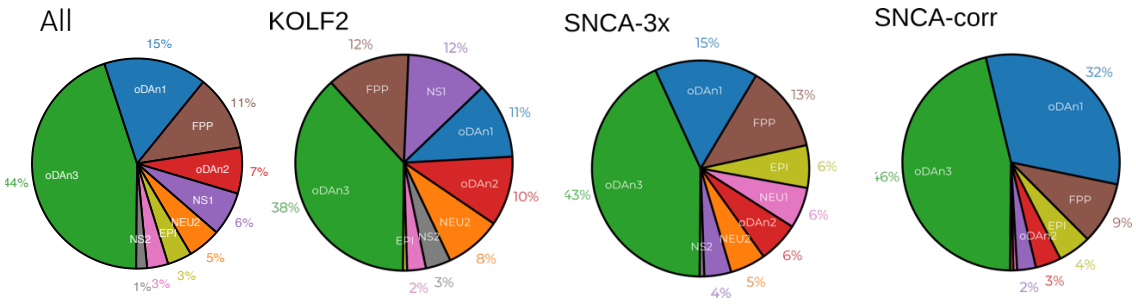



**Figure 3.13**: Pie charts illustrating the distribution of cell types within the unified dataset, encompassing all three cell lines, along with the breakdown of cell proportions specific to each individual cell line. **Note:** These plots were generated by Nikos Patikas.

Analysis of cell proportions revealed that within the HMOs, 66% of the composition consisted of dopaminergic neurons, 11% were floor plate progenitors, 7% constituted neural stem cells, 11% were non-dopaminergic neurons, and 3% represented ependymal cells (Figure 3.13). Examining the individual cell line HMOs, it was evident that most cell types were consistently present across all three HMOs, with the exceptions of EPI and NEU2(Figure 3.13). This observation highlighted the robustness of the differentiation protocol. EPI cells were identified in SNCA-3x and SNCA-corr HMOs but absent in KOLF2 HMOs (Figure 3.13). Conversely, NEU2 cells were found in KOLF2 and SNCA-3x HMOs but were notably absent in SNCA-corr HMOs (Figure 3.13). In terms of dopaminergic neuron proportions, SNCA-corr HMOs exhibited the highest percentage at 81%, followed by 64% in SNCA-3x HMOs, and 59% in KOLF2 HMOs (Figure 3.13).

### 3.2.10 HMOs integration with in-vitro datasets

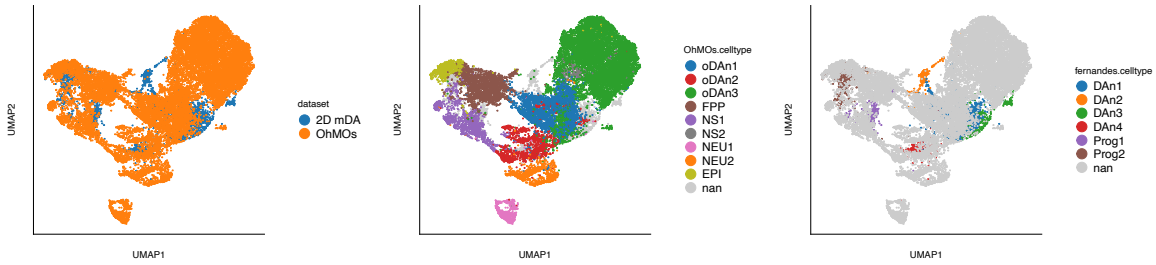



**Figure 3.13**: UMAP plots revealing the overlap between monolayer dopaminergic neuronal cultures (Fernandes et al., 2020) and HMO cultures. The Fernandes et al. (2020) dataset included 2645 single cells, featuring 4 distinct dopaminergic neuronal populations (DAn1, DAn2, DAn3, and DAn4), as well as 2 progenitor populations (Prog1 and Prog2). **Note:** These plots were generated by Nikos Patikas.

To comprehensively evaluate the similarities between the identified cell types within HMOs and those present in monolayer dopaminergic neuronal cultures (as previously described by Fernandes et al. (2020), we integrated scRNA-seq datasets from both culture systems using an autoencoder-based dimensionality reduction method (Lopez et al., 2018).

The results visualized through UMAP revealed a notable degree of overlap among the cell types between the two culture systems (Figure 3.13). A particularly significant finding was the overlap observed between oDAn1 in HMOs and DAn1 neurons in the monolayer cultures (Figure 3.13). Notably, monolayer DAn1 neurons have been characterized as bearing resemblance to *in vivo* substantia nigra dopaminergic neurons. Conversely, the presence of monolayer DAn2 and DAn3 neurons was not observed within the HMOs (Figure 3.13). It is worth mentioning that monolayer DAn1 and DAn3 neurons were reported to express serotonergic, GABAergic, and glutamatergic markers, respectively (Fernandez et al.,2020).

Interestingly, HMOs DAn2 population exhibited overlap with monolayer DAn4 neurons, which were shown to express markers indicative of ventral tegmental area (VTA) identity, such as *OTX2* and *CALB1*(Figure 3.13). Conversely, certain cell types within HMOs, specifically oDAn3, NEU1, NEU2, EPI, and NS2, were not observed within the monolayer dopaminergic neuronal culture (Figure 3.13).

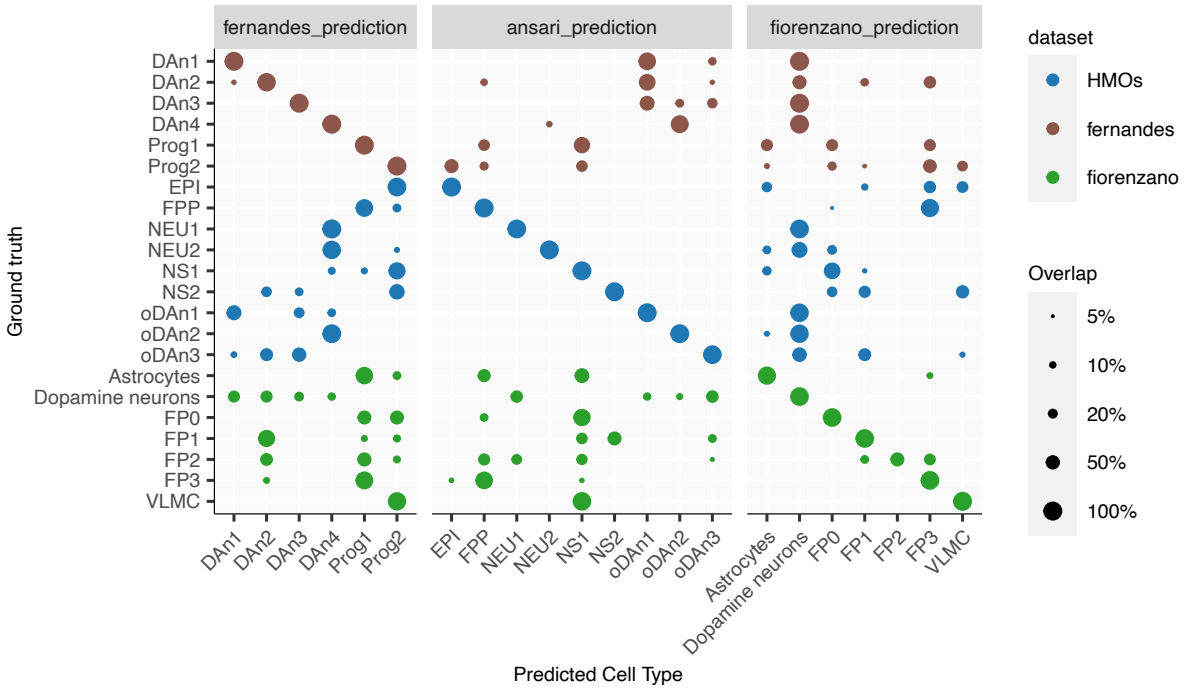



**Figure 3.14:** The matrix plot exhibits the similarities between actual cell types (ground truth) presented in rows and their corresponding predicted cell types (columns). Each plot panel corresponds to a distinct classifier trained to predict the label of an individual dataset. Points along the diagonal signify the accuracy of the classifier in predicting a cell's identity based on its training set. For evaluating the overlap between two cell types originating from different datasets, we examine both perspectives: (d1-ct1, d2-ct2) and the reciprocal overlap (d2-ct2, d1-ct1). This comprehensive approach enables the assessment of outcomes from separate and independent classifiers. **Note:** These plots were generated by Nikos Patikas.

To explore the resemblance of cell types across distinct *in vitro* models, a comprehensive investigation involving cell type label transfer was carried out among three separate *in vitro* models. These models included the previously discussed monolayer model by Fernandes, our midbrain organoids (HMO), and an established Day 60 midbrain organoid scRNA-seq dataset as presented by Fiorenzano et al. (Fiorenzano et al., 2021). To achieve this, the integration of these three datasets was performed, yielding the establishment of three distinct label transfer models. Each model was trained on the cell type annotations from its respective dataset and consequently projected these labels onto the other two *in vitro* datasets.

HMO's oDAn1 was found to exhibit a broad mapping that included monolayer DAn1, DAn2, and DAn4 populations (Figure 3.14). It's worth noting that Fiorenzano et al. annotated

a singular population as dopamine neurons, and intriguingly, this annotation was found to broadly map across all dopamine neuron populations within both HMOs and the monolayer culture (Figure 3.14). Furthermore, a strong mapping observed between Fiorenzano's astrocyte population and HMO's FPP and NS1 cell types (Figure 3.14). HMO's oDAn2 mapped with monolayer DAn4, thereby corroborating the findings derived from the UMAP dimensionality reduction analysis (Figure 3.14). Additionally, the cell types NEU1 and NEU2 within HMOs displayed an overlap with the monolayer DAn2 population (Figure 3.14).

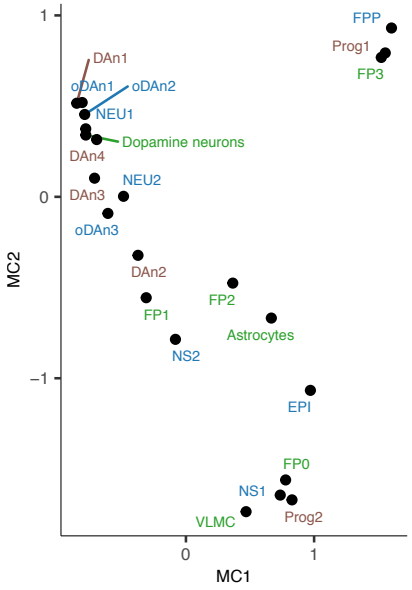



**Figure 3.15**: Multiple Correspondence Analysis (MCA) illustrating the relations among cell types through cell type label transfers. Utilizing the predicted cell types assigned to each cell types across various dataset annotations, a dimensionality reduction analysis was conducted that depicts the relationships of diverse cell types. Text annotations are color-coded according to their respective cell types and the dataset colors presented in figure 1.7. **Note:** This plot was generated by Nikos Patikas.

A Multiple Component Analysis (MCA) dimensionality reduction was performed to provide enhanced visualization of the interrelationships between cell types across diverse model classifications. This visualization technique facilitated the grouping of cell types co-existing within single cells across different classification schemes into closer proximity. HMO's oDAn1 and monolayer DAn1 were observed to be present in close proximity (Figure 3.15). Similarly, the close proximity of HMO's oDAn2 and monolayer DAn4 within the visualization pointed to their close association (Figure 3.15). Furthermore, the MCA visualization suggested relationships between various cell types. Specifically, HMO's FPP, monolayer Prog1, and

Fiorenzano's FP3 appeared to be related (Figure 3.15). Similarly, HMO's NS1, monolayer Prog2, and Fiorenzano's FP0 were also identified as being closely related (Figure 3.15). However, for the remaining cell types within the datasets, distinct boundaries were less defined, indicating a higher degree of variation between the *in vitro* models.

### 3.2.11 Comparing HMO Neuronal Cell Types to Human SNpc *In-Vivo* Neuronal Nuclei

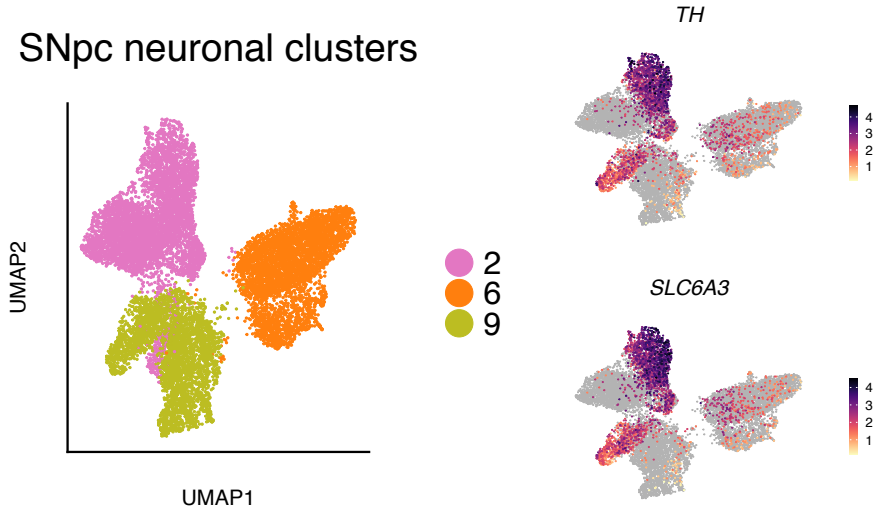



**Figure 3.16**: This UMAP plot illustrates the neuronal clusters derived from the clustering of *SYT1*-positive cells within the human post-mortem SNpc single nuclei dataset (Wang et al., 2022). Additionally, UMAP heatmaps present the log2-normalized expression of *TH* and *SLC6A3* within these neuronal clusters. **Note:** These plots were generated by Nikos Patikas.

To understand the resemblance of neuronal cell types within HMOs (NEU1, NEU2, oDAn1, oDAn2, and oDAn3) to the neurons typically affected in PD, a comparative analysis was undertaken with substantia nigra pars compacta (SN) tissue. To achieve this goal, the single cell dataset of HMOs was integrated with a single-nuclei dataset derived from a postmortem brain tissue sample of healthy aged individuals (Wang et al.,2022). For this analysis, solely neuronal nuclei expressing *SYT1* were selected from the *in vivo* dataset. Through clustering, the *SYT*1-expressing neuronal nuclei unveiled three distinct clusters, namely Cluster 2, 6, and 9(Figure 3.16). A subset of Cluster 2 and Cluster 9 exhibited notable levels of *TH* and *SLC6A3*(Figure 3.16).

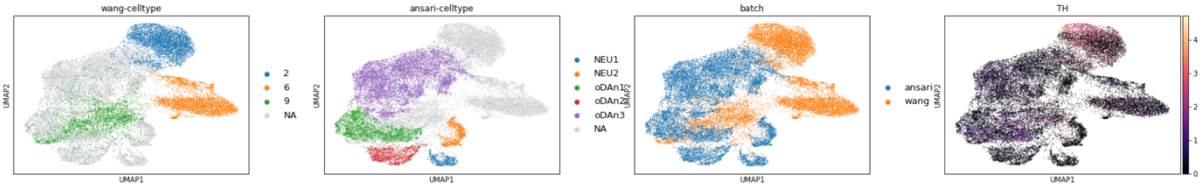



**Figure 3.17**: UMAP plots demonstrating the overlap of HMO cell types with SNpc neuronal clusters. **Note:** These plots were generated by Nikos Patikas.

The integration of HMOs neuronal cell types and SN neuronal clusters using UMAP dimensionality reduction showed overlap of HMO's oDAn1 to *in vivo* cluster 9(Figure 3.17). This overlap highlights the presence of similarities between these two clusters. Conversely, the remaining neuronal cell types within HMOs did not exhibit substantial overlap with any of the *in vivo* neuronal clusters (Figure 3.17). Furthermore, it's noteworthy that the high *TH*-expressing *in vivo* cluster failed to find a meaningful mapping to any of HMOs' neuronal cell types.

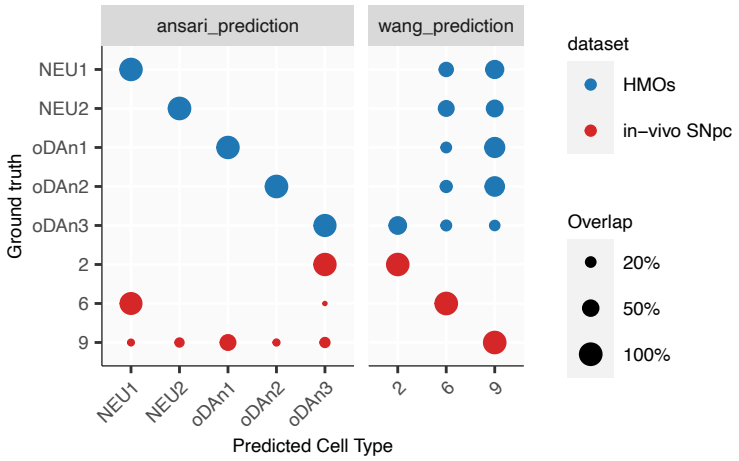



**Figure 3.18**: The matrix plot exhibits the similarities between actual cell types (ground truth) presented in rows and their corresponding predicted cell types (columns). Each plot panel corresponds to a distinct classifier trained to predict the label of an individual dataset. Points along the diagonal signify the accuracy of the classifier in predicting a cell's identity based on its training set. **Note:** This plot was generated by Nikos Patikas.

As performed on the *in vitro* datasets, we conducted a label transferring process (Xu et al., 2021) on the *in vivo* dataset and HMO's neuronal cell types, thereby facilitating the reciprocal transfer of labels between these two domains. The outcome of this analysis showed that in the context of label transfer from *in vivo* to HMOs, cluster 9 showed a comprehensive

mapping across all HMO's neuronal cell types, with a particularly robust overlap to oDAn1(Figure 3.18). While cluster 9 displayed a mapping to NEU1, the reverse analysis indicated that cluster 6 had resemblance to all HMO's neuronal cell types (Figure 3.18). In contrast, the mapping was distinctly clear in the case of cluster 2, which precisely matched with HMO's oDAn3 during the two-way analysis (Figure 3.18). It is important to highlight the significance of training the model using well characterized cell types, as the performance can otherwise suffer. We believe that the subpar performance observed in the label transfer from *in vivo* to HMOs may have been influenced by the heterogeneity present within the neuronal clusters in the *in vivo* dataset. This highlights the complexities presented in transferring labels across different domains and the need for a improved approach to ensure accurate and informative results.

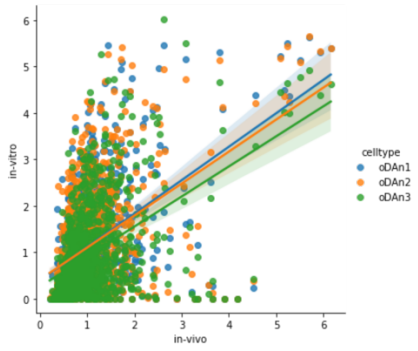



**Figure 3.19**: Gene correlation analysis between HMOs and in-vivo SNpc data. Only *TH*+ nuclei were extracted from the in-vivo dataset (totaling 2274 cells). In the plot, we focused on the top 1000 genes expressed within these nuclei, forming a gene correlation matrix between the same set of 1000 genes in our in-vitro cells. oDAn1 (1358 cells), oDAn2 (1358 cells), and oDAn3 (8587 cells) were tested against the *TH*+ positive *in vivo* group. R scores of this analysis were as follow: oDAn1 Pearson's R: 0.4655, oDAn2 Pearson's R: 0.4387, oDAn3 Pearson's R: 0.4590. Note: This plot was generated by Nikos Patikas.

To delve deeper into the assessment of similarities between HMOs and *in vivo* neurons, a focused analysis was conducted. *TH*-positive cells were extracted from the *in vivo* dataset, and the gene correlation profiles of the top 1000 highest expressing genes were plotted in comparison with HMO's dopamine neuronal cell types (oDAn1, oDAn2, and oDAn3).

The results of this analysis unveiled compelling correlations. Notably, oDAn1 emerged as displaying the highest correlation coefficient (Pearson's R: 0.465) among these cell types.

This correlation surpassed those observed in oDAn2 (Pearson's R: 0.438) and oDAn3 (Pearson's R: 0.459) (Figure 3.19). The results of this analysis consistent with the findings garnered through the UMAP dimensionality reduction analysis. Evidently, this analysis provides confidence that oDAn1 appear to be the most strongly correlated cell type with the *in vivo* dataset. This connection is especially emphasized by its superior correlation with *TH*-positive nuclei, contributing to a more comprehensive understanding of the dynamic relationships between HMOs and *in vivo* neurons.

### 3.2.12 Assessment of Dopaminergic Marker Genes Across In-Vivo and In-Vitro Datasets

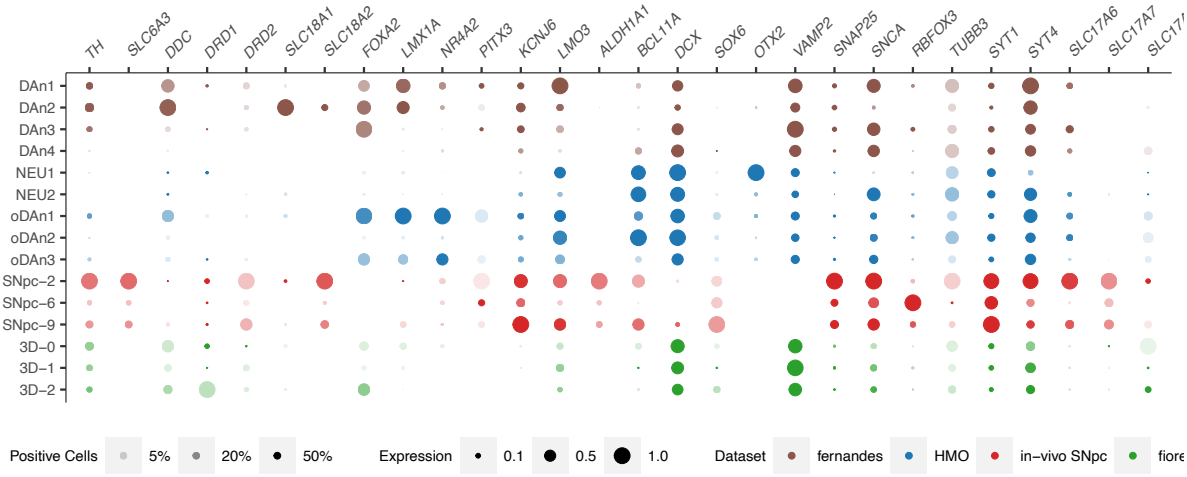



**Figure 3.20**: This matrix plot shows the expression of genes involved in dopaminergic neurons development, dopamine metabolism and functions within neuronal populations of HMOs (oDAn1, oDAn2, oDAn3, NEU1, NEU2), monolayer dopaminergic cultures (DAn1, DAn2, DAn3), Fiorenzano's midbrain organoids (3D-0, 3D-1, 3D-2), and *in vivo* datasets (SNpc-2, SNpc-6, SNpc-9). **Note:** These plots were generated by Nikos Patikas.

Single cell integration workflows provide valuable insights into the similarities in cell types between datasets. However, they fall short in offering insights at the individual gene expression level. To address this limitation, a focused gene expression analysis was conducted on dopaminergic (DA) neurons derived from various sources, including human midbrain organoids (HMOs), monolayer cultures, *in vivo* substantia nigra tissue, and the Fiorenzano dataset. In the case of the Fiorenzano DA neurons, they were categorized into three clusters for the purpose of this analysis.

The analysis concentrated on selected genes associated with dopaminergic neuronal identity and dopamine metabolism. Notably, dopaminergic neuronal markers such as *TH* and *SLC6A3* exhibited high expression levels in the *in vivo* cluster 2, while such expression was absent in the *in vitro* dopamine neurons derived from both 2D and 3D cell culture models (Figure 3.20). Additionally, the *in vitro* dopamine neurons failed to express key markers linked to dopamine metabolism, such as *SLC6A3* and *ALDH1A1*(Figure 3.20). Another marker, *SCL18A2*, displayed significant expression in *in vivo* DA neurons, in contrast to the limited expression observed in most *in vitro* DA neurons (Figure 3.20). The dopaminergic lineage markers *LMX1A* and *FOXA2* exhibited robust expression in HMO-derived DA neurons compared to *in vivo* DA neurons, suggesting a relative immaturity of the HMO-derived DA neurons(Figure 3.20).

### 3.2.8 Neuromelanin accumulation in Human Midbrain Organoids

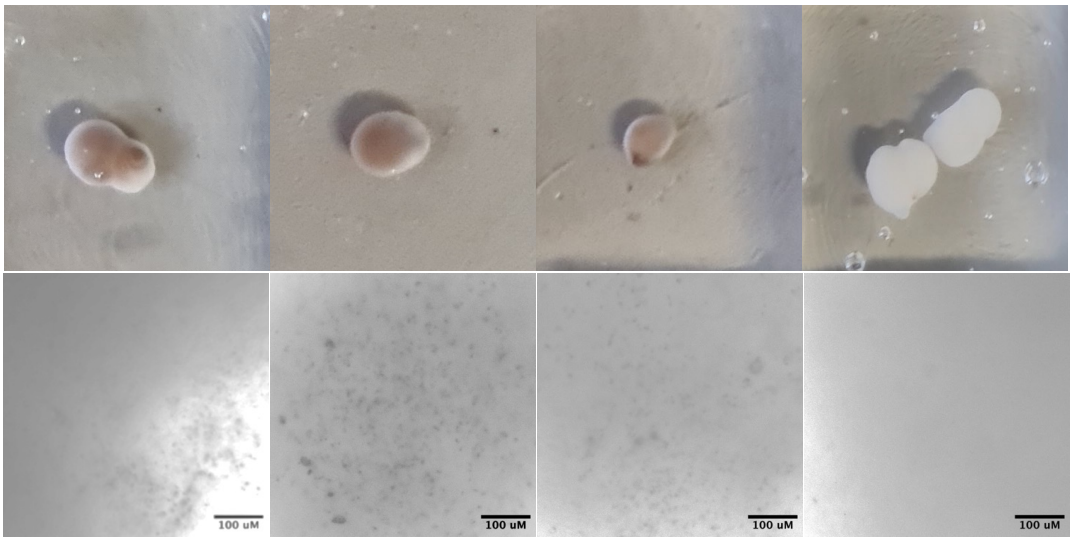



**Figure 3.21**: Brightfield images illustrating the colour transformation of KOLF2, SNCA-corr, SNCA-3x HMOs following L-DOPA treatment (100µM), accompanied by the emergence of neuromelanin-like pigmentation.

Neuromelanin, characterized by its insoluble black/brown granular appearance, accumulates notably within the substantia nigra pars compacta (SNpc) in both humans and primates. It arises as a by-product of dopamine synthesis within midbrain dopaminergic neurons. Intriguingly, studies have shown that the exposure of exogenous dopamine or L-

DOPA (a dopamine precursor) accelerates neuromelanin accumulation within in young midbrain organoids (Jo et al., 2016).

To delve deeper into the capacity of neurons within our HMOs to generate neuromelanin, a strategic approach was undertaken. Our HMOs were subjected to a 15-day treatment with maturation media supplemented with L-DOPA (100uM). Following the treatment, all L -DOPA treated HMOs appeared black/brown coloured to the naked eyes compared to the untreated HMOs (Figure 3.21). Subsequent examination under a brightfield microscope unveiled the presence of neuromelanin-like dark pigmentation exclusively within the treated HMOs (Figure 3.21). The appearance of neuromelanin indicate resemblance of HMOs with human midbrain.

### 3.2.13 Evaluating Ansari Protocol Reproducibility

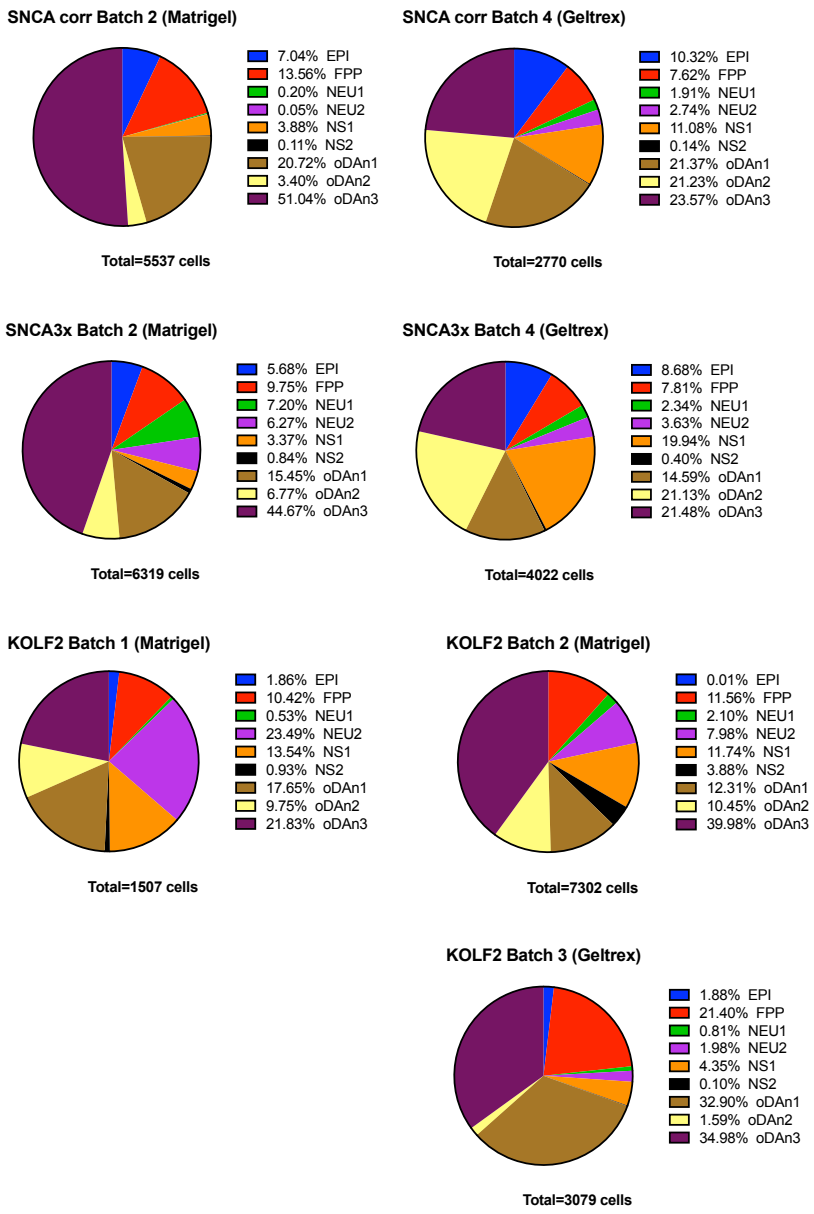



**Figure 3.22**: Pie chart illustrating the distribution of cell types in KOLF2, SNCA-corr, and SNCA-3x HMOs from various batches. **Note:** These plots were generated by Nikos Patikas.

To assess the reproducibility of the Ansari protocol, cellular heterogeneity within HMOs derived from multiple cell lines across different batches was analysed. The analysis showed the consistent presence of all major cell types across various batches. However, it's noteworthy that the percentages of these distinct cell types exhibited variations from one batch to another(Figure 3.22).

### 3.2.14 Comparing Cellular Heterogeneity and Gene Expression in HMOs Cultured with Matrigel and Geltrex

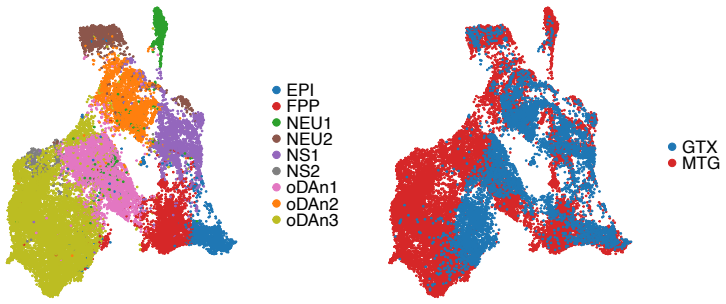



**Figure 3.22**: UMAP plots highlighting the overlap between Matrigel (MTG) and Geltrex HMOs (GTX). On the left, the UMAP plot displays the combined dataset of Matrigel and Geltrex HMOs, accompanied by cell type annotations. On the right, the UMAP plot displays cell annotations categorized by matrix type: Geltrex cells in blue and Matrigel cells in red. **Note:** These plots were generated by Nikos Patikas.

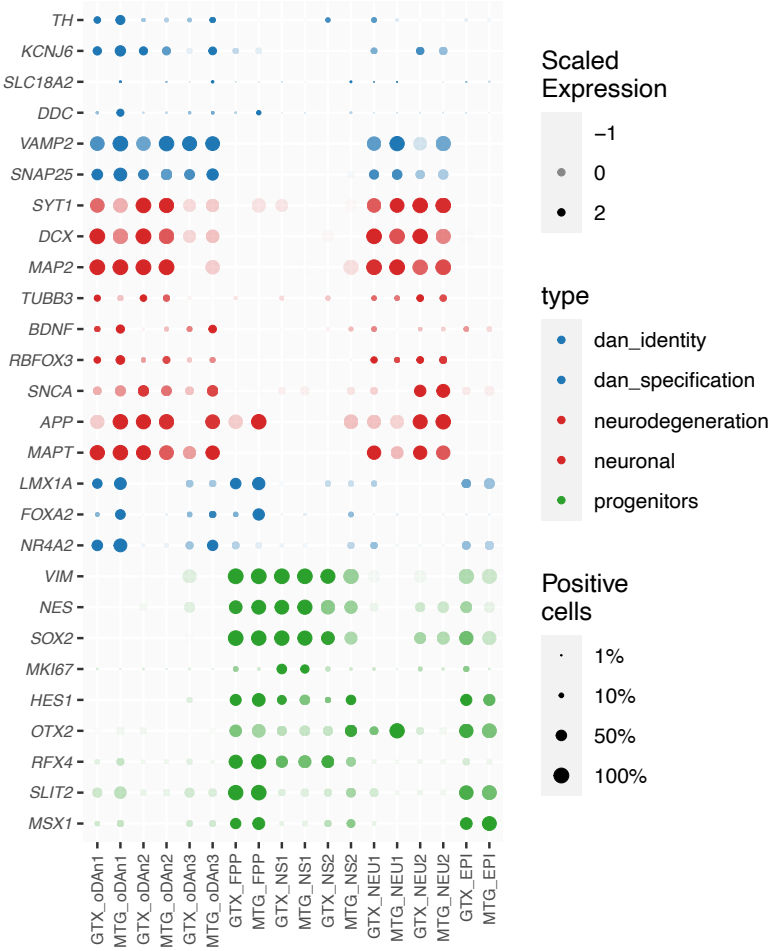
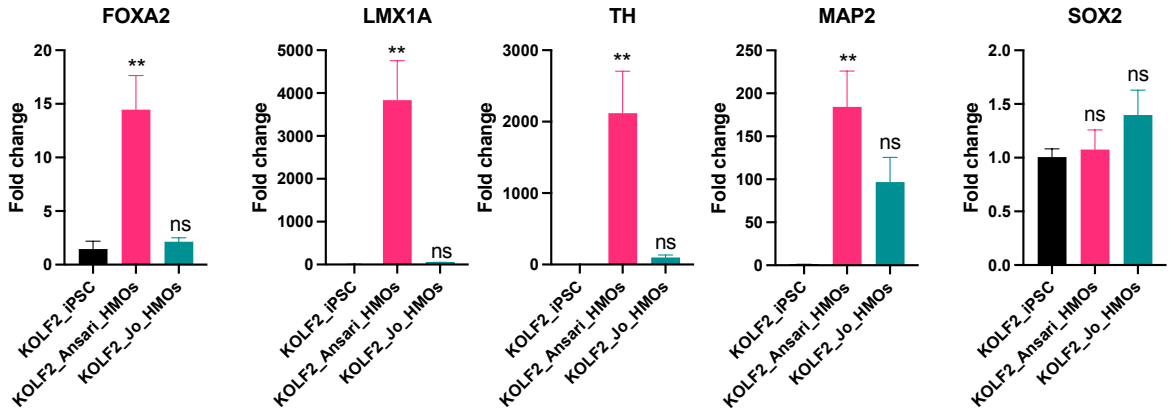


**Figure 3.23**: Matrix plot illustrating the gene expression of markers associated with various stages of dopaminergic neuron development across distinct cell types within Matrigel and Geltrex HMOs. **Note:** These plots were generated by Nikos Patikas.

To elucidate the difference between HMOs developed using Matrigel and Geltrex matrices in terms of cellular heterogeneity and gene expression levels, a comprehensive approach was employed. The scRNA-seq datasets from both culture conditions were integrated through autoencoder dimensionality reduction (Lopez et al., 2018), complemented by a label transfer workflow (Xu et al., 2021).

UMAP analysis showed a noteworthy overlap of cell types between the two cultures (Figure 3.22). Further insights emerged through cell proportion analysis, revealing a higher proportion of oDAn3 cells in Matrigel HMOs, accompanied by fewer oDAn2 cells, and a comparable count of oDAn1 cells (Figure 3.22). Gene expression analysis showed that Matrigel oDAn1 cells exhibited higher expression of the dopaminergic neuronal marker *FOXA2(*Figure 3.23). Notably, markers associated with neurodegenerative diseases such as *SNCA* and *APP* were significantly elevated in Matrigel oDAn1, oDAn2, and oDAn3 populations (Figure 3.23).

### 3.2.15 Comparison of Ansari and Unmodified Jo et al. 2016 Protocols



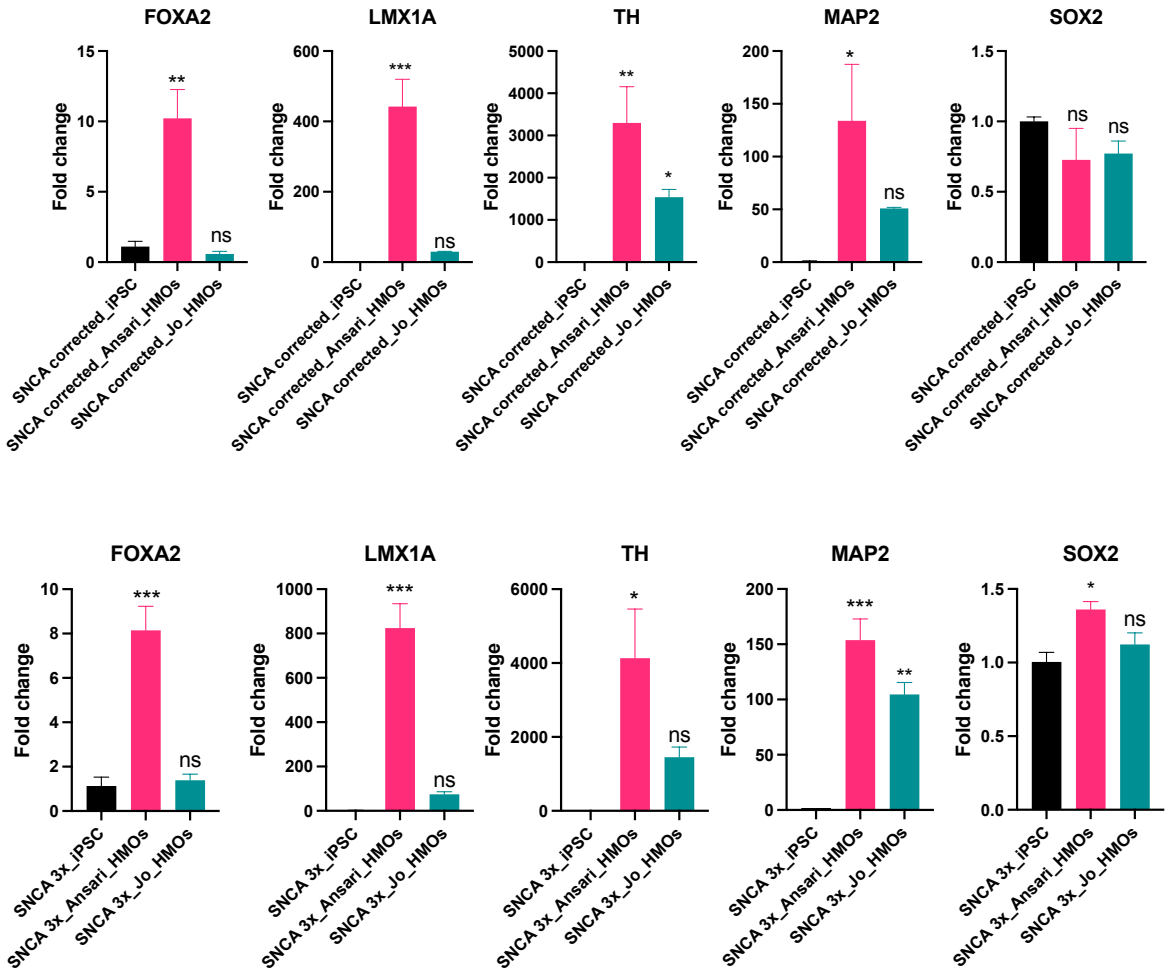



**Figure 3.24**: Bar plot depicting the qPCR analysis outcomes for the expression of dopaminergic markers (*LMX1A, FOXA2,* and *TH*), mature neuronal marker (*MAP2*), and stem cell marker (*SOX2*) in 4-week-old KOLF2, SNCA-corr, and SNCA-3x HMOs generated using both the Ansari protocol and unmodified Jo et al., 2016 protocol. The data are represented as ± SEM for n = 3 samples. Statistical analysis was conducted using one-way ANOVA with Dunnett's multiple comparisons test. Significance levels: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$, compared to the control (iPSCs).

The objective of this study was to compare HMOs generated through the optimized Ansari protocol with the original Jo et al., 2016 protocol (Jo protocol). The unmodified Jo protocol relies on the utilization of Noggin for SMAD inhibition and Shh protein to activate sonic hedgehog signaling. For this experiment, KOLF2, *SNCA*-3x, and *SNCA*-corr were utilized.

HMOs were cultured and harvested on day 35 for comprehensive gene expression assessment through qPCR analysis. The qPCR analysis showed that the HMOs generated via

the Ansari protocol exhibited significantly elevated levels of gene expression of critical dopaminergic markers, such as *LMX1A, FOXA2,* and *TH*, in addition to the mature neuronal marker, *MAP2*(Figure 3.24). In contrast, the HMOs developed through the Jo et al. protocol exhibited comparatively lower levels of expression for these pivotal markers (Figure 3.24).

## 3.3 Discussion and Future Directions

Here, we introduce an innovative protocol for generating human midbrain organoids (HMOs) from hiPSCs. This protocol has demonstrated both reproducibility and efficiency in producing midbrain organoids from both healthy and PD mutant hiPSC lines. Through immunohistochemical analysis, we observed the presence of midbrain progenitors, dopaminergic neurons, and neural stem cells within the midbrain organoids, meticulously organized in an *in vivo* like tissue-specific manner. To delve deeper into the cellular diversity and gene expression patterns of these cell types within the midbrain organoids, we conducted single-cell RNA sequencing (scRNA-seq). This analysis unveiled multiple subpopulations of dopaminergic neurons, including a subpopulation closely resembling *in vivo* human substantia nigra neurons. The remarkable resemblance of midbrain organoids to human midbrain cellular heterogeneity and tissue architecture enables us to model aspects of PD that cannot be accurately evaluated in 2D models.

Several protocols for deriving human midbrain organoids are available in the literature (Galet et al., 2020). These protocols primarily differ in the concentration of specific molecules and the timing used to direct cells toward a midbrain floor plate fate. The derivation of midbrain organoids can be broken down into four stages: 1) Neural induction through dual SMAD inhibition and Wnt modulation. 2) Midbrain floor plate specification using SHH and FGF8 modulation. 3) Tissue induction. 4)Organoid development and maturation.

Some protocols employ a simultaneous patterning approach, where neural inducers and midbrain floor plate patterning factors are applied simultaneously. Matrigel is commonly used as a matrix for tissue induction, facilitating tissue growth and organization. To enhance nutrient and gas exchange throughout the organoids, researchers have utilized orbital shakers, spinner flasks, and microfluidic systems.

The protocol described by Jo et al. in 2016 utilized a sequential patterning approach. Initially, human embryonic stem cells (hESCs) were induced toward a neuroectoderm fate using SB431542 and Noggin. Midbrain patterning was then achieved using SHH-C25II and FGF8, while Wnt signaling was activated using CHIR99021 starting from Day 0. We adapted the Jo et al. 2016 protocol with some modifications to derive midbrain organoids from human iPSCs. Notably, we replaced Noggin and SHH-C25II with LDN193189 and SAG, respectively. Characterization of the midbrain organoids generated using this modified protocol revealed a low number of LMX1A+/FOXA2+ positive cells. To address this misspecification, we conducted titration experiments with small molecules CHIR-99021 (a Wnt signaling activator) and SAG (a Shh signaling agonist). Improved conditions were identified that enhanced FOXA2 and LMX1A expression, though tissue cytoarchitecture remained lacking.

To address these limitations, we developed a protocol named "Ansari" by combining the widely accepted Kriks et al. 2011 midbrain floor plate patterning strategy with the Jo et al. 2016 tissue induction and organoid development procedures. In this approach, neural induction and midbrain floor patterning occurred simultaneously. Dual SMAD inhibition was achieved using SB431542 and LDN193189, and Shh signaling was activated with purmorphamine and SAG. Unlike the original Jo et al. 2016 protocol, Wnt signaling was activated on day 3, and the midbrain floor plate patterning stage was prolonged compared to the Jo et al. protocol.

### 3.3.1 Neurospheres Generation Using Kriks et al. Protocol

Previously, researchers generated midbrain-like neurospheres based on the Kriks et al. protocol (Tieng et al., 2014). These neurospheres contained A9 dopaminergic neurons but lacked tissue organization. The absence of tissue organization in these midbrain-like neurospheres may have result de to lack of a tissue induction step using hydrogel in the protocol. In our combined protocol, we introduced a tissue induction step using Matrigel/Geltrex, resulting in the production of dopaminergic neurons, neural stem cells, and midbrain floor plate progenitor cells organized in an *in vivo* tissue-like manner.

### 3.3.2 Absence of Glial Cells (Astrocytes & Oligodendrocytes) In Midbrain Organoids

Like any *in vitro* cell culture model, human midbrain organoid models have their limitations and challenges. In our day 49 midbrain organoids, single transcriptomic analysis did not reveal enrichment for glial cell markers, indicating the absence of glial cells. The lack of glial cells in midbrain organoid model represents a significant limitation of the model, as these cells have implications in PD pathogenesis. However, the absence of glial cells in the midbrain at the day 49 time point is consistent with other published midbrain organoid models. Kwak et al. and Fiorenzano et al.'s midbrain organoid models demonstrate the appearance of glial cells (astrocytes and oligodendrocytes) post day 60. These findings align with the developmental timeline of the midbrain, where neurogenesis precedes gliogenesis. Ideally, midbrain organoid models should be utilized in experiments at later time points when glial cells have had the opportunity to form. Unfortunately, there is a drawback, as the dopaminergic neuronal population diminishes over time, as shown by Fiorenzano et al. Their research indicates that the dopaminergic neuronal population peaks around day 60, coinciding with the appearance of glial cells, but significantly decreases by day 120, comprising less than 5% of the total cell population. Therefore, further optimization of midbrain organoid protocols is necessary to address these challenges effectively.

### 3.3.3 Absence of Microglia in Midbrain Organoids

The midbrain organoid model also lacks microglia, which represents a significant limitation since activated microglia have been linked to the demise of dopaminergic neurons in PD. Microglia in the brain perform various crucial functions, including pruning neuronal synapses and phagocytosing apoptotic cells and debris (Anderson et al., 2022; Lewis et al., 2011). The absence of microglia in midbrain organoid models is expected because neuronal cells originate from the ectodermal lineage, whereas *in vivo*, microglia cells come from the mesoderm lineage. Midbrain organoid differentiation tends to favour the ectoderm lineage while suppressing the mesoderm and endoderm lineages. Therefore, it is not surprising that microglia cells are absent in midbrain organoid models due to their non-neuroectodermal origin (Zhang et al., 2022).

To address this limitation, researchers have made efforts to integrate exogenously generated microglia into midbrain organoids. Recently, Sabate-Soler et al. described the stable integration of iPSC-derived microglia into midbrain organoids (Sabate-Soler et al., 2023). They

demonstrated that integrated microglia in midbrain organoids expressed genes associated with phagocytosis and released chemokines and cytokines. Additionally, they showed that midbrain organoid models with integrated microglia exhibited reduced expression of genes associated with oxidative stress and cell death. Neurons in the midbrain organoid-microglia co-culture model displayed enhanced electrophysiological properties. This integrated midbrain organoid-microglia co-culture model holds promise as a more suitable platform for studying brain development, neuroinflammation, and neurodegenerative diseases.

### 3.3.4 Necrotic Core in Midbrain Organoids

One of the foremost challenges associated with the brain organoid model is the presence of a necrotic core. This issue arises due to the limited diffusion of nutrients and gases within the organoids and the absence of microvascularization within these brain organoids (Lancaster et al., 2018; Mansour et al., 2018). The existence of this necrotic core imposes constraints on both the size and long-term viability of the brain organoids. Moreover, it significantly impacts processes such as neuronal differentiation, maturation, and connectivity. Studies have revealed that brain organoids exhibit an increased reliance on glycolysis and undergo elevated stress in the endoplasmic reticulum (ER) and mitochondria (Galat et al., 2020; Bhaduri et al., 2020). Several techniques, including bioreactors, orbital shakers, and microfluidic systems, have been employed in attempts to alleviate the formation of the necrotic core. However, these approaches have encountered limited success.

### 3.3.5 Vascularization of Midbrain Organoids

An alternative strategy to address the issue of the necrotic core involves co-culturing brain organoids with endothelial cells, which has been shown to enhance the functional maturation of neurons (Missy T et al., 2018; Cakir et al., 2019). Missy T et al. achieved this by generating co-cultured iPSC-derived ECs and embedding them in Matrigel along with the brain organoid on day 34. This Matrigel/ECs embedding led to remarkable vascularization of the brain organoids within 3-5 weeks (Missy T et al., 2018).

In a different approach, Cakir et al. induced differentiation of brain organoids by adding 20% iPSCs carrying an inducible ETV2 transcription factor. The ectopic expression of ETV2 induced iPSCs to differentiate into ECs, resulting in the formation of vascular-like structures within the brain organoids (Cakir et al., 2019).

Another promising avenue is achieving vascularization within brain organoids through *in vivo* transplantation into rodents (Lancaster et al., 2018; Mansour et al., 2018). *In vivo* vascularized brain organoids exhibit superior neuronal differentiation and maturation, gliogenesis, microglia integration, and axonal growth into various areas of the brain (Mansour et al., 2018). Unfortunately, the scalability and affordability of this approach currently make it impractical for widespread use.

### 3.3.6 Midbrain Organoids on Air-Liquid Interface (ALI)

Taking a fundamentally different approach, researchers have adapted the organotypic slice culture method at the air-liquid interface (ALI) to cultivate brain organoids (Giandomenico et al., 2019). These brain organoids cultured at the air-liquid interface have shown enhanced neuronal survival and axon outgrowth.

While the mentioned methods aim to improve nutrient and oxygen delivery, it is important to note that neither of them is currently as easily expandable as the conventional organoid culture method, which remains the primary approach in organoid research. Consequently, the field still faces ongoing challenges related to tissue health and cellular stress until a universally applicable and easily scalable experimental solution becomes available.

### 3.3.7 Identification and Removal of Stressed Cell Populations In Brain Organoids

Vertesy et al. conducted an extensive analysis of a large single-cell RNA sequencing dataset involving brain organoids and fetal tissue (Vertesy et al., 2022). In their study, they identified a distinct subpopulation of stressed cells within brain organoids characterized by a unique gene expression profile. These stressed cells exhibited the upregulation of genes associated with endoplasmic reticulum stress, hypoxia, and glycolysis. Notably, terms related to "neurogenesis" and "gliogenesis" were notably absent in this stressed subpopulation. The authors hypothesized that these cells might be in a metabolic survival mode, lacking specific differentiation into neurons or glial cells. Furthermore, the stressed cell population displayed low expression levels of pan-neuronal markers such as *MAP2, DCX,* and *NCAM1*. Importantly, the presence of this subpopulation did not seem to affect the neuronal specification or maturation of other non-stressed neurons. To address this issue, the authors introduced a computational algorithm called Gruffi (Granular functional filtering) to identify and remove

stressed cells from brain organoid samples. This innovative method holds promise for future studies involving brain organoid cultures.

### 3.3.8 oDAn3 Dopaminergic Neuronal Population in Midbrain Organoid Model

In our midbrain organoid model, we identified three distinct dopamine (DA) neuronal populations. Of particular interest were the oDAn3 cells, which exhibited expression of genes associated with the unfolded protein response, including HSPA1A, *HSPA1B, DNAJB1,* and *HSPA6*. While these oDAn3 cells displayed robust expression of key dopaminergic markers such as *TH, SNCA, NURR1,* and *SNAP25*, their expression of neuronal functional markers, such as *SYT1, DCX,* and *MAP2*, was relatively low. Furthermore, when subjected to Pearson's analysis, comparing the gene correlation of the top 1000 highest expressing genes between these oDAn3 cells and *in vivo* substantia nigra TH+ positive neurons, they scored the lowest. Interestingly, these oDAn3 cells did not correspond to any cell types found in monolayer midbrain dopaminergic neuronal cultures. Collectively, these findings suggest that oDAn3 cells may represent a unique stressed population specific to our midbrain organoid model.

### 3.3.9 Mechanical Cutting of Brain Organoids

A recent study has demonstrated a simple yet effective method for addressing the issue of necrotic cores in brain organoids. By mechanically cutting the brain organoids into multiple pieces using a scalpel blade, researchers achieved almost complete reduction of the necrotic core (Choe et al., 2021). Upon immunohistochemical analysis, the cut brain organoids exhibited a similar cellular composition and tissue organization compared to their uncut counterparts. Importantly, the cut brain organoids did not display a central region expressing hypoxia-inducible factor (HIF)-1α or positive staining with terminal deoxynucleotidyl transferase dUTP nick end-labeling (TUNEL), both of which are indicative of cellular stress and apoptosis. Furthermore, the researchers observed reduced expression of the pro-apoptotic protein *BAX* and cleaved caspase 3 in the cut brain organoids when compared to the uncut ones. These findings suggest that the mechanical cutting approach could serve as a feasible addition to the routine culture of brain organoids, significantly improving their quality and addressing the issue of necrotic cores.

### 3.3.10 Maturity of Midbrain Organoid Model

In our study, we conducted a focused investigation into the expression of genes associated with dopaminergic identity and metabolism in dopaminergic neurons derived from our midbrain organoids. We compared the findings with aged human post-mortem substantia nigra samples, Fiorenzano's midbrain organoids, and monolayer dopaminergic neuronal cultures. The results revealed some notable differences.

Specifically, our midbrain organoid dopaminergic neurons exhibited relatively lower expression of *TH*, and there was low to no expression of *SLC6A3* and *SLC18A2*, which are associated with dopaminergic neurotransmitter transport. On the other hand, we observed higher expression levels of *FOXA2* and *LMX1A*, genes associated with early dopaminergic neuron development, in our midbrain organoid dopaminergic neurons when compared to *in vivo* substantia nigra neurons. Similar trends in gene expression were observed in dopaminergic neurons from other *in vitro* cultures, although Fiorenzano's dopaminergic neurons expressed *FOXA2* and *LMX1A* at slightly lower levels than our midbrain organoids.

It is worth noting that our comparative single-cell RNA sequencing (scRNA-seq) analysis did not detect *SLC6A3* expression in Fiorenzano's dopaminergic neurons. However, Fiorenzano et al. reported DAT (dopamine transporter) expression through immunohistochemical analysis at a similar timepoint. This suggests that the lack of expression of these mature markers in our scRNA-seq analysis could be a technical artifact. Further investigations using immunohistochemical analyses are warranted to clarify the expression of these genes in our midbrain organoid model.

Our midbrain organoids dopaminergic neurons expressed high levels of *FOXA2* and *LMX1A*, suggesting immature identity. This is consistent with the findings of other published midbrain organoids models. Jo et al. midbrain organoids model shows 30% *FOXA2* expressing cells at day 45. In general, brain organoids models have been shown to resemble human fetal brain (Lancaster et al., 2016; Jo et al., 2016; Nascimento et al., 2019). The study of disease associated aging mechanisms have been attempted using iPSCs derived models using pharmacological inhibition of telomerase, and overexpression of Progerin (Vera et al., 2016; Miller et al., 2013). The aged models have shown to display hallmarks of aging such as expression of γH2AX (marker of DNA damage) and increased oxidative stress. Further

advancement of brain organoid model is necessary to accurately model aging associated neurodegenerative diseases.

# 4. PD DISEASE MODELLING

## 4.1 Background

Various genetic mutations have been associated with familial Parkinson's disease (PD), including mutations in GBA, LRRK2, PINK1, PARK7, and SNCA. The SNCA gene encodes alpha-synuclein, a protein integral to processes like vesicle priming, docking, fusion, trafficking, neurotransmitter release, and axonal transport. However, its specific role in PD pathogenesis remains unclear. Notably, alpha-synuclein is a component of Lewy bodies, a critical pathological feature of PD. Elevated SNCA gene copy numbers are linked to early-onset PD. The revelation of mutations leading to extra SNCA copies is especially intriguing, as it implies that an excessive amount of normal alpha-synuclein protein could be a driving factor in PD. Research has shown that excess alpha-synuclein expression contributes to nigral region degeneration and motor impairments. Numerous studies utilizing animal and iPSC-derived neuronal models have identified mitochondrial dysfunction associated with alpha-synuclein mutations.

Rotenone, a pesticide linked to PD and mitochondria complex I inhibitor, is employed to emulate PD characteristics. Animal models exposed to rotenone replicate PD traits like nigral neurodegeneration and motor deficits. *In vitro* research has also highlighted the heightened vulnerability of dopaminergic neurons to rotenone. The precise mechanism by which increased ROS trigger dopaminergic neuronal degeneration remains unclear.

By delving into the impact of SNCA triplication mutation and oxidative stress on dopaminergic neurons, it may provide significant insights into the roles these factors play in PD pathology. This exploration serves as a platform for investigating how excess alpha-synuclein and oxidative stress contribute to neurodegeneration, with the potential to steer the development of targeted PD therapies.

For this study, HMOs were generated using SNCA triplication patient derived iPSCs (4 copies of SNCA gene) and its isogenic control (2 copies of SNCA gene), which allowed us to

study effect of excess alpha-synuclein in a controlled manner. To induce oxidative stress, HMOs were treated with media without any antioxidant supplemented with rotenone.

First, we evaluated the impact of SNCA triplication on various dopaminergic neuron subtypes. This assessment revealed associations between increased SNCA expression and crucial processes like translation, oxidative phosphorylation, and ER-stress.

Furthermore, our in-silico analysis of dopaminergic neurons pinpointed several processes relevant to the sensitivity and resilience of these neurons in response to rotenone induce oxidative stress. These processes encompassed synaptic signalling, cholesterol biosynthesis, and other vital factors. Moreover, we identified genes that are likely influenced by rotenone in dopaminergic neurons. Subsequently, we examined treated dopaminergic neurons and observed the upregulation of the unfolded protein response (UPR), alongside the downregulation of genes linked to cholesterol biosynthesis and mitochondria DNA-encoded genes. These findings shed light on the intricate molecular changes occurring within dopaminergic neurons following rotenone treatment.

In summary, our study employed in depth single cell transcriptomics to comprehensively investigate the composition and behaviour of cell types within HMOs, highlighting the significance of dopaminergic neurons and their responses to SNCA triplication and external factors like rotenone.

Numerous *in vitro* and *in vivo* studies have consistently reported the transfer of alpha-synuclein between cells and its prion-like propagation. Once alpha-synuclein aggregates form within neurons, they undergo intra-axonal transportation and are subsequently released via exosomes to various brain regions. These liberated alpha-synuclein aggregates can be taken up by neighbouring neurons through endocytosis, thereby triggering a cascade of further aggregation of the native alpha-synuclein. This intricate process suggests that the accumulation of alpha-synuclein within a limited cell population may progressively lead to the pathological engagement of diverse brain regions over time.

These discoveries have added a new dimension to therapeutic research for Parkinson's disease (PD). However, the intricate mechanisms underlying the prion-like spread of alpha-

synuclein necessitate comprehensive elucidation before they can be leveraged to develop targeted therapeutic interventions.

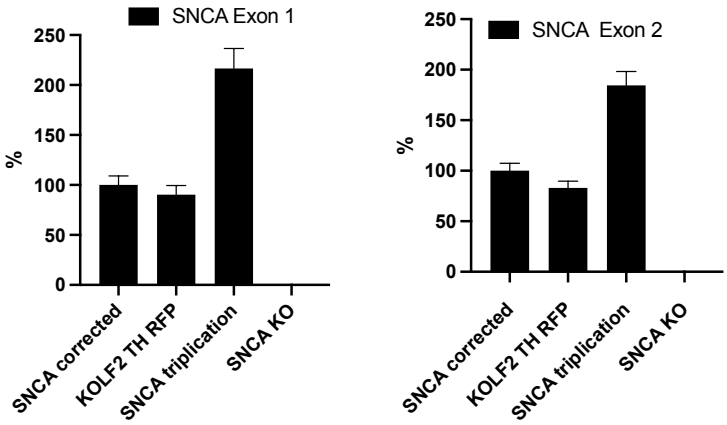
To model the non-cell-autonomous effects on dopaminergic neurons, we generated chimeric organoids comprising both healthy and disease-derived cells. This was achieved by co-culturing iPSCs from multiple donors. The utilization of chimeric organoids presents several advantages over monolayer dopaminergic neuronal cultures. The incorporation of cells from multiple donors within a single tissue diminishes technical variability, lowers experimental costs, and facilitates scalability.

These chimeric organoids underwent assessment through scRNA-seq, and the cell identities were determined via an in-silico approach as outlined by Heaton et al., 2020. The analysis revealed that when healthy dopaminergic neurons were cultured alongside SNCA-3x cells, they did not entirely replicate the mutational signatures of SNCA-3x. Nevertheless, the co-cultivation of these healthy neurons with SNCA-3x cells exhibited indications of deleterious effects on the former.

## 4.2 Results

### 4.2.1 Elevated SNCA Expression and α-Synuclein Protein Observed in SNCA-3x Organoids

(A)



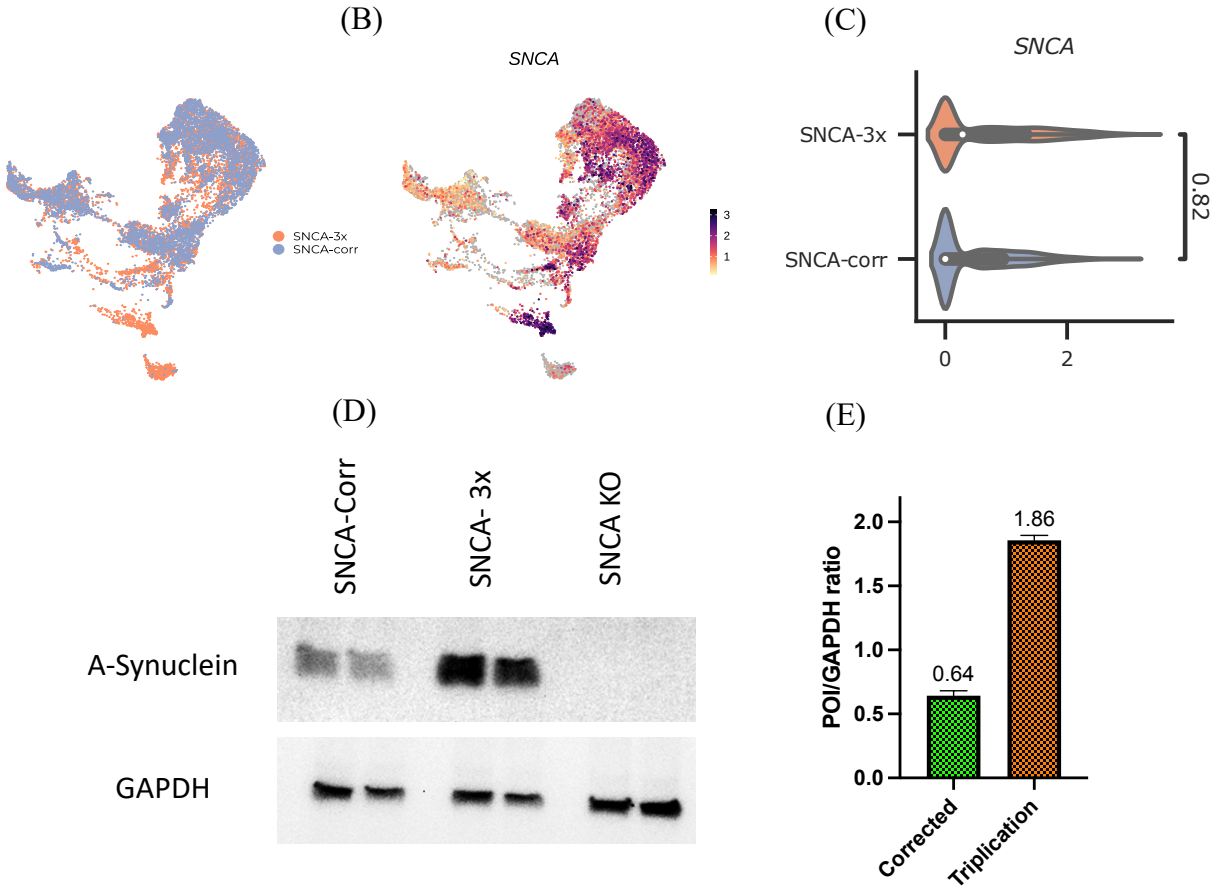



**Figure 4.1**: (A) Determination of SNCA gene dosage through qPCR analysis of SNCA EXON 1 and 2 in healthy KOLF2, SNCA-3x, SNCA-corr, and SNCA-KO (knockout) iPSCs. This procedure was conducted according to Devine et al. 2011, employing β-globin (HBB) as the diploid control gene. (B) UMAP heatmap displaying the log-normalized expression of the SNCA gene in SNCA-3x and SNCA-corr HMOs. (C) SNCA violin plots presenting the cultural differences between SNCA-3x and SNCA-corr. The x-axis represents log2-normalized expression, and annotations denote log2 fold change. (D) Western blot analysis of alpha-synuclein protein expression in SNCA-corr, SNCA-3x, and SNCA-KO HMOs. (E) Quantification of alpha-synuclein protein expression, obtained from western blot analysis in (D). The relative expression of alpha-synuclein was calculated by dividing the arbitrary blot density values of the protein of interest (POI) by GAPDH protein for each HMO sample. **Note:** Plots in B and C were generated by Nikos Patikas.

The human midbrain organoids (HMOs) derived from iPSCs of PD patients bear an intriguing genetic alteration: a triplication mutation of the SNCA gene (SNCA-3x), which is linked to elevated expression and aggregation of alpha-synuclein, contributing to early onset PD. To unravel the ramifications of the SNCA triplication mutation, SNCA-3x HMOs were

meticulously contrasted with their corresponding isogenic controls, wherein two out of the four SNCA gene copies were knocked out. Genotyping using quantitative PCR-based methods confirmed the presence of additional gene copies in SNCA-3x HMOs relative to the control group, comprising healthy KOLF2 HMOs and SNCA knockout HMOs. Analysis using single-cell RNA sequencing (scRNA-seq) unveiled differential overexpression of SNCA in SNCA-3x HMOs. Moreover, western blot analysis definitively verified a substantial increase (2.88-fold change) in the expression of the alpha-synuclein protein in SNCA-3x HMOs compared to the control group.

### 4.2.2 SNCA-3x impacts Endoplasmic Reticulum Protein Folding, Translation, and Oxidative Phosphorylation in oDAn1 Cells

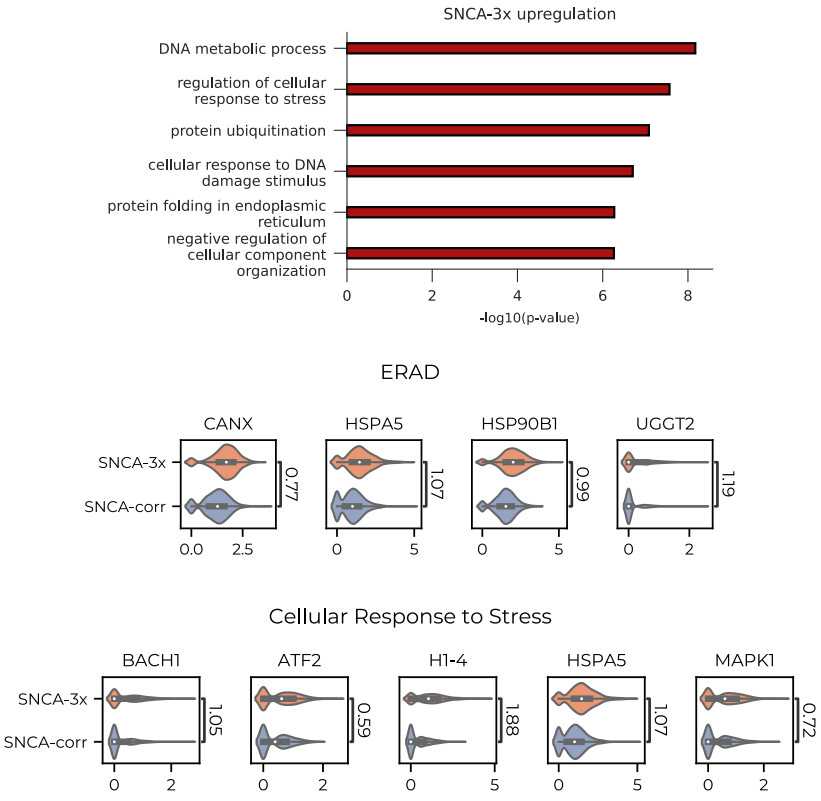



**Figure 4.2**: Bar plot illustrates enriched pathways corresponding to the top 200 differentially upregulated genes in SNCA-3x oDAn1 cells, contrasted with SNCA-corr oDAn1 cells. The x-axis indicates -log10(P-value). Violin plots present the log2-normalized expression levels of genes linked to these enriched pathways. Annotations between SNCA-3x and SNCA-corr conditions refers to log2 fold change. Note: These plots were generated by Nikos Patikas.

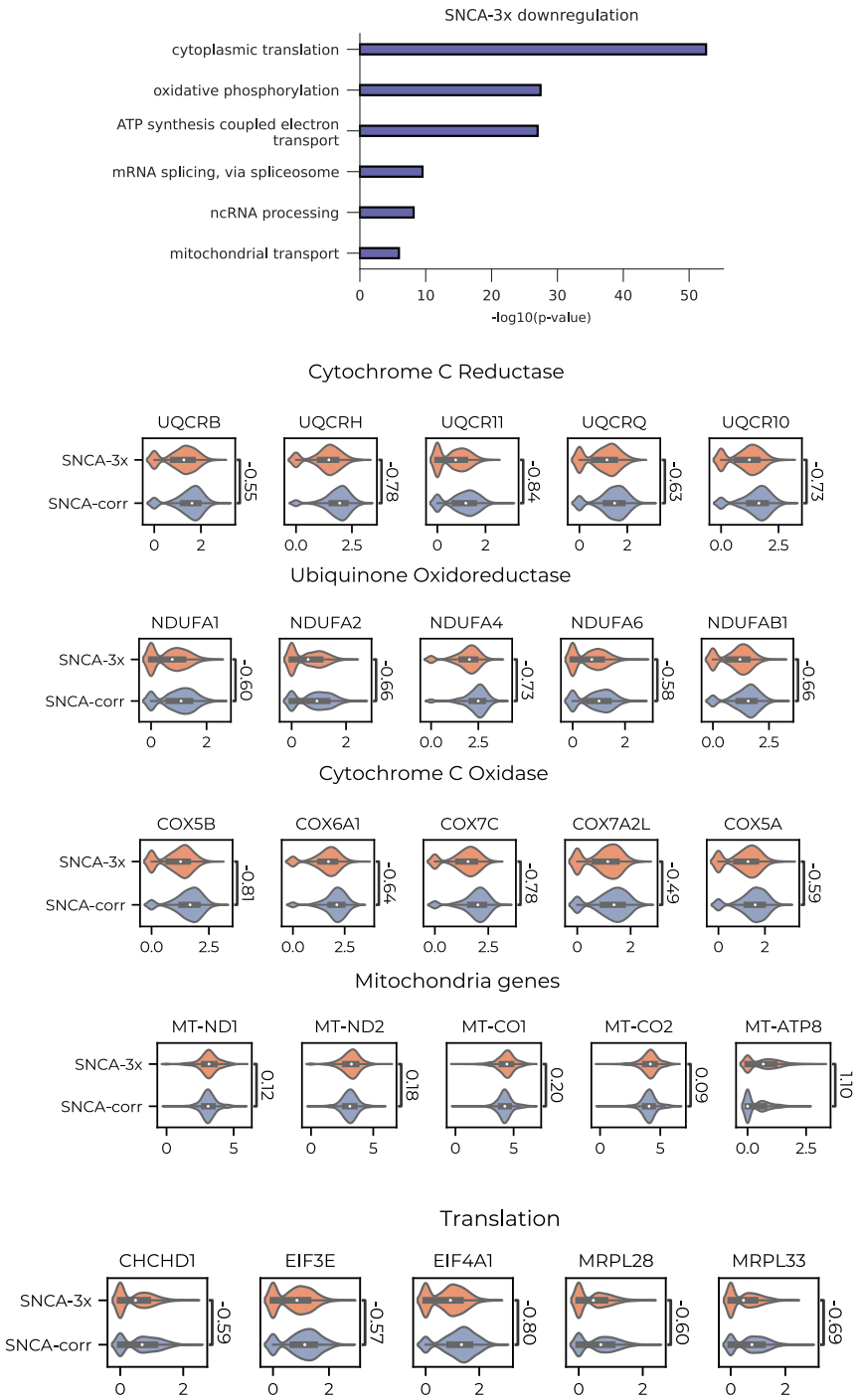



**Figure 4.3**: Bar plot illustrates enriched pathways corresponding to the top 200 differentially downregulated genes in SNCA-3x oDAn1 cells, contrasted with SNCA-corr oDAn1 cells. The x-axis indicates -log10(P-value). Violin plots present the log2-normalized expression levels of genes linked to these enriched pathways. Annotations between SNCA-3x and SNCA-corr conditions refers to log2 fold change. **Note:** These plots were generated by Nikos Patikas.

To investigate the implications of SNCA triplication phenotypes on dopaminergic (DA) neurons, we selected oDAn1 cells due to their expression of several markers linked to DA neuron identity and their resemblance to both *in vivo* and *in vitro* DA neurons, particularly *in vivo* SN DA neurons. SNCA-3x oDAn1 cells exhibited notably elevated expression of SNCA.

Differential gene expression analysis was carried out, uncovering numerous upregulated and downregulated genes in SNCA-3x DA neurons. Pathway overrepresentation analysis of these differentially expressed genes (DEGs) highlighted several affected pathways. Specifically, SNCA-3x oDAn1 cells displayed an upregulation of genes associated with protein folding in the endoplasmic reticulum (e.g., HSPA5, HSP90B1, CANX, and PDIA3), cellular response to stress (e.g., UBE2V1, DDX39B, PARG, and KDM1A), and response to oxidative stress (e.g., ATP2A2, FER, MAPK1, HDAC6, and SIRT1). Conversely, the downregulated genes were enriched for translation-related processes (e.g., CHCHD1, RPL34, RPL35, RPL23, and RPL27) and oxidative phosphorylation (OXPHOS) (e.g., NDUFB3, COX5B, UQCRH, MT-ND1, and MT-ND6).

Furthermore, essential protein complexes implicated in oxidative phosphorylation, including ubiquinone oxidoreductase (e.g., NDUFA2, NDUFA4, and NDUFA6), cytochrome C oxidase (e.g., COX5B, COX6A1, COX7C, COX5A), and cytochrome C reductase (e.g., UQCRB, UQCRH, UQCRQ, and UQCR10), demonstrated downregulation in SNCA-3x oDAn1 neurons compared to the control. Intriguingly, only mitochondria-nuclear-encoded genes exhibited downregulation, whereas mitochondria DNA-encoded genes such as MT-ND1, MT-ND2, and MT-CO1 remained unaffected or were even upregulated in SNCA-3x oDAn1 neurons.

Remarkably, MOSPD2, a gene associated with ER membrane contact sites, emerged as the most differentially upregulated gene in SNCA-3x oDAn1 cells. Conversely, ARMCX6, a gene involved in neuronal mitochondria trafficking, was the most prominently downregulated in SNCA-3x oDAn1 cells.

### 4.2.3 Effect of SNCA Triplication on oDAn2 and oDAn3

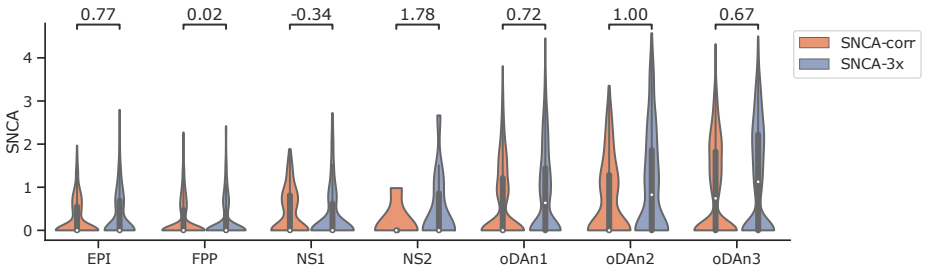



**Figure 4.4**: Violin plots depicting variations in SNCA gene expression between SNCA-3x and its isogenic control (SNCA-corr) across various identified cell types. The vertical axis represents log2-normalized expression levels. To account for possible discrepancies in cell counts among the groups, the log2 fold change was computed using resampling with replacement, generating a dataset of 1000 cells for each group. Due to their limited presence in SNCA-corr organoids, NEU1 and NEU2 cell types are omitted from this plot. **Note:** These plots were generated by Nikos Patikas.

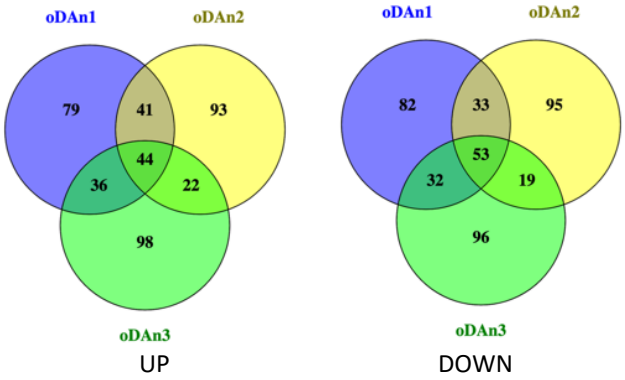



**Figure 4.17:** Venn diagram illustrating the overlap of the top 200 differentially upregulated and downregulated genes among distinct HMO DA neuron populations.

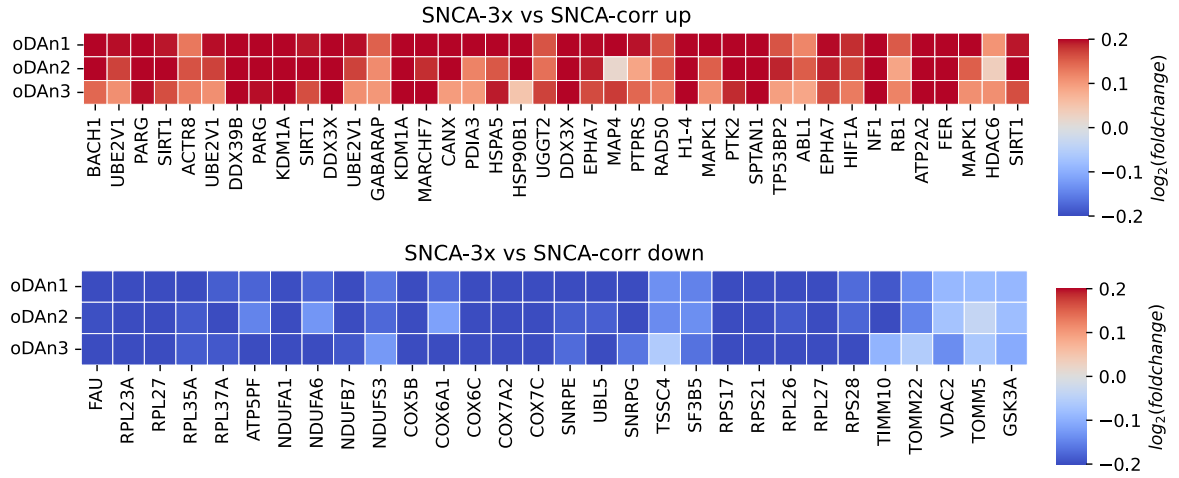


**Figure 4.5**: Depiction of the expression patterns of specifically chosen upregulated and downregulated genes in SNCA-3X oDAn1, compared to their expression in SNCA-3x oDAn2 and oDAn3. The log2-fold change for each cell type was computed through a comparison between SNCA-3x and SNCA-corr. **Note:** These plots were generated by Nikos Patikas.

Differential gene expression analysis was also conducted on oDAn2 and oDAn3 cells. Elevated SNCA expression was observed in both oDAn2 and oDAn3 cells when compared to their corresponding SNCA-corrected dopaminergic (DA) neurons. The overlap of the top 200 upregulated and downregulated genes across the various DA cell types (oDAn1, oDAn2, oDAn3) revealed 44 shared upregulated genes and 53 shared downregulated genes among these DA cell types. The commonly upregulated genes exhibited enrichment for processes such as cellular response to stress, dopamine metabolic processes, Golgi vesicle transport, and DNA damage response. Similarly, the shared downregulated genes were associated with biological functions including oxidative phosphorylation, translation, axon guidance, and non-coding RNA processing.

Furthermore, the overlap analysis identified 79 genes specifically upregulated, and 82 genes specifically downregulated in SNCA-3x oDAn1 cells. The SNCA-3x oDAn1 specific upregulated genes demonstrated enrichment for processes related to the unfolded protein response (UPR) and oxidative stress. Notably, genes associated with cellular component organization (MAP4 and PTPRS), endoplasmic reticulum stress (PDIA3 and HSPA5), protein ubiquitination (UBE2V1), and apoptosis (MAPK1) exhibited higher expression levels in SNCA-3x oDAn1 compared to SNCA-3x oDAn2 and oDAn3 cells.

Many of the upregulated and downregulated genes observed in SNCA-3x oDAn1 cells displayed a similar pattern in the other DA cell types, suggesting a shared response to the SNCA-3x mutation across all three DA cell types.

### 4.2.4 Validation of SNCA-3x Mutation Signatures Using qPCR

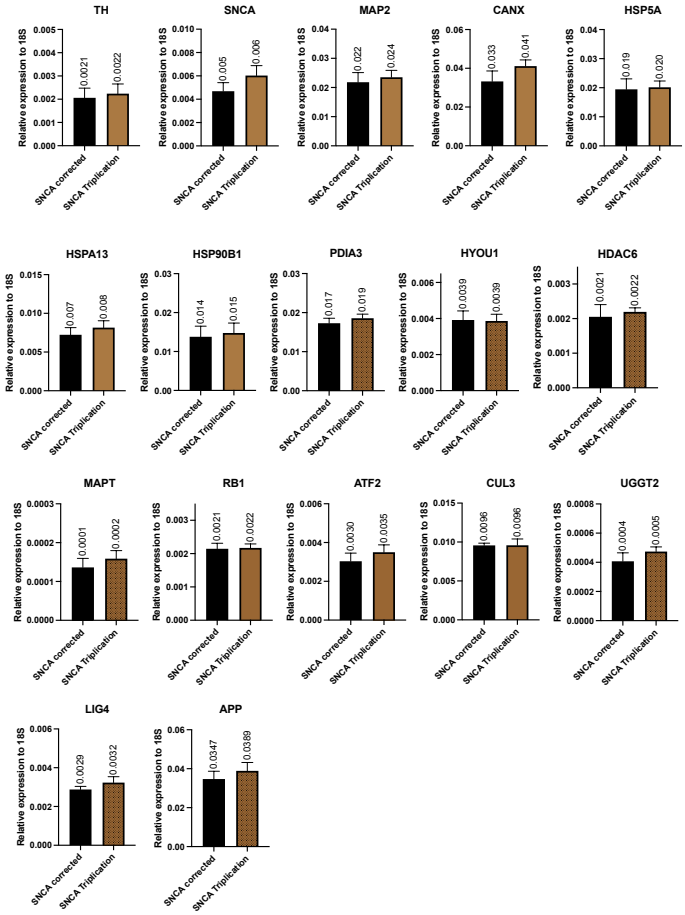



**Figure 4.6**: Bar plots illustrating the expression levels of chosen genes linked to SNCA triplication mutations within HMOs, as determined through qPCR analysis.

To validate the phenotypic effects associated with the identified SNCA mutation using single-cell RNA sequencing (scRNA-seq), a subset of genes that were observed to be upregulated in dopaminergic (DA) neurons was chosen for validation through quantitative polymerase chain reaction (qPCR). These selected genes were linked to cellular responses related to protein folding in the endoplasmic reticulum (CANX, PDIA3, HSPA5, HYOU1, and HSP90B1), the management of misfolded proteins (HSPA5, HSPA13, CUL3, HDAC6, UGGT2), as well as neuron death and apoptotic processes (APP, ATF2, LIG4, and RB1).

As qPCR involves the analysis of total RNA extracted from organoid samples, the higher expression levels of the selected genes were validated at the total RNA level, consistent with the scRNA-seq dataset. Human midbrain organoids (HMOs) were generated and harvested on day 34 for the qPCR analysis. The qPCR results indicated elevated expression levels of CANX, HSPA5, HSPA13, HSP90B1, PDIA3, and HDAC6. Despite these trends, none of the observed results reached statistical significance.

### 4.2.5 Validation of SNCA-3x Mutation Signatures Using scRNA-seq

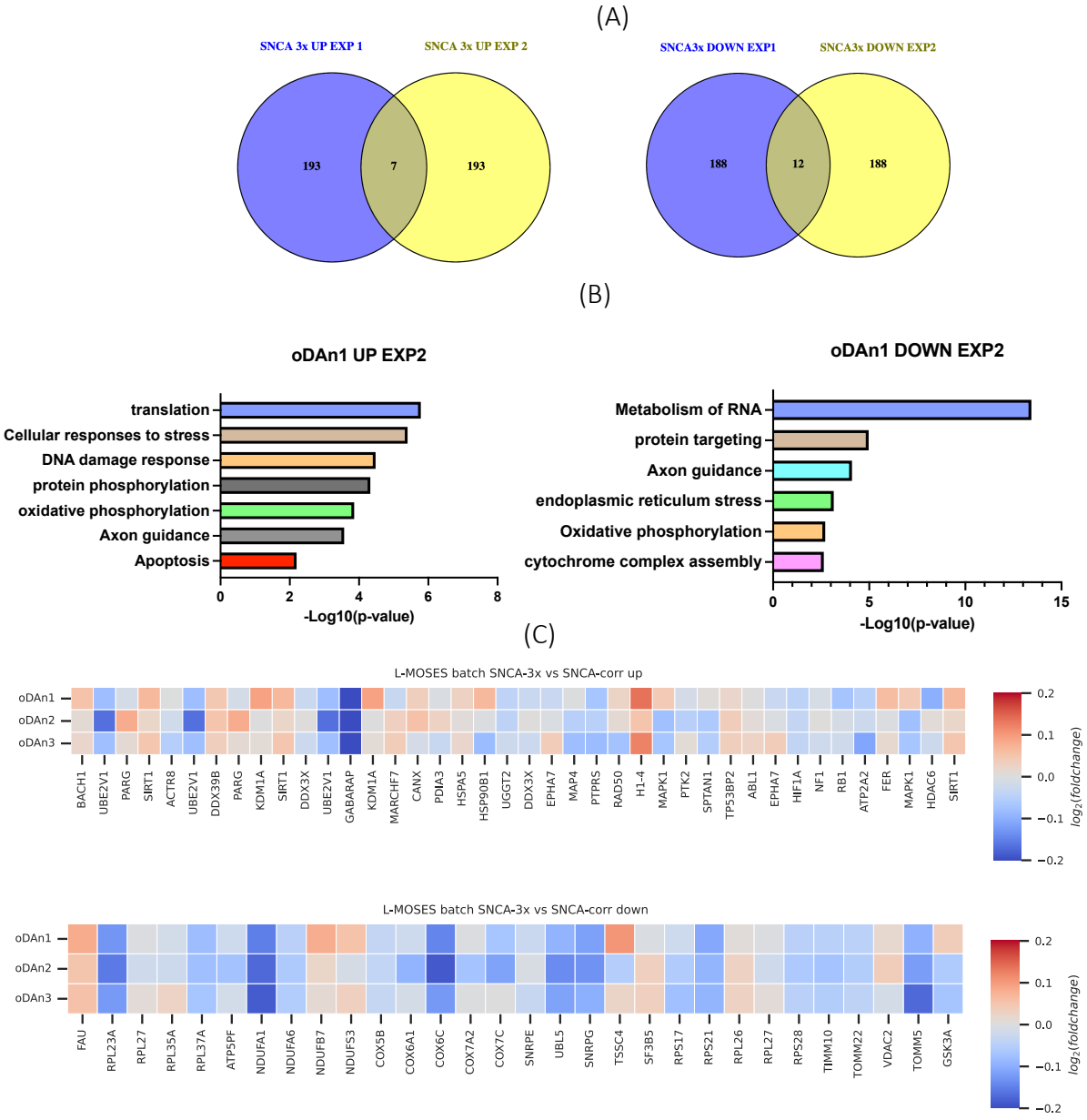



**Figure 4.7**: (A) Venn diagram illustrating the overlap of the top 200 differentially upregulated and downregulated genes among Matrigel and Geltrex SNCA-3x oDAn1 cells. (B) Bar plot illustrates enriched pathways corresponding to the top 200 differentially upregulated and downregulated genes in Geltrex SNCA-3x oDAn1 cells. The x-axis indicates -log10(P-value). (C) Matrix plot showing expression patterns of previously identified upregulated and downregulated genes in Geltrex SNCA-3X oDAn1, oDAn2 and oDAn3. The log2-fold change for each cell type was computed through a comparison between SNCA-3x and SNCA-corr. **Note:** These plots were generated by Nikos Patikas.

For this validation experiment, human midbrain organoids (HMOs) were generated using Geltrex matrix due to the global shortage of Matrigel matrix. HMOs were collected at the day 49 timepoint, consistent with the previous experiment, and processed for single-cell RNA sequencing (scRNA-seq). Differential gene expression analysis was conducted between SNCA-3x and SNCA-corrected (SNCA-corr) oDAn1 cells, revealing the top 200 differentially upregulated and downregulated genes. Pathway overrepresentation analysis was performed on these genes, highlighting both upregulated and downregulated pathways in SNCA-3x oDAn1 cells.

The upregulated genes were associated with processes such as cellular response to stress (e.g., CSNK2B, H1-4, HSP90AB1, and IGFBP7), DNA damage response (e.g., PARP4, BAK1, CCND1, and PRKDC), and apoptosis (e.g., PAK2, TJP1, and PSMD5). Additionally, several genes linked to axon guidance, translation, and oxidative phosphorylation exhibited differential expression in oDAn1 cells. Notably, genes related to endoplasmic reticulum stress (ATF4, SCTA, AIFM1, and EDEM1) were downregulated in SNCA-3x oDAn1 cells.

The overlap of the top 200 upregulated and downregulated genes between the two batches revealed seven shared upregulated genes and twelve shared downregulated genes in SNCA-3x oDAn1 cells. Among these, MOSP2 was identified as the most differentially upregulated gene, while ARMCX6 was the most differentially downregulated gene in both batches.

Although a majority of the top 200 differentially upregulated and downregulated genes did not coincide between the two experiments, the expression patterns of numerous genes previously associated with SNCA-3x were found to be consistent in oDAn1 cells. Genes related to endoplasmic reticulum stress (CANX, HSPA5, HSP90B1), cellular stress response (DDX39B, KDM1A, SIRT1), apoptosis (H1-4 and MAPK1), and response to oxidative stress (FER) exhibited upregulation in SNCA-3x oDAn1 cells. Similarly, genes associated with translation (RPL23A and RPL37A), oxidative phosphorylation (NDUFA1 and NDUFA6), and mitochondria transport (TIMM10, TOMM22, and TOMM5) were downregulated.

In summary, the differences observed between the results of the two experiments may have arisen from well-known batch-to-batch variations in organoids and the alteration in the matrix used for HMO generation.

### 4.2.6 Variations in Alpha-synuclein Protein Levels Over Time and Across Batches

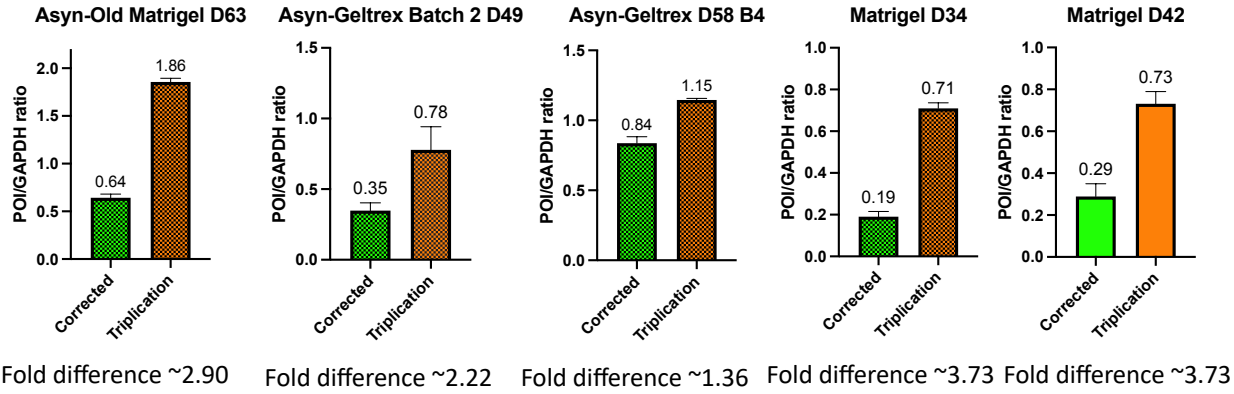



**Figure 4.8**: Bar plot illustrating the quantification of alpha-synuclein protein expression in different batches of HMOs at different time points, as determined through western blot analysis. The fold difference between SNCA-corr and SNCA-3x HMOs is indicated above the respective plots. The relative expression of alpha-synuclein was determined by dividing the arbitrary blot density values of the protein of interest (POI) by those of GAPDH protein for each HMO sample.

To investigate the subpar reproducibility of the SNCA-3x associated signatures in validation experiments, we examined the alpha-synuclein expression burden in human midbrain organoids (HMOs) at various timepoints and across different batches. Western blot analysis was performed on HMOs obtained from both early and late timepoints in two separate batches, produced using Geltrex and Matrigel matrices. The results indicated a reduction in alpha-synuclein levels in later timepoints.

In Geltrex HMOs, alpha-synuclein levels in SNCA-3x HMOs showed a 2.22-fold increase at day 49, which decreased to a 1.36-fold increase by day 58. Similarly, Matrigel HMOs exhibited a 3.73-fold increase at day 34 in alpha-synuclein levels for SNCA-3x HMOs, followed by a decrease to a 2.51-fold increase by day 42. The observed disparities in alpha-synuclein burden among HMOs may have contributed to the lack of reproducibility in the results.

### 4.2.7 Effect of L-MOSES on Alpha-synuclein Protein Levels and SNCA-3x Associated Transcriptomic Profile

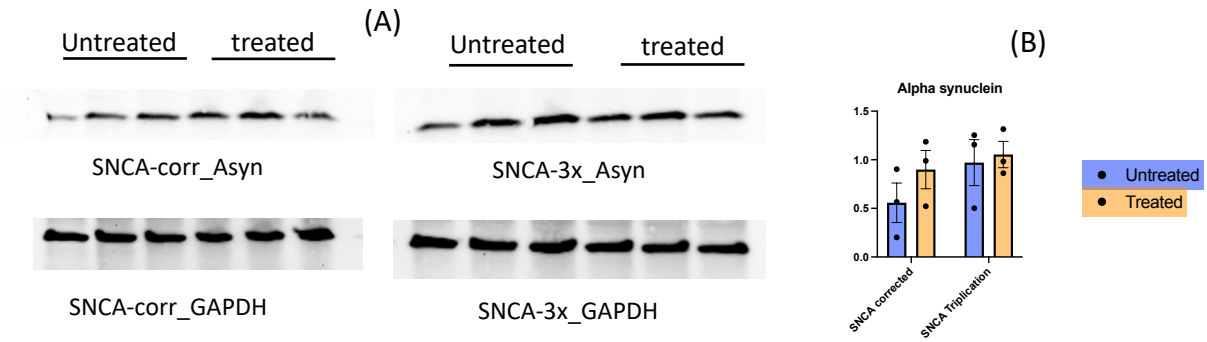



**Figure 4.9**: (A) Western blot analysis of alpha-synuclein protein expression in L-moses treated and untreated SNCA-corr and SNCA-3x HMOs. Housekeeping gene GAPDH was used as a loading control. (B) Quantification of alpha-synuclein protein expression, obtained from western blot analysis in (A). The relative expression of alpha-synuclein was calculated by dividing the arbitrary blot density values of the protein of interest (POI) by GAPDH protein for each HMO sample.

L-MOSES, an inhibitor of the lysine acetyltransferase KAT2B, has demonstrated its ability to reduce endoplasmic reticulum (ER) stress, attenuate unfolded protein response (UPR) activation, and prevent neuronal cell death in both cortical and dopaminergic neurons. Given that the SNCA-3x mutation results in the accumulation of misfolded alpha-synuclein within dopaminergic neurons—known to be detrimental to these cells—42-day-old SNCA-3x and SNCA-corrected control organoids were treated with L-MOSES for 7 days to assess its impact on alpha-synuclein protein levels. Western blot analysis revealed that L-MOSES did not lower the alpha-synuclein levels in either SNCA-3x or Control HMOs.

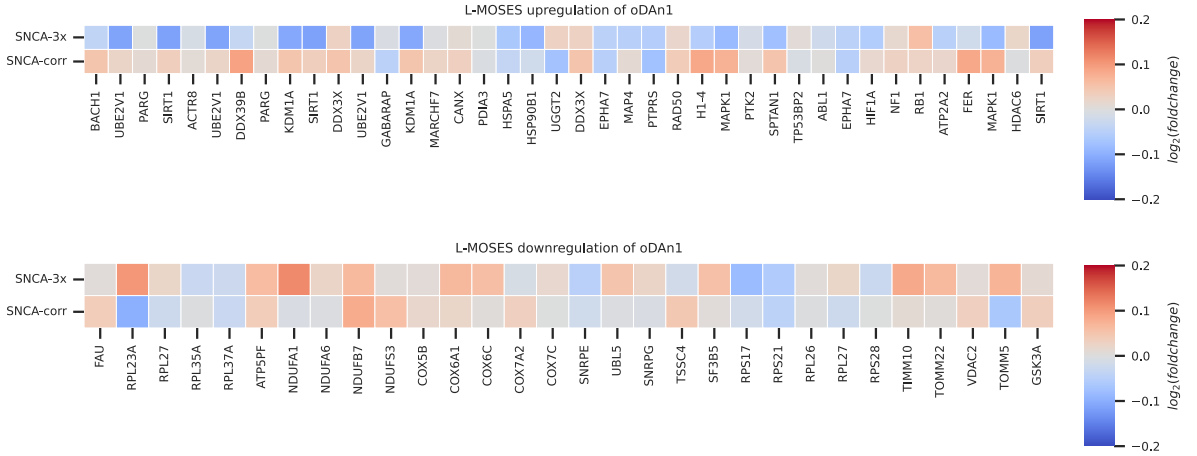


**Figure 4.10**: Matrix plots show expression patterns of previously identified SNCA triplication mutation associated genes in L-MOSES treated SNCA-corr and SNCA-3X oDAn1. The log2-fold change for each condition was computed by comparing treated and untreated HMOs. **Note:** These plots were generated by Nikos Patikas.

We investigated the impact of L-MOSES on the previously identified genes associated with SNCA-3x in oDAn1 cells. L-MOSES-treated SNCA-3x oDAn1 cells exhibited downregulation of genes linked to responses to endoplasmic reticulum stress (CANX, HSPA5, HSP90B1, UGGT2), cellular responses to oxidative stress (SIRT1, MAPK1, and FER), and apoptosis (H1-4 and SPTAN1). However, an opposite pattern was observed for some of these genes in L-MOSES-treated SNCA-corr oDAn1 cells.

L-MOSES-treated SNCA-3x oDAn2 cells displayed upregulation of genes involved in translation (RPL23A, RPL27, RPL37A, RPL27), oxidative phosphorylation (NDUFB7, NDUFB7, and NDUFS3), and mitochondria transport (TIMM10, TOMM22, and TOMM5). These results suggest that L-MOSES treatment appears to mitigate the transcriptional changes mediated by SNCA-3x mutation in SNCA-3x oDAn1 cells.

## 4.2.8 Effect Of L-MOSES On SNCA-3x and SNCA-Corr oDAn1 Cells

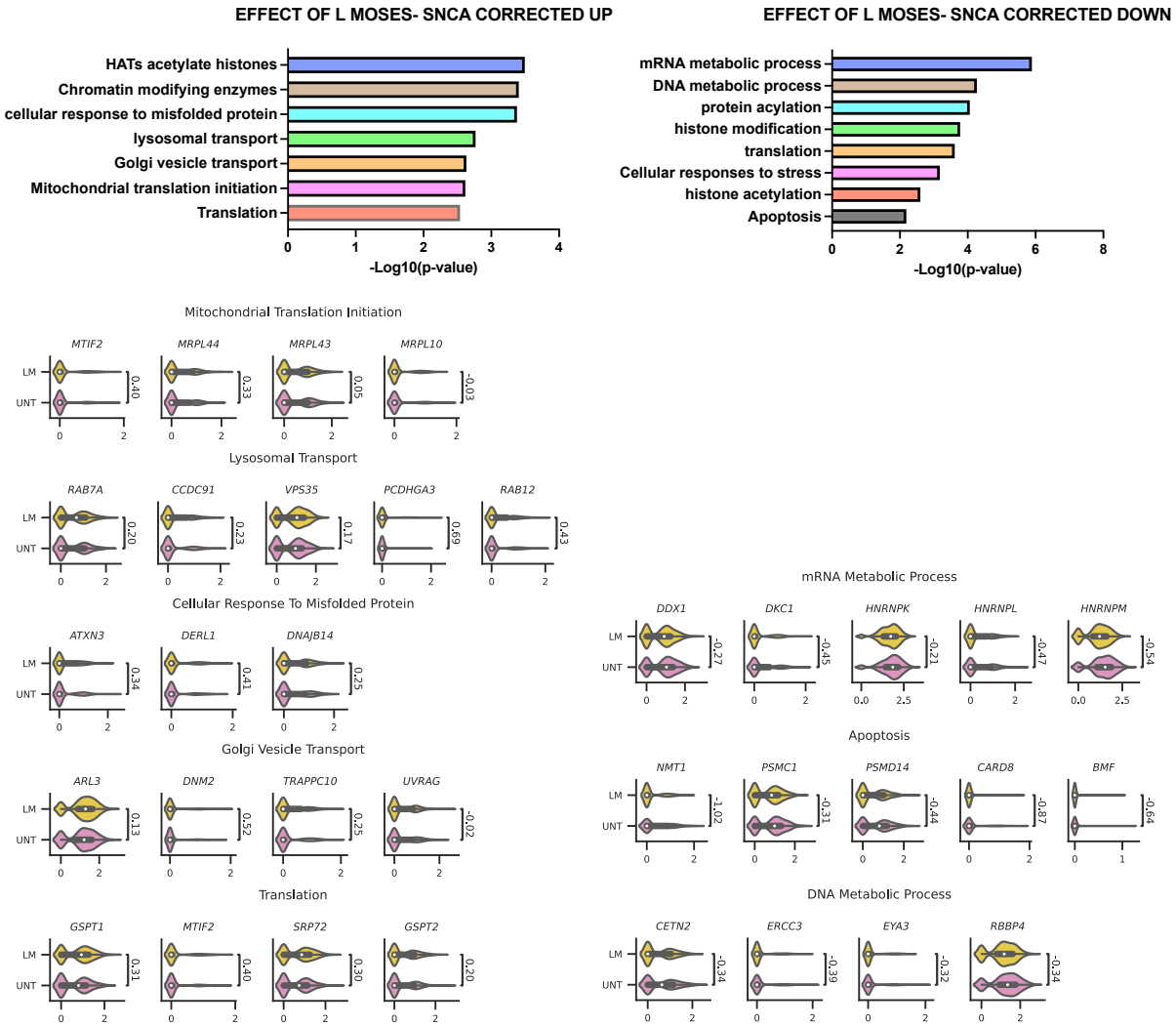
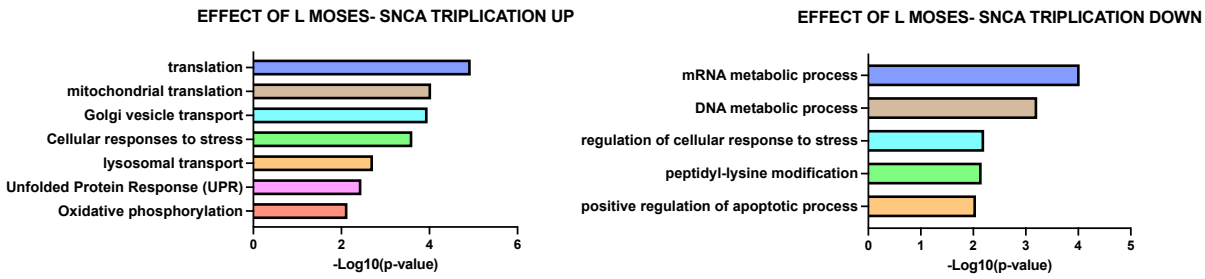



**Figure 4.11:** Bar plot illustrates enriched pathways corresponding to the top 200 differentially upregulated and downregulated genes in L-MOSES treated SNCA-corr oDAn1 cells, contrasted with untreated SNCA-corr oDAn1 cells. The x-axis indicates -log10(P-value). Violin plots present the log2-normalized expression levels of genes linked to these enriched pathways. Annotations between L-MOSES treated (LM) and untreated (UNT) conditions refers to log2 fold change. **Note:** These plots were generated by Nikos Patikas.



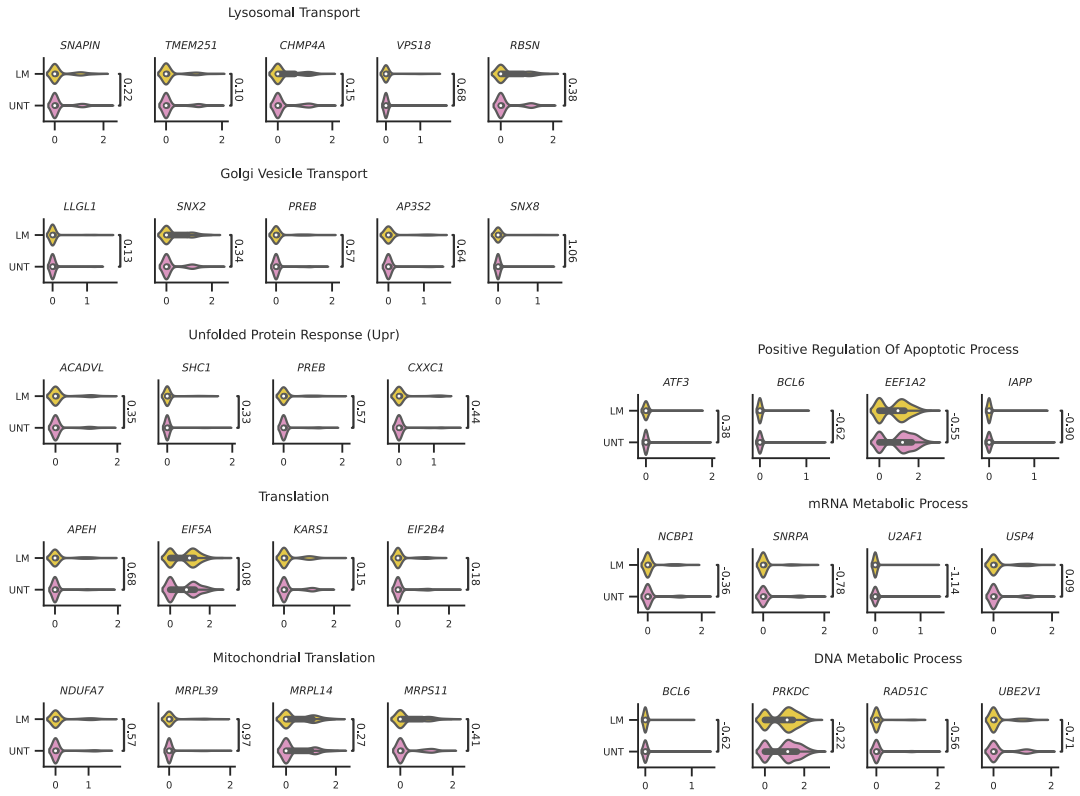



**Figure 4.12**: Bar plot illustrates enriched pathways corresponding to the top 200 differentially upregulated and downregulated genes in L-MOSES treated SNCA-3x oDAn1 cells, contrasted with untreated SNCA-3x oDAn1 cells. The x-axis indicates -log10(P-value). Violin plots present the log2-normalized expression levels of genes linked to these enriched pathways. Annotations between L-MOSES treated (LM) and untreated (UNT) conditions refers to log2 fold change. **Note:** These plots were generated by Nikos Patikas.

We also examined the impact of L-MOSES on SNCA-3x and SNCA-corrected (SNCA-corr) oDAn1 cells. Differential gene expression analysis between treated and untreated samples revealed several upregulated and downregulated genes. Pathway overrepresentation analysis on the top 200 differentially upregulated (UP) and downregulated (DOWN) genes highlighted the biological processes influenced by L-MOSES.

Both treated SNCA-3x and SNCA-corr oDAn1 cells demonstrated the upregulation of processes such as lysosomal transport, Golgi vesicle transport, translation, mitochondria translation, and the unfolded/misfolded protein response. Additionally, genes associated with oxidative phosphorylation were upregulated in treated SNCA-3x oDAn1 cells. On the other hand, the downregulated genes were linked to mRNA metabolic processes, DNA metabolic processes, and apoptosis in both treated SNCA-3x and SNCA-corr oDAn1 cells. Overall, these

results suggest that L-MOSES might exert a significant effect on protein processing and the degradation of misfolded proteins.

### 4.2.9 Identification of Rotenone-Sensitive and Resilient oDAn1 Cells through In-Silico Analysis

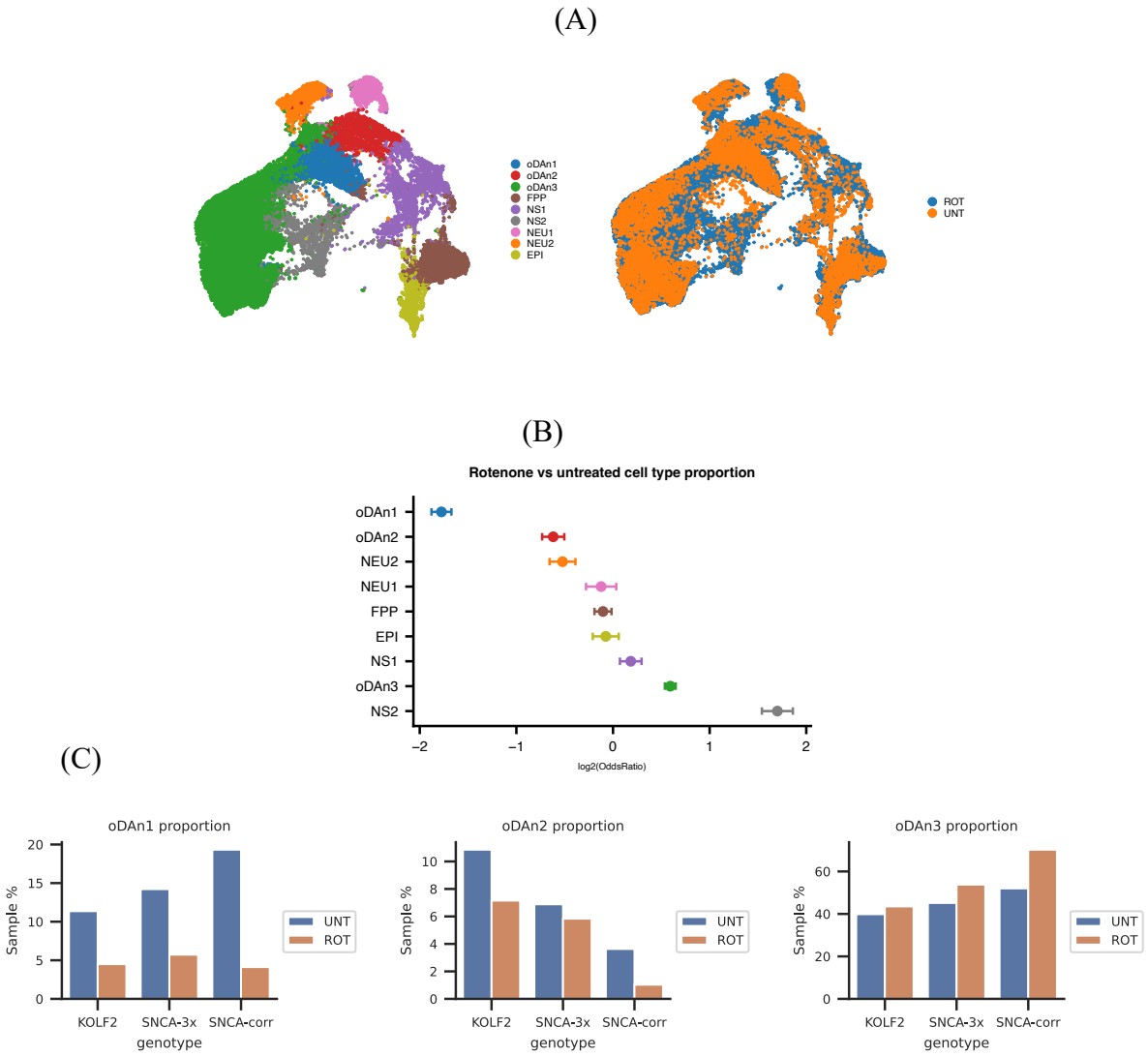



**Figure 4.13**: (A) UMAP plot showing overlap of rotenone-treated and untreated HMOs. (B) Evaluation of cell type proportions in rotenone-treated HMOs. Negative values indicate a reduction in the proportion of cells in the rotenone samples. Error bars depict the 95% confidence intervals derived from MASC. (C) Bar plots illustrating the proportion of dopaminergic subpopulations within HMOs in both rotenone-treated and untreated samples. **Note:** These plots were generated by Nikos Patikas.

To study the impact of oxidative stress on the HMOs, they were treated with a medium lacking antioxidants and supplemented with a widely used oxidative stress inducing compound

rotenone for 24hours. Rotenone functions by inhibiting complex I of the mitochondrial electron transport chain, thereby inducing oxidative stress through electron leakage and subsequent heightened production of reactive oxygen species (ROS). This excessive ROS generation overwhelms the cell's antioxidant defence system, consequently causing damage to various cellular components.

The rotenone-treated and untreated samples were subjected to integration on UMAP space, enabling the identification of cell types affected by oxidative stress. All previously identified cell types were identified in both the treated and untreated HMOs. Examination of cell type proportions unveiled a noteworthy reduction in oDAn1 cells within the rotenone-treated HMOs, a pattern consistently observed across all three lines used in this experiment, including SNCA-3x HMOs. Following oDAn1 cells, the next most affected cell type was oDAn2, trailed by NEU2. Conversely, oDAN3 and NS2 cells exhibited the least pronounced response to the rotenone condition. These encouraging findings offer further validation to the hypothesis that oDAn1 cells accurately replicate pivotal characteristics of Parkinson's disease, including their vulnerability to oxidative stress.

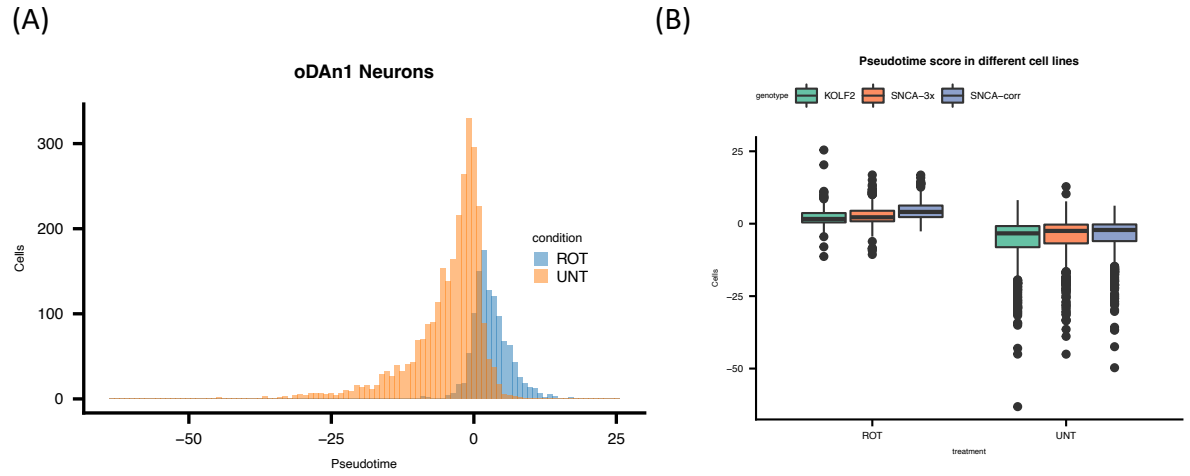



**Figure 4.14**: (A) a histogram displays the distribution of pseudo-time scores across 2845 untreated and 1145 rotenone-treated oDAn1 cells. (B) On the right, a boxplot exhibits the pseudo-time oDAn1 cell scores, categorized by genotype and treatment.

Traditional differential expression methods fall short in connecting changes in RNA expression with potential alterations in cell type proportions under experimental conditions. These two variables can be effectively explored using single-cell transcriptomic techniques,

such as Pseudotime analysis. This approach establishes a correlation between the likelihood of a single untreated cell's absence following rotenone treatment, assisting in the identification of rotenone-sensitive profiles within the untreated cell population. Additionally, this metric underscore transcriptomic disparities between surviving rotenone-treated cells and untreated cells sharing similar cell states. Employing this single-cell metric, pseudotime scores were calculated for both untreated and rotenone-treated oDAn1 cells, subsequently used as pseudotime values in our analysis.

The analyzed oDAn1 cells were derived from three distinct cell lines, including SNCA-3x, comprising a total of 3970 cells. Both the treatment and genotype exhibited a significant influence on the Pseudotime scores, with the treatment demonstrating a p-value of $< 2e-16$ and the genotype showing a p-value of 1.42e-05. However, the interaction term between treatment and genotype did not exhibit a statistically significant difference (p-value 0.057), suggesting that the effect of rotenone is not specific to genotype.

The Pseudotime score of an untreated oDAn1 cell serves as a measurement of its sensitivity to rotenone. Lower and higher values correspond to susceptibility and resilience upon treatment, respectively. The Switchde package (Campbell 2022) was employed to ascertain the transcriptomic profiles of these trajectories. Subsequently, a pathway overrepresentation analysis was conducted to identify the pathways enriched in both rotenone-sensitive and resilient oDAn1 cells.

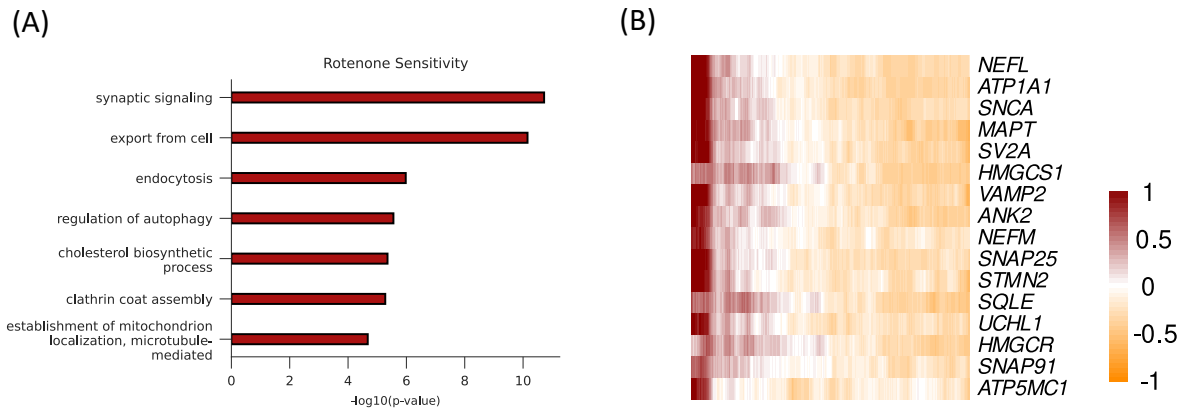



**Figure 4.15**: (A) The bar plot depicts enriched pathways for the top 200 genes identified in rotenone-sensitive oDAn1 cells. These genes were sorted by the switched p-value, with the x-axis representing -log10 (P value). (B) Smoothed heatmaps provide an illustration of the shifts in gene expression along the pseudo-time trajectory within untreated oDAn1 cells. The color intensity reflects the standard deviation of mean gene expression. The selected genes in this heatmap correspond to the pathways highlighted in panel A. **Note:** These plots were generated by Nikos Patikas.

Enrichment analysis of rotenone-sensitive oDAn1 cells revealed an abundance of genes associated with various functions, including synaptic signaling (e.g., NEFL, NEFM, SNAP91, SNAP25), cholesterol biosynthesis (e.g., SQLE, HMGCR, and HMGCS1), endocytosis, and cellular export. Furthermore, these sensitive oDAn1 cells exhibited elevated expression of genes linked to Parkinson's disease (PD), such as SNCA, UCHL1, and ANK2. Another significant gene observed in these cells was MAPT, which also demonstrated high expression. However, these enriched pathways were not unique to oDAn1 cells, as seen by the instance of oDAn3 cells displaying high expression of cholesterol biosynthesis-associated genes like SQLE and HMGCR. Nevertheless, the overlap between cell markers of oDAn1 cells and rotenone sensitivity-associated genes showcased eight shared genes, including PPP2R2B, CISD1, HRAS, PTS, and SYT5. PPP2R2B, identified through PD GWAS, is involved in protein dephosphorylation. SYT5, with altered expression reported in the brains of individuals with PD, presents another interesting candidate.

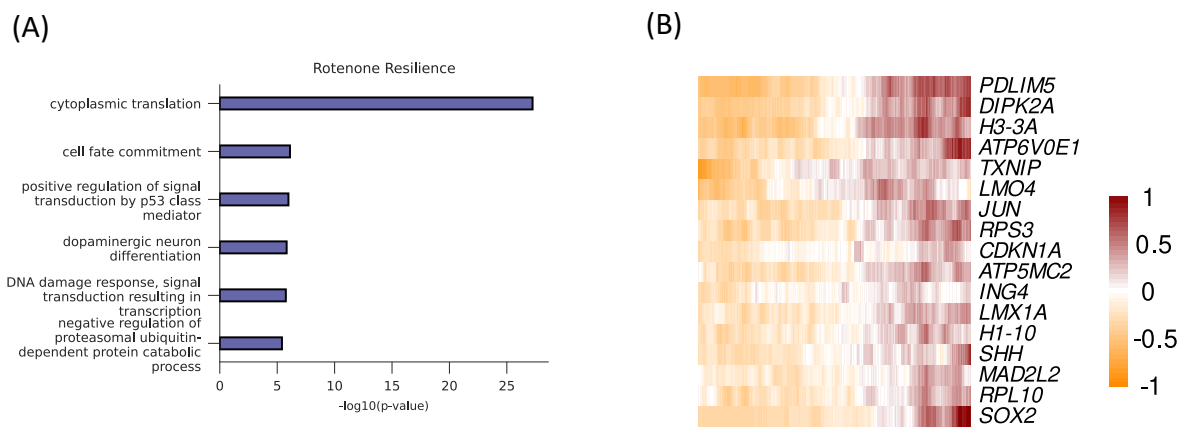



**Figure 4.16**: (A) The bar plot depicts enriched pathways for the top 200 genes identified in rotenone-resilient oDAn1 cells. These genes were sorted by the switched p-value, with the x-axis representing -log10 (P value). (B) Smoothed heatmaps provide an illustration of the shifts in gene expression along the pseudo-time trajectory within untreated oDAn1 cells. The color intensity reflects the standard deviation of mean gene expression. The selected genes in this heatmap correspond to the pathways highlighted in panel A. **Note:** These plots were generated by Nikos Patikas.

Rotenone-resilient oDAn1 cells exhibited enrichment in genes associated with various functions, including translation (e.g., RPS3 and RPL10), dopaminergic neuron differentiation (e.g., LMX1A and SHH), and DNA repair (e.g., ING4, MAD2L2, and CDKN1A). Interestingly, the resilient cells also demonstrated higher expression of SOX2, LMX1A, SHH, and LMO4, suggesting that immature DA neurons had a higher survival rate under rotenone conditions, as expected. Additionally, JUN, a gene involved in regulating cell survival and apoptosis, exhibited high expression in the resilient cells, potentially contributing to the survival of DA neurons.

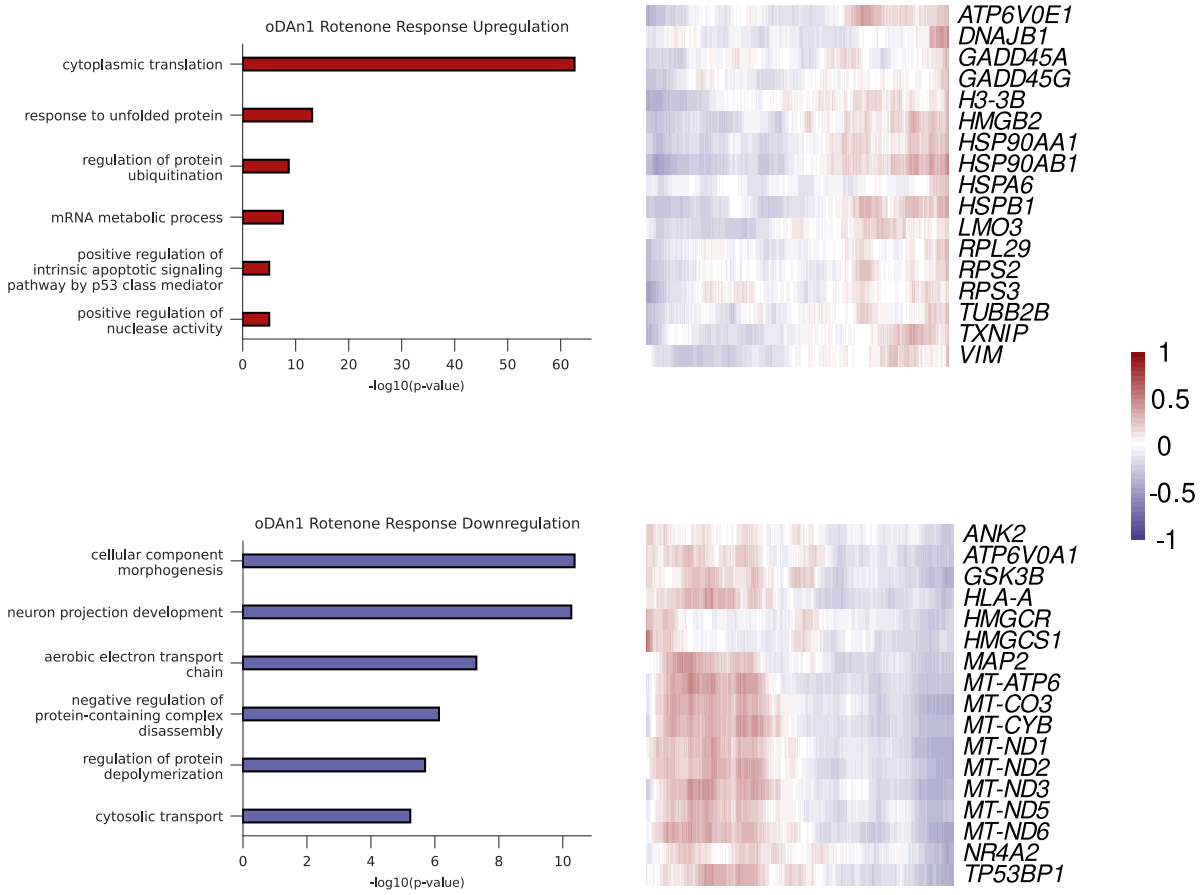



**Figure 4.17**: (A) The bar plot depicts enriched pathways for the top 200 genes upregulated and downregulated in rotenone-treated oDAn1 cells. These genes were sorted by the switched p-value, with the x-axis representing -log10 (P value). (B) Smoothed heatmaps provide an illustration of the shifts in gene expression along the pseudo-time trajectory within treated oDAn1 cells. The color intensity reflects the standard deviation of mean gene expression. The selected genes in this heatmap correspond to the pathways highlighted in panel A. **Note:** These plots were generated by Nikos Patikas.

The pseudotime score assigned to a rotenone-treated oDAn1 cell indicates its transcriptome divergence from untreated cells within the latent space. By applying this pseudotime score to the rotenone-treated cells, genes that are likely to be influenced by rotenone were discerned.

Pathways upregulated in response to rotenone treatment include translation-related processes (e.g., RPL29, RPS29, and RPS3), and the unfolded protein response (UPR) pathway (e.g., HSP90AA1, HSP90AB1, HSPB1, and HSPA6). Notably, TXNIP also displayed high expression in response to rotenone. TXNIP encodes the protein thioredoxin interacting protein,

a crucial regulator of cellular redox balance. Furthermore, VIM, a gene associated with DA progenitor cells, exhibited prominent expression in rotenone-treated cells.

As anticipated, several mitochondria DNA-encoded genes (e.g., MT-ATP6, MT-CO3, MT-CYB, and MT-ND6) experienced downregulation in response to rotenone in oDAn1 cells. Genes associated with cholesterol biosynthesis (HMGCR and HMGCS1) also demonstrated downregulation. Intriguingly, the dopaminergic neuronal marker NR4A2 and the neuronal maturity marker MAP2 were found to be repressed in rotenone-treated oDAn1 cells. The combined upregulation of VIM and downregulation of MAP2 suggest that rotenone had a notable impact on mature neurons.

### 4.2.10 Derivation and Characterization of Human Chimera Midbrain Organoids

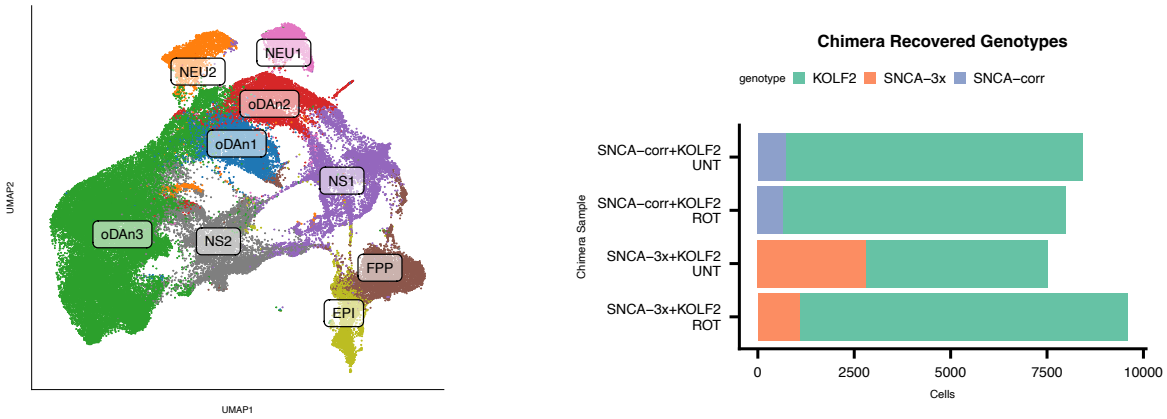



**Figure 4.18**: (A) UMAP embedding of a combined dataset comprising cells from both chimera and non-chimera HMOs, totaling 76,000 cells. (B) Representation of the cell count retrieved post-genotype demultiplexing in chimera organoids. **Note:** These plots were generated by Nikos Patikas.

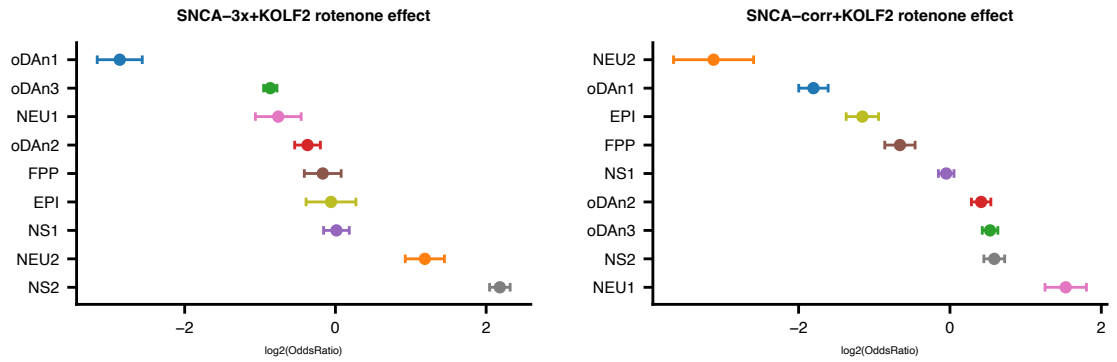


**Figure 4.19**: Effect of rotenone on the distinct cell types identified within individual chimera samples. The error bars represent 95% confidence intervals obtained from MASC. Note: These plots were generated by Nikos Patikas.

Chimera HMOs were generated to assess the feasibility of cultivating cells from different donors within a single tissue structure and to study non-autonomous effects. Two types of chimera organoids were created using previously established healthy KOLF2, SNCA-3x, and SNCA-corrected control iPSCs. The identity of cells within chimera organoids was determined by comparing them with monoculture samples as a genetic reference.

Unified datasets of chimera HMOs demonstrated that the presence of multiple donor cells within the same culture did not affect differentiation. All previously identified cell types were also observed in the chimera organoids. However, analysis of cell proportions revealed a significant shift in donor distribution within both types of chimera HMOs. Although cells were initially seeded at a 1:1 ratio on day 0, examination of HMOs on day 49 indicated a distribution of 65% KOLF2 and 35% SNCA-3x cells in the KOLF2 + SNCA-3x chimera, and a distribution of 90% KOLF2 and 10% SNCA-corrected cells in the KOLF2 + SNCA-corrected chimera.

Upon rotenone treatment, the KOLF2 + SNCA-corrected chimera showed no significant changes, whereas the KOLF2 + SNCA-3x chimera exhibited a reduced proportion of SNCA-3x cells compared to untreated conditions. At the level of cell types, similar to monocultures, oDAn1 cells were diminished in response to rotenone in both types of chimera cultures. Discrepancies in the reduced proportion of EPI and NEU2 cells in rotenone-treated KOLF2 + SNCA-corrected chimera may be attributed to the absence of EPI cells in KOLF2 HMOs and NEU2 cells in SNCA-corrected HMOs. Notably, NS2 cells demonstrated no detrimental effects in response to rotenone in both chimera HMOs, mirroring their behavior in monoculture conditions.

### 4.2.11 Investigation of non-cell autonomous effect on healthy and disease cells

Significant research in Parkinson's disease (PD) has been focused on exploring "cell-autonomous" mechanisms, while the study of "non-cell-autonomous" mechanisms has received comparatively less attention. Although it is acknowledged that non-cell-autonomous effects on

dopaminergic neurons hold a substantial role in PD pathogenesis, the underlying mechanisms are still not fully elucidated.

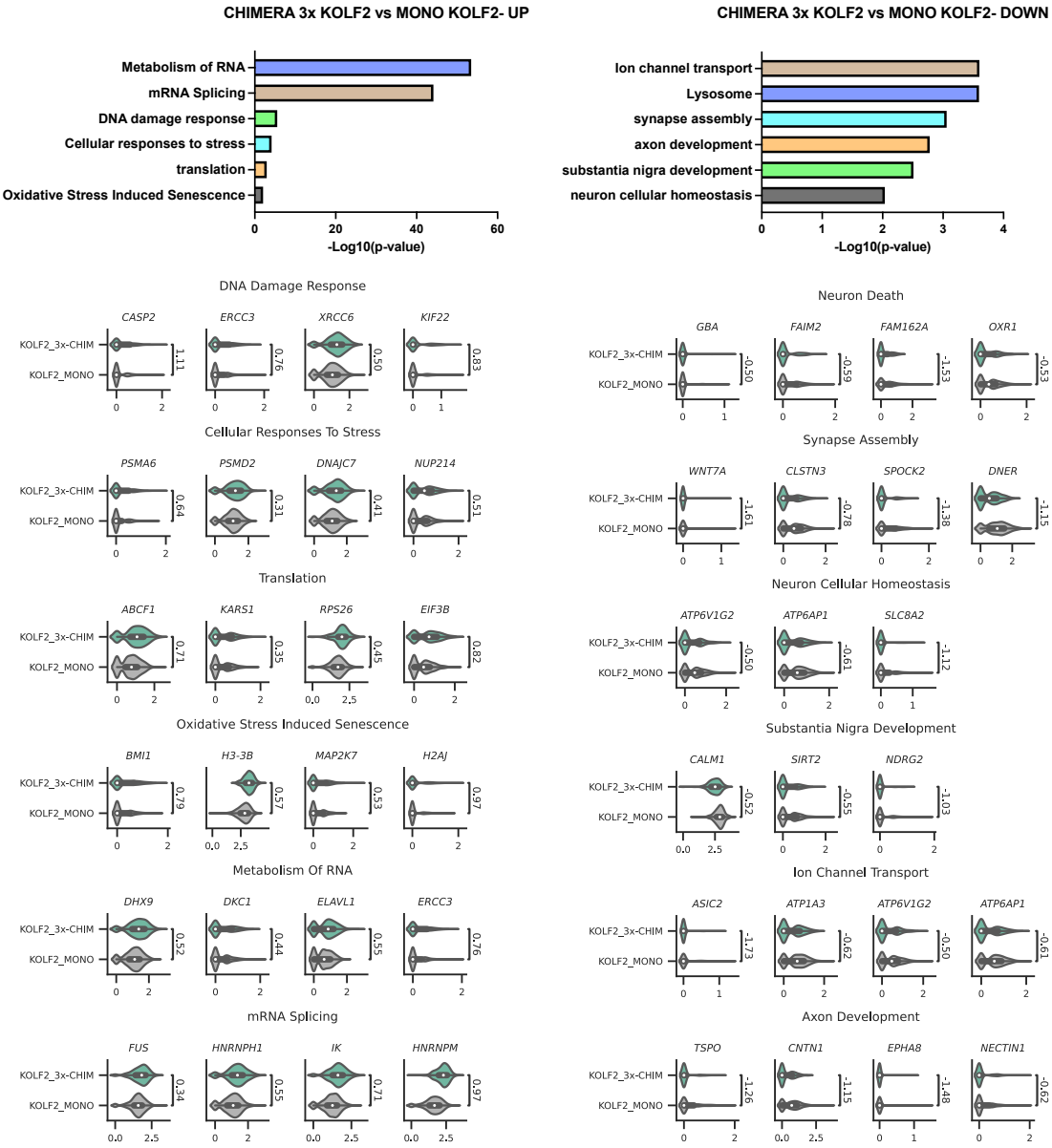



**Figure 4.20**: Bar plot illustrates enriched pathways corresponding to the top 200 differentially downregulated genes in chimera KOLF2-SNCA-3x oDAn1 cells, contrasted with monolayer KOLF2 oDAn1 cells. The x-axis indicates -log10(P-value). Violin plots present the log2-normalized expression levels of genes linked to these enriched pathways. Annotations between SNCA-3x and SNCA-corr conditions refers to log2 fold change. **Note:** These plots were generated by Nikos Patikas.

Differential gene expression analysis was conducted between healthy KOLF2 oDAn1 cells cultured with disease-associated SNCA-3x cells (chimera KOLF2) and monoculture KOLF2 oDAn1 cells to investigate the non-cell-autonomous effect of the SNCA-3x mutation on healthy dopaminergic (DA) neurons when co-cultured with diseased cells. While many previously identified signatures, such as the upregulation of genes linked to the unfolded protein response (UPR) and protein folding in the endoplasmic reticulum, as well as the downregulation of genes associated with translation and oxidative phosphorylation, were not observed in chimera KOLF2 oDAn1 cells, distinct patterns emerged.

Genes associated with DNA damage response (e.g., CASP2, ERCC3, XRCC6, KIF22) and stress response (e.g., PSMA6, PSMD2, DNAJC7, NUP214) displayed heightened expression. Conversely, genes linked to synapse assembly (e.g., WNT7A, CLSTN3, SPOCK2, and DNER), axon guidance (e.g., TSPO, CNTN1, EPHA8, and NECTIN1), substantia nigra development (e.g., CALM1, SIRT2, NDRG2), and neuron cellular homeostasis (e.g., ATP6V1G2, ATP6AP1, and SLC8A2) exhibited downregulation.

In summary, these findings suggest that although chimera KOLF2 oDAn1 cells did not entirely replicate the mutational signatures of SNCA-3x, the co-culturing of these cells with SNCA-3x-mutating cells could potentially induce detrimental effects on the healthy cells.

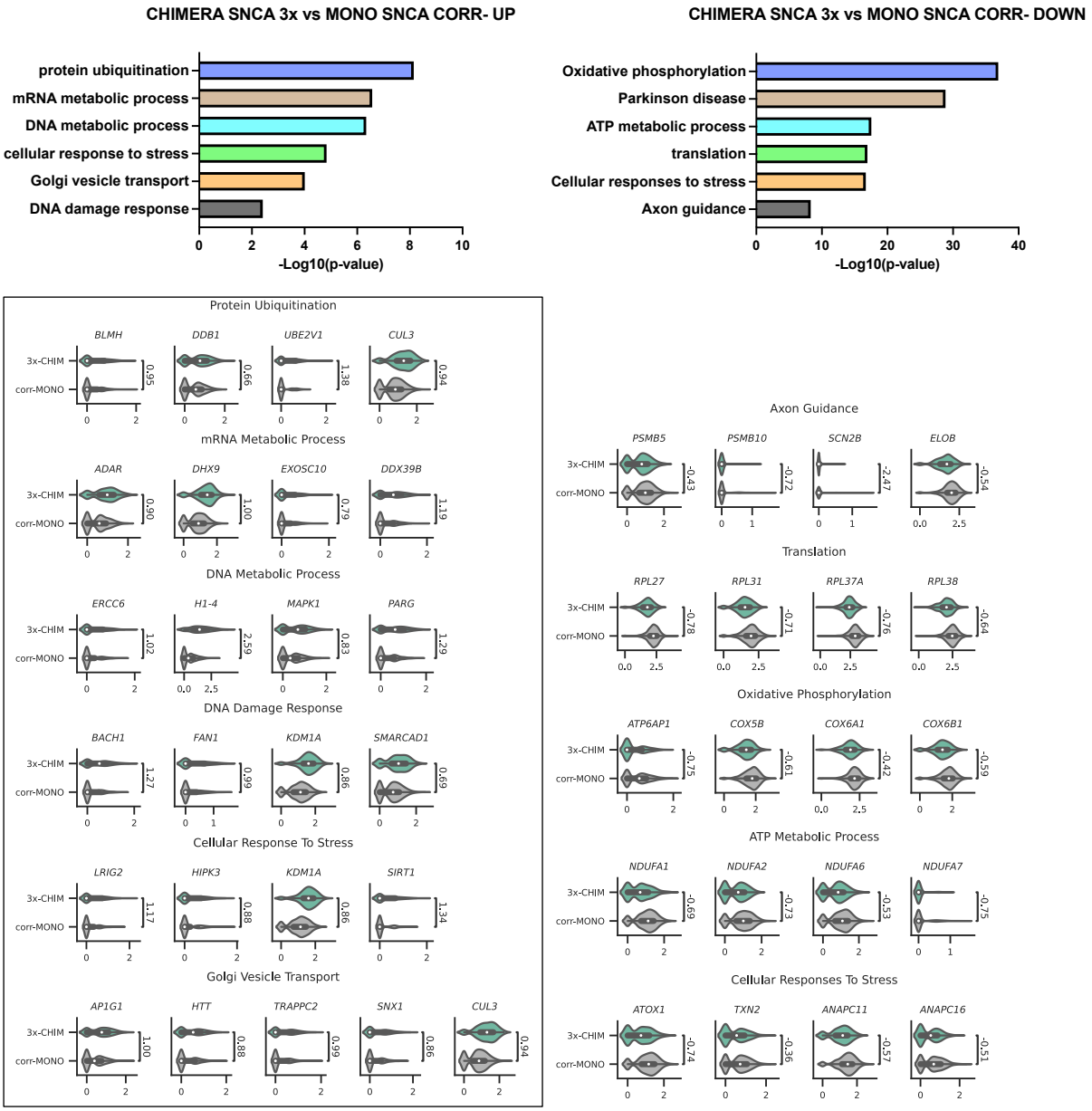



**Figure 4.21**: Bar plot illustrates enriched pathways corresponding to the top 200 differentially downregulated genes in chimera SNCA-3x oDAn1 cells, contrasted with monolayer SNCA-corr oDAn1 cells. The x-axis indicates -log10(P-value). Violin plots present the log2-normalized expression levels of genes linked to these enriched pathways. Annotations between SNCA-3x and SNCA-corr conditions refers to log2 fold change. **Note:** These plots were generated by Nikos Patikas.

Differential gene expression analysis was conducted between SNCA-3x oDAn1 cells cultured with healthy KOFL2 cells (chimera SNCA-3x) and monoculture control oDAn1 cells to explore the impact of the SNCA-3x mutation on disease-relevant dopaminergic neurons when co-cultured with healthy cells. The downregulated genes in chimera SNCA-3x oDAn1

cells exhibited enrichment for processes like oxidative phosphorylation (e.g., ATP6AP1, COX5B, COX6A1, COX6B1), translation (e.g., RPL27, RPL31, RPL37A, RPL38), and axon guidance (e.g., PSMB5, PSMB10, SCN2B, ELOB), similar to what was observed in monolayer SNCA-3x oDAn1 cells.

The upregulated genes in chimera SNCA-3x oDAn1 cells displayed enrichment for processes such as cellular response to stress (e.g., LRIG2, HIPK3, KDM1A, SIRT1), DNA damage response (e.g., BACH1, FAN1, KDM1A, SMARCAD1), protein ubiquitination (e.g., BLMH, DDB1, UBE2V1, CUL3), and DNA metabolic processes (e.g., ERCC6, H1-4, MAPK1, PARG). These Gene Ontology (GO) terms were also enriched in monoculture SNCA-3x oDAn1 neurons. However, notable differences emerged: GO terms related to protein folding in the endoplasmic reticulum or the unfolded protein response, which were observed in monolayer SNCA-3x oDAn1 neurons, were not evident in chimera SNCA-3x oDAn1 cells. These findings imply that chimera SNCA-3x oDAn1 cells, surrounded by healthy cells, may have an influence on the cellular response to the SNCA-3x mutation.

## 4.3 Discussion and Future Directions

To investigate the impact of the SNCA triplication mutation on dopaminergic neurons, we compared SNCA-3x HMOs with their corresponding isogenic controls, where two out of the four SNCA gene copies were knocked out. The triplication mutation of the SNCA gene (SNCA-3x) is associated with elevated expression and aggregation of alpha-synuclein, contributing to early onset PD. Utilizing single-cell RNA-sequencing and western blot analysis, we observed an overexpression of SNCA mRNA levels (0.82-fold change) and alpha-synuclein (2.82-fold change) in SNCA-3x HMOs when compared to the Control. We observed SNCA-3x mutation did not lead to a linear overexpression of its mRNA level. This finding is inconsistent with the study by Mohamed et al., which reported a 2.5-fold increase in mRNA levels in SNCA-3x midbrain organoids compared to the Control group, albeit using qPCR (Mohamed et al., 2021). Additionally, while a 2.82-fold change in alpha-synuclein protein expression was detected, this may not necessarily result from a dose-dependent linear translation of alpha-synuclein protein due to the SNCA gene dosage. Rather, it could be attributed to the accumulation of misfolded or unfolded proteins, resulting from the excessive production of alpha-synuclein and subsequent disruption of protein homeostasis.

We have identified a consistent downregulation of oxidative phosphorylation and translation in response to the SNCA-3x mutation across all subtypes of dopamine neuron populations in our model. This observation suggests a shared response to the SNCA-3x mutation within all three dopamine neuron subtypes. A previous single-cell RNA-seq study conducted by our group on A53T monolayer dopaminergic neuronal cultures reported a downregulation of synuclein family genes such as SNCA and SNCG in the SNCA A53T mutant culture when compared to the isogenic controls (Fernandez et al., 2020). Despite the downregulation of SNCA expression, SNCA A53T dopamine neurons exhibited concurrent downregulation of oxidative phosphorylation and translation. This finding implies that the downregulation of oxidative phosphorylation and translation may be a phenotype associated with SNCA mutations, and further investigation through additional experiments is warranted.

### 4.3.1 Dysregulation in Protein Translation in SNCA-3x Dopaminergic Neurons

Numerous genes associated with PD have been implicated in the protein translation process. This connection manifests in two distinct ways: firstly, the products of these genes may actively participate in protein synthesis, and secondly, specific phenotypes associated with these genes can potentially lead to regulatory alterations in translation (Kim et al., 2017).

Recent investigations into familial forms of PD have unveiled a compelling link between PD and protein translation. Studies have recently illustrated that the G2019S mutation in LRRK2 results in an upsurge in overall protein synthesis, consequently leading to neurotoxicity (Martin et al., 2014). Furthermore, GWAS have identified mutations in the eukaryotic initiation factor 4 gamma-1 (eIF4G1) that are associated with PD, and there have been indications suggesting that PTEN-induced putative kinase 1 (PINK1) plays a role in the synthesis of mitochondrial proteins (Gehrke et al., 2015; Chartier-Harlin et al., 2011). While mounting evidence substantiates the concept that disrupted protein synthesis is entwined with PD, the precise mechanisms underlying this phenomenon are actively under investigation.

It has been postulated that the unfolded protein response (UPR), initiated by protein aggregates, results in elevated levels of phosphorylated eIF2α, consequently leading to a reduction in overall protein translation (Kaufman et al., 2002; Colla et al., 2019). It is highly conceivable that the observed downregulation of global translation in our SNCA-3x PD model may stem from the UPR response. Nevertheless, specific facets of the translation process, such

as the ribosome, mitoribosome, and translation initiation factors, which have been reported, may be unique to the signatures associated with the SNCA-3x triplication mutation. Thus, further investigation and validation is warranted to delineate the pathogenesis linked to SNCA-3x associated PD.

### 4.3.2 ER Stress and UPR in SNCA-3x SNCA-3x Dopaminergic Neurons

Our examination of SNCA-3x DA neurons revealed an increase in the expression of genes associated with ER stress. Although we observed an elevation in the levels of ER protein folding chaperones (CANX, HSPA5, HSP90B1), the typical signatures of the unfolded protein response (UPR) were conspicuously absent in our SNCA-3x- PD model. Instead, we noted a downregulation of ATF4, a gene pivotal in the UPR response and a component of the PERK branch. PERK phosphorylates eIF2α, which in turn dampens global protein translation while selectively boosting the translation of specific proteins, such as ATF4. The downregulation of ATF4 in our SNCA-3x DA neurons suggested that the PERK-mediated UPR response might not be active in these cells. The observed reduction in global translation in our SNCA-3x neurons might involve an alternative mechanism.

ATF4 has been shown to play a role in re-establishing ER homeostasis by regulating the expression of chaperones, foldases, and components of autophagy (Hetz, 2018). During prolonged ER stress, ATF4 induces the expression of proapoptotic factors, including the production of reactive oxygen species (ROS), upregulation of the CHOP transcription factor, and modulation of the BCL-2 family (Tabas and Ron, 2011; Pihan et al., 2017). While ATF4 was identified as a downregulated regulon in our analysis, several of its target genes also exhibited downregulation. This phenomenon may help elucidate why a UPR was not evident in SNCA-3x DA neurons, despite their heightened levels of ER protein folding chaperones.

Heman-Ackah et al. reported the induction of the IRE1α/XBP1 axis of the unfolded protein response (UPR), culminating in terminal UPR activation in neurons derived from SNCA-3x iPSCs (Heman-Ackah et al., 2017). They substantiated the clinical relevance of this phenotype by demonstrating elevated levels of pIRE1α in neurons containing Lewy bodies from post-mortem PD brain tissue. Zombon et al. also reported a significant increase in IRE1α protein, a mediator of the UPR response, in iPSCs DA neurons carrying A53T and SNCA-3x mutation.

Conversely, a recent study by Stojkovska, employing patient-derived SNCA-3x DA neurons, demonstrated that alpha-synuclein overexpression induces ER fragmentation and promotes the aggregation of immature lysosomal GCase. Significantly, they observed a lack of the unfolded protein response in SNCA-3x DA neurons, suggesting that alpha-synuclein may interfere with the cell's ability to detect or transduce signals related to protein misfolding in the ER (Stojkovska et al., 2022). They also noted a mild increase in ER stress chaperones, such as GRP78, GRP94, and CANX, in the absence of UPR activation. In our study, we also observed an upregulation of protein folding chaperones, including CANX. The findings from both our study and Stojkovska's suggest that SNCA-3x DA neurons may lack the capability to initiate a UPR response in a disease environment.

### 4.3.3 Mitochondrial Dysfunction and Oxidative Phosphorylation in SNCA-3x dopaminergic Neurons

We have identified significant downregulation of various cellular processes associated with mitochondrial function in SNCA-3x DA neurons. These processes include oxidative phosphorylation (OXPHOS), ATP synthesis coupled with electron transport, and mitochondrial transport. These findings strongly suggest mitochondrial dysfunction in SNCA-3x DA neurons.

Animal models of PD harbouring SNCA A53T or A30P mutations have demonstrated the development of neuronal mitochondrial degeneration and subsequent cell death (Lee et al., 2006). Functional deficiencies in complex I have also been pinpointed in the substantia nigra and frontal cortex of PD patients (Li et al., 2021). Notably, similar complex I deficiencies have been observed in non-brain tissues such as skeletal muscles, platelets, and fibroblasts, indicating a systemic occurrence (Parker et al., 1989; Wiedemann et al., 1999). Furthermore, there have been varied and inconsistent reports regarding the function of complex III and complex IV in PD (Li et al., 2021).

A recent study by Choi et al. demonstrated that human induced pluripotent stem cell (iPSC)-derived neurons carrying the SNCA-A53T mutation exhibit impaired respiration, depolarized mitochondrial membrane potential, and increased oxidative stress. They also revealed that A53T oligomers induce the early opening of the mitochondrial permeability transition pore (mPTP), leading to subsequent toxicity (Choi et al., 2022). A prior study by our

group suggested downregulation of OXPHOS in iPSC derived A53T DA neurons (Fernandez et al., 2020). Bulk RNA sequencing performed on SNCA A53T and SNCA-3x iPSC-derived neurons confirmed perturbations in mitochondrial function. These findings were further supported by observed changes in mitochondrial morphology and membrane potential in these PD DA neurons compared to controls (Zombon et al., 2019).

Numerous genes implicated in PD, such as PINK1, PRKN, DJ1, LRRK2, and SNCA, have shown associations with OXPHOS, particularly complex I. Disruption of complex I leads to reduced ATP production in these neurons, ultimately triggering apoptosis (Zainul Ali et al., 2022). The aggregation of α-Synuclein has been linked to decreased complex I activity, increased ROS generation, inhibition of mitochondrial fusion, elevated unfolded protein levels, and increased Ca2+ influx, all contributing to cell death (Zainul Ali et al., 2022). Interestingly, ATP synthase operates more efficiently in the presence of monomeric synuclein, as it binds to and enhances ATP synthase efficiency (Ludtmann et al., 2016). The loss of α-Synuclein results in decreased ATP formation and membrane depolarization (Ludtmann et al., 2016).

PD is increasingly recognized as a mitochondrial disease, specifically involving complex I of the electron transport chain (ETC). However, recent studies have expanded our understanding, implicating UQCRC1 and CHCHD2 in PD, which are critical for the functions of complex III and complex IV, respectively (Li et al., 2021). In our study, consistent with these findings, we observed the downregulation of genes associated with multiple complexes of OXPHOS, including Complex I (ubiquinone oxidoreductase), Complex III (cytochrome c reductase), and Complex IV (cytochrome c oxidase) in PD SNCA-3x DA neurons. Notably, this is the first study to demonstrate the impact of the alpha-synuclein triplication mutation on multiple OXPHOS complexes. Furthermore, we determined that OXPHOS downregulation was primarily due to the downregulation of nuclear DNA-encoded mitochondrial genes, while mitochondrial DNA-encoded genes remained unaffected or even upregulated. Collectively, these results strongly support the notion that mitochondria may play a pivotal role in the pathogenesis of PD.

### 4.3.4 oDAn1 vs oDAn2 vs oDAn3

Many of the genes that were upregulated and downregulated in SNCA-3x oDAn1 cells exhibited a similar expression pattern in the other dopamine (DA) subsets, including an

upregulation of protein folding in the endoplasmic reticulum (ER) and a downregulation of oxidative phosphorylation (OXPHOS) and translation. This suggests a shared response to the SNCA-3x mutation across all three DA cell types. However, it is worth noting that several genes were specifically dysregulated in oDAn1 cells. These genes are associated with cellular component organization (such as MAP4 and PTPRS), endoplasmic reticulum stress (PDIA3 and HSPA5), protein ubiquitination (UBE2V1), oxidative stress (SIRT1, MSH2, and EPHA7), and apoptosis (MAPK1). These genes displayed higher expression levels in SNCA-3x oDAn1 cells compared to SNCA-3x oDAn2 and oDAn3 cells. The role of these genes in the vulnerability of oDAn1 cells, which were found to be selectively sensitive to rotenone, could be further evaluated.

### 4.3.5 Validation Experiments

To validate the phenotypic effects associated with the identified SNCA mutation using scRNA-seq, a subset of genes that were observed to be upregulated in dopamine (DA) neurons was selected for validation through qPCR. However, the qPCR analysis confirmed the upregulation of these genes in midbrain organoids, but the results obtained were statistically insignificant. This lack of validation through qPCR may be attributed to the fact that qPCR analysis was conducted on bulk RNA extracts from whole organoids, whereas the scRNA-seq study was performed at the single-cell level, specifically focusing on a particular subset of DA neurons. Furthermore, qPCR validation was carried out using a few replicates, and due to variations between organoid samples, the significance values were affected. In contrast, the scRNA-seq analysis involved the examination of thousands of individual cells, which provided a more robust statistical analysis.

We conducted a repeated scRNA-seq SNCA-3x experiment in an effort to replicate the SNCA-3x associated signatures. Upon comparing the top 200 upregulated and downregulated genes from both experiments, we found that only seven upregulated genes and twelve downregulated genes were common to both in SNCA-3x oDAn1 cells. Although the majority of the top 200 differentially upregulated and downregulated genes did not align between the two experiments, we did observe consistent expression patterns in numerous genes previously linked to SNCA-3x in oDAn1 cells. The disparities observed between the results of the two experiments may be attributed to well-known batch-to-batch variations in organoids and differences in the matrix used for HMO generation. In the first experiment, Matrigel was

employed in the derivation of midbrain organoids, while Geltrex was utilized in the second experiment. Notably, in both SNCA-corr and SNCA-3x midbrain organoids, the proportion of oDAn3 cells in the second experiment (Geltrex organoids) was almost half of that in the first experiment (Matrigel organoids). oDAn3 cells exhibited elevated levels of SNCA and ER stress markers. The discrepancy in the cell proportion of oDAn3 cells between the two experiments may have contributed to the incomplete validation of SNCA-3x disease signatures in the second experiment.

### 4.3.6 The Impact Of L-MOSES On Alpha-Synuclein Levels and Transcriptomic Characteristics of The SNCA-3x Mutation

In a previous research endeavour conducted by our group, we identified L-MOSES as an inhibitor of the lysine acetyltransferase KAT2B. L-MOSES showcased its ability to alleviate endoplasmic reticulum (ER) stress, attenuate the activation of the unfolded protein response (UPR), and prevent neuronal cell demise in cortical and dopaminergic neuronal model subjected to Tunicamycin treatment (Pavlou et al., 2022).

Given that the SNCA-3x mutation leads to the accumulation of misfolded alpha-synuclein within dopaminergic neurons, a condition known to be detrimental to these cells, we subjected 42-day-old SNCA-3x and SNCA-corr control organoids to a 7-day treatment with L-MOSES to evaluate its impact on alpha-synuclein protein levels. The Western blot analysis revealed that L-MOSES did not reduce alpha-synuclein levels in either SNCA-3x or Control HMOs.

However, at the transcriptomic level, L-MOSES exhibited a beneficial effect on genes associated with the SNCA-3x mutation in DA neurons. For instance, SNCA-3x neurons treated with L-MOSES exhibited a reduction in the expression of several genes linked to endoplasmic reticulum stress, oxidative stress, and apoptosis. Nevertheless, a contrasting pattern was observed for some of these genes in L-MOSES-treated SNCA-corr DA neurons. The underlying cause of the differing effects of L-MOSES in Control SNCA-corr DA neurons remains elusive. Concurrently, L-MOSES-treated SNCA-3x DA neurons displayed an elevation in the expression of several genes associated with translation, oxidative phosphorylation, and mitochondrial transport. These findings suggest that L-MOSES treatment appears to mitigate the transcriptional changes induced by the SNCA-3x mutation in SNCA-

3x DA neurons. Nevertheless, it is imperative to validate these results in future experiments to ensure their robustness.

In a broader context, both L-MOSES-treated SNCA-3x and SNCA-corr neurons exhibited an upregulation of lysosomal transport and Golgi vesicle transport. This implies that L-MOSES may exert a significant influence on protein processing and the degradation of misfolded proteins. Further investigation is warranted to evaluate the efficacy of L-MOSES in reversing changes associated with the SNCA-3x mutation.

### 4.3.7 Impact of Oxidative Stress on Dopaminergic Neurons

To investigate the impact of oxidative stress on HMOs, we subjected them to a medium lacking antioxidant, and supplemented with rotenone, a commonly used oxidative stress-inducing compound, for a 24-hour period. Rotenone, a naturally occurring pesticide, functions by inhibiting complex I of the mitochondrial electron transport chain, consequently increasing the production of reactive oxygen species (ROS) (Heinz et al., 2017). Rotenone has been demonstrated to induce Parkinson's disease-like pathological features, including oxidative damage, mitochondrial impairment, microglia activation, α-synuclein accumulation and aggregation leading to the formation of Lewy body-like inclusions, selective nigrostriatal dopaminergic degeneration, and L-DOPA-responsive motor deficits in animal models (Cannon et al., 2009).

Analysis of cell type proportions in rotenone-treated midbrain organoids revealed a significant reduction in oDAn1 cells within the treated HMOs. This trend was consistently observed across all three lines used in this experiment, including SNCA-3x HMOs. This finding aligns with our prior research and the study by Jerber et al., which both showed a significant depletion of dopaminergic (DA) neurons upon exposure to rotenone (Fernandez et al., 2020; Jerber et al., 2021).

Both the treatment and genotype factors exhibited a significant influence on Pseudotime scores, with the treatment demonstrating a p-value of < 2e-16 and the genotype showing a p-value of 1.42e-05. However, the interaction term between treatment and genotype did not reach statistical significance (p-value 0.057). Notably, the SNCA-3x mutation did not exacerbate the loss of oDAn1 cells caused by rotenone. This raises the possibility that the pre-existing

mitochondrial dysfunction associated with SNCA-3x may affect the production of ROS triggered by rotenone. To assess the genetic impact of the SNCA-3x mutation under these circumstances, oxidative stress should be induced independently of mitochondrial processes.

Enrichment analysis of rotenone-sensitive oDAn1 cells revealed an enrichment of genes associated with cholesterol biosynthesis, consistent with our previous study on SNCA A53T iPSC-derived DA neurons (Fernandez et al., 2020). Cholesterol dysregulation has been implicated in PD (Jin et al., 2019). Studies have shown that cholesterol synthesis, export, and utilization increase in the context of complex I deficiency. Schirris et al. demonstrated that boosting cholesterol production in mice with mitochondrial complex I deficiency reduces reductive stress, leading to improved motor function and increased survival (Schirris et al., 2021). High cholesterol has also been linked to dopaminergic neuronal loss in the MPTP animal model of PD (Paul et al., 2017). Given that rotenone is lipophilic and dopaminergic neurons are rich in cholesterol-rich membranes, this may contribute to the selective vulnerability of dopaminergic neurons to rotenone.

Rotenone-resistant oDAn1 cells exhibited enrichment in genes associated with various functions, including translation, dopaminergic neuron differentiation, and DNA repair. Intriguingly, these resilient cells also displayed higher expression of SOX2, LMX1A, SHH, and LMO4, suggesting that immature DA neurons had a higher survival rate under rotenone conditions.

Pathways upregulated in response to rotenone treatment included translation-related processes and the unfolded protein response (UPR) pathway. Previous studies have shown that rotenone induces the UPR response in cellular and animal models (Tong et al., 2016; Ryu et al., 2002; Han et al., 2014). In oDAn1 cells, several mitochondrial DNA-encoded genes and cholesterol biosynthesis-related genes were downregulated in response to rotenone. Interestingly, mitochondrial DNA-encoded genes were unaffected or even upregulated in response to the SNCA-3x mutation, nuclear DNA-encoded OXPHOS genes were observed to be downregulated. Notably, the dopaminergic neuronal marker NR4A2 and the neuronal maturity marker MAP2 were repressed in rotenone-treated oDAn1 cells. The concurrent upregulation of VIM and downregulation of MAP2 suggest that rotenone had a significant impact on mature neurons.

### 4.3.8 Chimera Midbrain Organoids

In this study, we have also demonstrated the feasibility of differentiating induced pluripotent stem (iPS) cells from multiple donors within the same tissue, which we have termed "chimera midbrain organoids." Human iPSC-derived neuronal models have been shown to exhibit high heterogeneity and variability (Volpato et al., 2018; Schwartzentruber et al., 2017). The utilization of chimeric organoids offers several advantages over monolayer dopaminergic neuronal cultures. The incorporation of cells from multiple donors within a single tissue reduces technical variability, lowers experimental costs, and enhances scalability. The chimera organoids generated in this study exhibited comparable immunohistochemistry and single-cell RNA sequencing (scRNA-seq) results to monogenic organoid cultures, suggesting that co-culturing human iPSCs from multiple donors does not impede the differentiation potential of the culture. The co-culture of human iPSCs from multiple donors for neuronal differentiation has been previously demonstrated by Jerber et al. (Jerber et al., 2021). The primary challenge associated with this approach is the drift in the proportion of donor cells from the starting point to the endpoint. In our study, the proportion of donor cells underwent significant changes compared to the initial 1:1 ratio. Technical improvements, such as refining differentiation methods and potentially aligning the growth rates of donor cells, must be considered to optimize the utilization of this chimera system.

To model the non-cell-autonomous effects on dopaminergic neurons, we generated chimeric organoids comprising both healthy and disease-derived cells. Despite the variation in the proportion of cells recovered from each donor within the chimera organoids, we obtained a sufficient number of cells to investigate cell-autonomous effects. scRNA-seq analysis revealed that when healthy dopaminergic neurons were co-cultured alongside SNCA-3x cells, they did not entirely replicate the mutational signatures of SNCA-3x, such as the upregulation of ER stress and the downregulation of translation and OXPHOS. Nevertheless, the co-cultivation of these healthy neurons with SNCA-3x cells exhibited indications of deleterious effects, such as the upregulation of stress and DNA damage response, as well as the downregulation of synapse assembly and axon guidance. These results suggest that chimera organoids have the potential to be used to study non-cell autonomous effects. Furthermore, the chimera brain organoid model would be an excellent tool for investigating the transfer of alpha-synuclein between cells and its prion-like propagation.

# 5. Induced Dopaminergic Neurons

## 5.1 Background

The iPSC-derived monolayer neuronal models have revealed a significant challenge due to undesired cellular heterogeneity and noticeable variability across different batches. In-depth analysis through single-cell RNA sequencing of the dopaminergic neuronal cultures, established using the well-known Kriks et al. 2011 protocol, highlighted substantial heterogeneity within the cells. Moreover, the distinct neuronal subtypes present in the culture exhibited diverse transcriptional profiles and varying sensitivity to genetic and toxin-induced stress.

Furthermore, the efficiency of differentiating dopamine neurons using existing protocols shows a notable range (10-70%). These protocols are financially demanding due to the requirement for small molecules/morphogens and are time-consuming to yield a satisfactory quantity of highly pure dopaminergic neurons. Intriguingly, the cultures resulting from these differentiation protocols do not specifically consist of substantia nigra A9 neurons; instead, they encompass various dopaminergic neuronal subtypes, including ventral tegmental A10 neurons. Despite the strides made in this field, achieving the efficient generation of a specific subtype of midbrain dopamine neurons remains a formidable challenge.

Researchers have explored the use of a forced expression approach to tackle these issues, but with limited success. The introduction of essential transcription factors for dopaminergic neurons, such as LMX1A, FOXA2, NURR1, PITX3, ASCL1, and MSX1, has been demonstrated to yield dopamine neurons. However, most of these protocols exhibit inefficiencies and necessitate the use of expensive small molecules/morphogens.

To address these persistent challenges, there is a pressing need to develop an effective protocol capable of generating highly homogeneous and fully characterized A9 substantia nigra dopaminergic neurons on a larger scale within a shorter timeframe.

Here, our study recapitulates a significant finding: the sole ectopic expression of ATOH1 is sufficient to initiate the emergence of TH-positive induced neurons (iNeurons). In fact, we

found it to be indispensable in generating TH-positive neurons in our study. Through the application of scRNA-seq analysis, both ATOH1+ 6TFs and NGN2+6TFs cultures revealed the presence of TH-positive iNeurons. However, it's noteworthy that the TH-positive iNeurons derived from the ATOH1+ 6TFs condition also exhibited markers characteristic of the peripheral nervous system. Conversely, the TH-positive iNeurons resulting from NGN2+6TFs displayed the expression of markers associated with GABAergic, glutamatergic, and serotonergic neurons.

When considering ATOH1 alone and its ectopic expression in conjunction with shh+ FGF8b treatment, the outcome is the induction of TH-positive iNeurons characterized by features of the peripheral nervous system. Interestingly, the co-expression of NGN2, FOXA2, LMX1A, and LMO3 appears to hinder the progression of neuronal differentiation. Furthermore, the ectopic expression of ATOH1 or NGN2 along with other TFs using inducible PiggyBac vectors did not lead to the generation of TH-positive iNeurons.

Our investigation included a screening process aimed at streamlining the number of TFs required for inducing TH-positive iNeurons. This screening process highlighted the ATOH1_LMX1A_NURR1_PITX3 combination as a promising candidate.

## 5.2 Results

### 5.2.1 Induced Dopaminergic Neurons Differentiation Using Standard PiggyBac Vectors

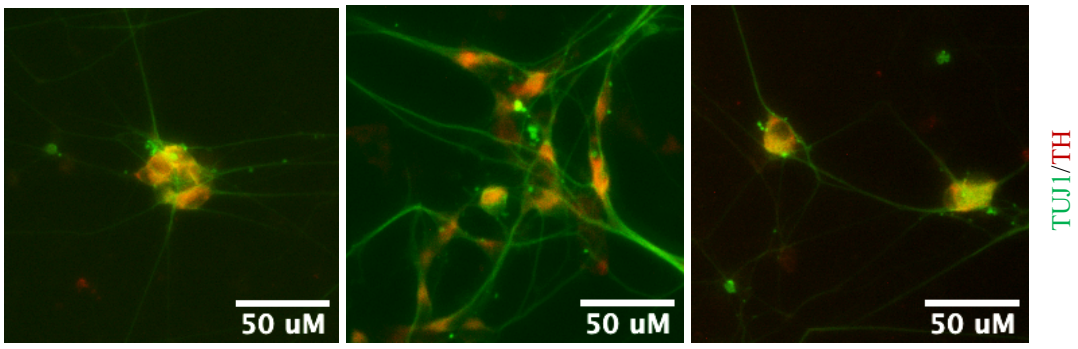



**Figure 5.1:** Immunocytochemical analysis of day 14 dopamine iNeurons for TH and TUJ1 under various conditions. The TF combinations include: 4 TFs - FOXA2, LMX1A, NGN2, and ATOH1; 8 TFs - FOXA2, LMX1A, NGN2, ATOH1, SOX6, NR4A2, MSX1, and PITX3; 7 TFs - FOXA2, LMX1A, NGN2, SOX6, NR4A2, MSX1, and PITX3.

Several transcription factors (TFs) were selected based on current knowledge of dopamine neuron developmental biology for screening, aiming to produce highly homogeneous substantia nigra A9 neurons through a forward reprogramming approach. The selected TFs include FOXA2, LMX1A, ATOH1, NGN2, MSX1, SOX6, NR4A2, LMO3, and PITX3. FOXA2, LMX1A, and MSX1 are required for midbrain floor plate specification. LMO3 and SOX6 play roles in substantia nigra A9 neuron specification. NGN2, NR4A2, and PITX3 are essential for dopamine neuron differentiation. ATOH1 was chosen based on the study by Sagal et al., 2014, wherein the authors demonstrated that ectopic expression of ATOH1 alone in pluripotent stem cells (PSCs) suffices to generate dopaminergic neurons.

Standard PiggyBac plasmids were designed, each containing a single TF per construct, and were obtained from the VectorBuilder company. The gene of interest (GOI) was driven by the EF1 alpha promoter. Additionally, a BFP reporter and Blasticidin selection marker were included to estimate transfection efficiency and select transfected cells, respectively. Subsequently, the hiPSCs were transfected and differentiated according to the experimental timeline depicted in the figure. The TF plasmids were utilized in various combinations to determine the most optimal set for differentiating hiPSCs into dopaminergic neurons. Three

experimental conditions were established:1) 4 TFs: ATOH1, NGN2, FOXA2, LMX1A, 2) 8 TFs: ATOH1, NGN2, FOXA2, LMX1A, NR4A2, PITX3, SOX6, and MSX1, 3)7 TFs: ATOH1, NGN2, FOXA2, LMX1A, NR4A2, PITX3, SOX6, and MSX1.

On day 14 post-transfection, the cells were fixed and subjected to immunostaining for TH (a dopamine neuron marker) and TUJ1 (a pan-neuronal marker) to confirm dopaminergic neuron differentiation. The immunostaining validated the presence of TH-positive cells across all three conditions (as shown in the Figure). These findings imply that hiPSCs could be differentiated into dopamine neurons by overexpressing lineage-specific TFs. However, due to the uniform selection marker and transient overexpression of all TF constructs, it remains challenging to pinpoint the precise combination of TFs responsible for the differentiation of dopaminergic neurons.

### 5.2.2 Induced Dopaminergic Neurons Differentiation Using NGN2 Polycistronic Constructs via Transient Expression

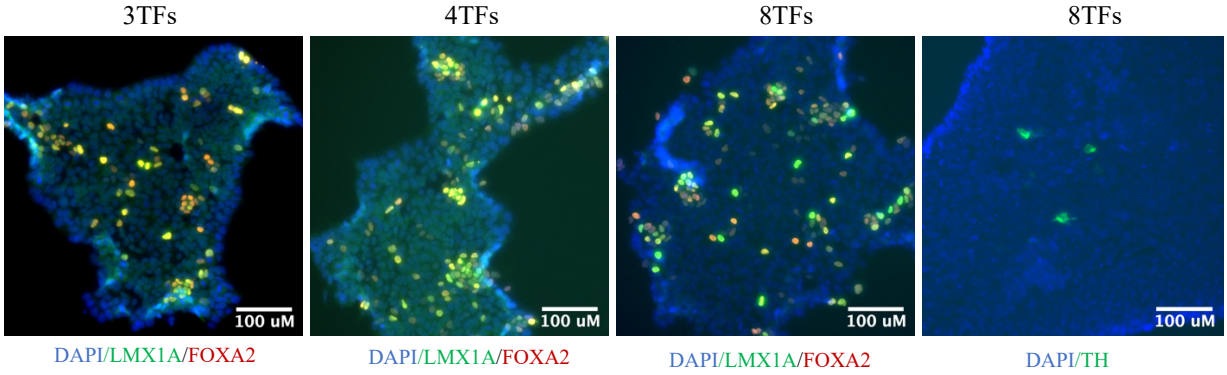



**Figure 5.2**: Immunostaining of dopamine iNeurons in the 8 TFs condition for TH expression on the sixth day following DOX induction. The three TF combination consists of FOXA2, LMX1A, and NGN2; the four TF combination comprises FOXA2, LMX1A, NGN2, and LMO3; and the eight TF combination includes FOXA2, LMX1A, NGN2, LMO3, SOX6, NR4A2, MSX1, and PITX3.

To further delve into the optimal combination of transcription factors (TFs) necessary for generating dopaminergic induced neurons (DA iNeurons), we devised a strategy employing low-leak tetracycline-inducible PiggyBac plasmids, encompassing multiple TFs. These specialized plasmids were intricately designed and procured from VectorBuilder company. The

genes of interest (GOIs) were driven by the TRE promoter, tetracycline-controlled transcriptional silencer (tTS) gene driven by a ubiquitous promoter, reverse tetracycline-responsive transcriptional activator M2 (rtTA) gene driven by an EF1 alpha promoter. Three distinct constructs were used in this experiment: 1) NGN2_FOXA2_LMX1A 2) NGN2_FOXA2_LMX1A_LMO3 3) SOX6, NR4A2, MSX1_PITX3. Notably, existing studies have unveiled the preferential expression of LMO3 in dopaminergic neurons within the substantia nigra pars compacta. hiPSCs were subjected to transfection with the constructs in diverse combinations: Condition 1: Only constructs 1; Condition 2: Only constructs 2; Condition 3: Constructs 2 and 3. The TFs expression was effectively induced using doxycycline (DOX), 24 hours after transfection. To validate the construct's expression, the transfected cells were fixed and immunostained two days after DOX treatment. Positive immunostaining of FOXA2 and LMX1A confirmed the expression of constructs 1 and 2, whereas the affirmative detection of SOX6 staining served as confirmation for the expression of construct 3. At this phase of differentiation, the transfection efficiency was estimated to range between 10% and 20%, as inferred from immunostaining results.

On day 6, the cells were fixed and immunostained for TH (a marker of dopamine neurons). TH-positive cells were only identified in condition 3, wherein cells were transfected with the full complement of 8 TFs. Evidently, the limited or absent TH-positive cells observed in the other conditions may be attributed to a combination of factors, including low transfection efficiency and the presence of proliferating cells that outpace neuronal cell growth.

### 5.2.3 Induced Dopaminergic Neurons Differentiation Using Subcloned iPSCs with NGN2 +3TFs

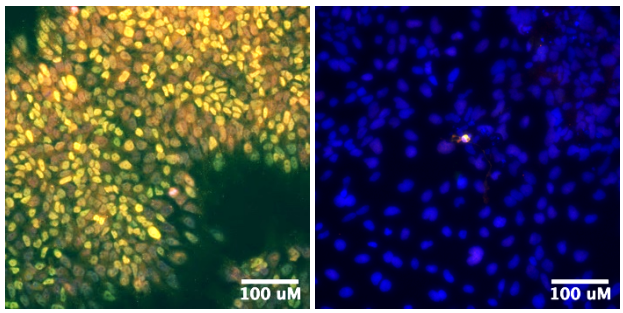



**Figure 5.4:** Immunocytochemical analysis of puromycin-selected iPSCs with NGN2+3TFs (FOXA2_LMX1A_LMO3) upon DOX induction. The left image depicts LMX1A and FOXA2

expression within the pool two days post DOX treatment. The right image illustrates TH expression on day 15.

An NGN2 polycistronic construct with a puromycin selection marker was devised to enhance the expression efficiency of the TFs throughout the culture. The puromycin selection marker gene was integrated downstream of the rtTA gene, separated by a 2A peptide sequence. iPSCs were transfected with the TRE_NGN2_FOXA2_LMX1A_LMO3 construct and were subsequently subjected to selection with puromycin. The iPSCs were subcloned and induced with DOX to initiate the expression of the TFs. Within 48 hours post DOX treatment, immunostaining was conducted targeting LMX1A and FOXA2, to confirm the expression of the construct. Strikingly, most of the cells exhibited robust FOXA2 and LMX1A expression.

Following a two-week differentiation period, the cells were fixed and subjected to TH marker staining to gauge the quantity of dopaminergic neurons. A small number of TH positive cells were detected, surrounded by non-neuronal cells. FOXA2 and LMX1A expression was observed in significant number of cells on day 2 which suggested that cells transfected with the construct unable to differentiate into dopaminergic neurons.

### 5.2.4 Induced Dopaminergic Neurons Differentiation Using Subcloned BON NGN2 iPSCs Utilizing NGN2+3TFs

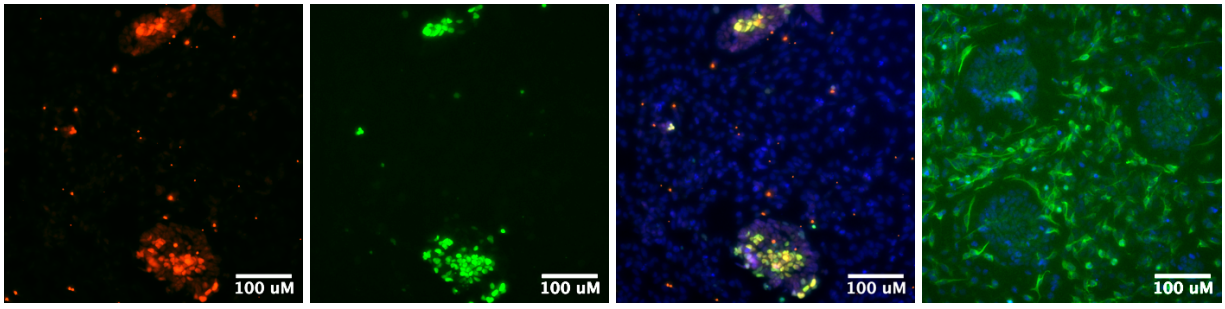



**Figure 5.5**: Immunocytochemical analysis of clonal BOB NGN2 iPSCs transfected with NGN2+3TFs (FOXA2_LMX1A_LMO3) following DOX induction. Representative images display the expression of LMX1A, FOXA2, and TUJ1 on the second day after DOX induction.

A well-established NGN2-OPTi-OX iPSC clonal line, known for its capacity to yield homogeneous cortical glutamergic neurons upon DOX induction, was harnessed to enhance

neuronal differentiation efficiency. The NGN2 cell line was subjected to transfection with a construct carrying NGN2_FOXA2_LMX1A_LMO3 and subsequently differentiated as detailed earlier. The introduction of DOX into the culture precipitated the overexpression of the TFs, including NGN2. After 48 hours of DOX treatment, cells were fixed and subjected to immunostaining for FOXA2, LMX1A, and TUJ1, effectively to confirm the expression of the transgenes and the initiation of neuronal differentiation. Distinct aggregates of cells displayed evident FOXA2 and LMX1A expression, encircled by TUJ1-positive cells that, intriguingly, lacked expression of the neuronal marker TUJ1. This compelling observation hints at the possibility that the combined presence of the TFs might potentially hinder the progression of iPSCs towards neuronal differentiation.

### 5.2.5 A Screen Aims to Identify Optimal Combination Of TF For Induced Dopaminergic Neurons Differentiation

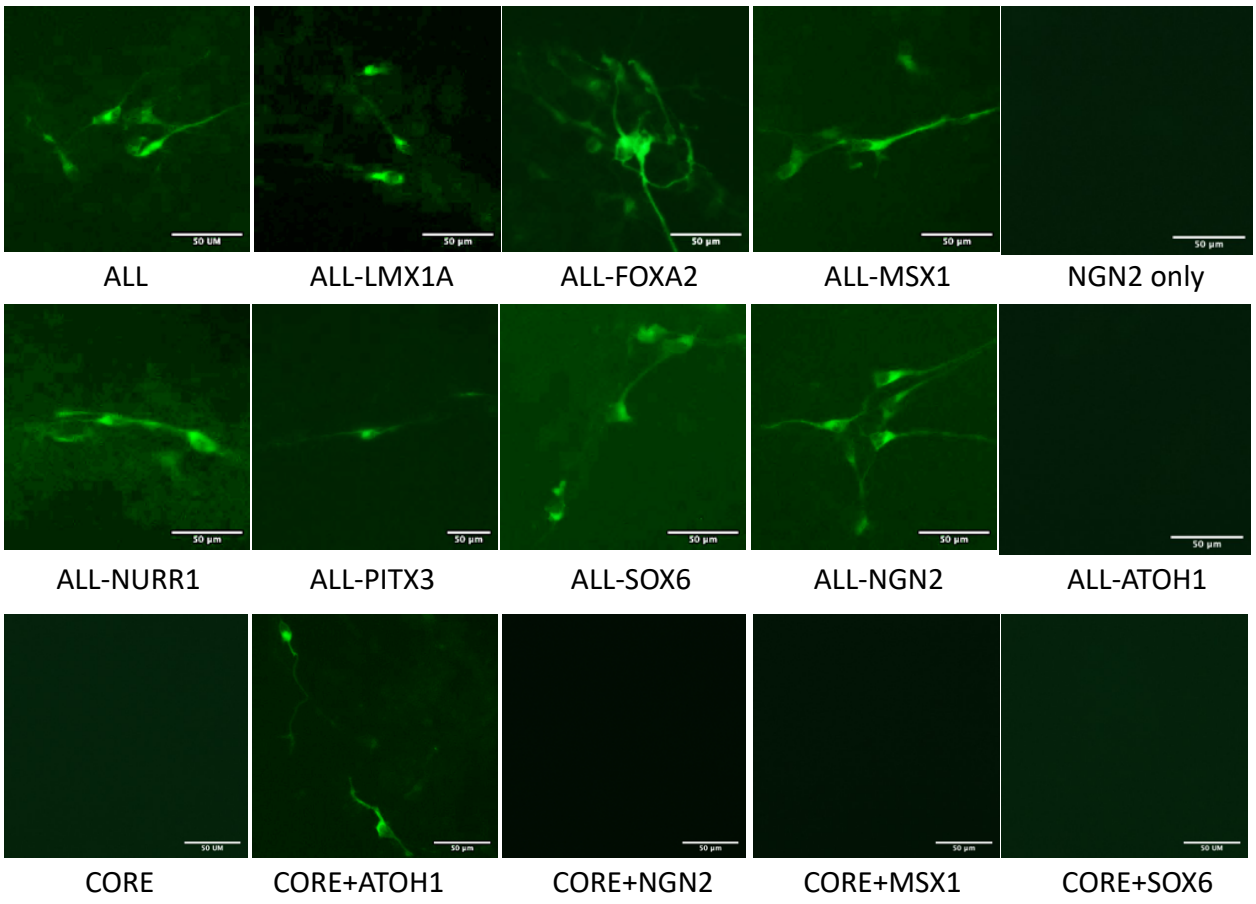



**Figure 5.6:** Immunostaining images depicting TH expression under diverse conditions aimed at identifying the most optimal TF combination. “All” condition includes: FOXA2, LMX1A,

PITX3, NURR1, SOX6, MSX1, NGN2, and ATOH1. “CORE” includes: FOXA2, LMX1A, NURR1, and PITX3.

A screening protocol was designed with the aim of pinpointing the minimum number of TFs required to generate TH expressing dopaminergic neurons. This comprehensive screening encompassed a total of 15 distinct conditions, which encompassed the presence or absence of specific TFs. Among these, there was an encompassing "All" condition consisting of ATOH1, NGN2, FOXA2, LMX1A, NURR1, PITX3, SOX6, and MSX1. The other conditions involved the exclusion of individual TFs, resulting in cells being transfected with 7 TFs. The experiment also included NGN2-only and ATOH1-only conditions, along with a “CORE” condition that comprised 4 TFs (FOXA2, LMX1A, NURR1, PITX3). Additionally, the remaining 4 TFs were sequentially integrated into the CORE group, one at a time, resulting in a total of 5 TFs.

TH-positive cells were observed in conditions encompassing the presence of all 8 TFs, as well as in those with 7 TFs (-FOXA2), 7 TFs (-LMX1A), 7 TFs (-NURR1), 7 TFs (-PITX3), 7 TFs (-SOX6), 7 TFs (-MSX1), CORE (+ATOH1), and ATOH1-only conditions. Conversely, no TH-positive cells were detected in conditions involving NGN2-only, CORE (+SOX6), CORE (+MSX1), and CORE (+NGN2). These results clearly underscores that conditions lacking ATOH1 overexpression were unable to yield TH-positive dopaminergic neurons. Moreover, the pivotal role of ATOH1 overexpression in driving dopaminergic neuron differentiation within this study becomes increasingly apparent.

### 5.2.6 Induced Dopaminergic Neurons Differentiation Using ATOH1 Polycistronic Vectors via Inducible Expression

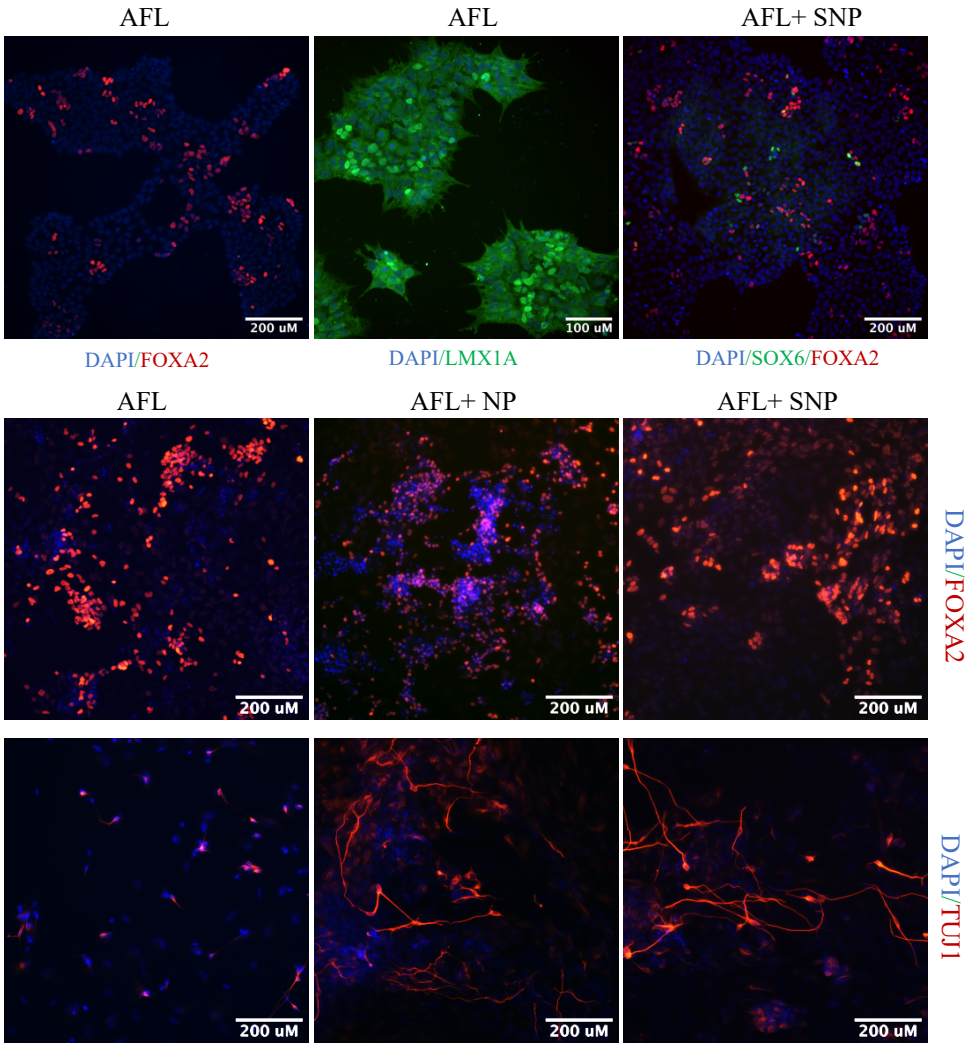



**Figure 5.7:** Immunostaining images illustrating the expression of FOXA2, LMX1A, and SOX6 on the second day after DOX induction (A), as well as FOXA2 and TUJ1 on day 15 (B), within puromycin-selected iPSC pools using various combinations of polycistronic piggyBac plasmids following DOX induction. AFL= ATOH1 + FOXA2_LMX1A; SNP= SOX6 + NURR1_PITX3; NP= NURR1_PITX3

Based on the outcomes of the TF screen, the proneuronal TF NGN2 was substituted with the ATOH1 gene. Tetracycline-inducible polycistronic PiggyBac plasmids were designed to incorporate ATOH1, while a puromycin selection marker was introduced downstream of the rtTA gene. Additionally, two supplementary constructs were conceived to enable NURR1, PITX3, SOX6 expression, each accompanied by a blasticidin selection marker. All these

plasmids were procured from the VectorBuilder company. In the experimental paradigm, iPSCs were subjected to transfection with these constructs under three distinct conditions: 1) ATOH1_FOXA2_LMX1A 2) ATOH1_FOXA2_LMX1A & NURR1_PITX3, 3) ATOH1_FOXA2_LMX1A & SOX6_NURR1_PITX3.

DOX was subsequently introduced to initiate the expression of the TFs. Cells were fixed and stained 48 hours after DOX treatment to validate the expression of the constructs. Within the ATOH1_FOXA2_LMX1A transfected cells, the presence of LMX1A and FOXA2 positive cells was evident. Similarly, the transfection of SOX6_NURR1_PITX3 showed SOX6 expression. However, the transfection efficiency remained suboptimal across all conditions.

To eliminate non-transfected cells, antibiotic selection was carried out. The selected cells were then subjected to a differentiation period of two weeks. Following differentiation, the cells were fixed and subjected to staining for FOXA2, TH, and TUJ1. The expression of FOXA2 was noticeable across all conditions, affirming the presence of transfected cells at this stage. Moreover, the expression of the neuronal marker TUJ1 was also observed, thereby confirming neuronal differentiation. Nonetheless, the absence of TH-positive cells indicated that the differentiated neurons failed to attain dopaminergic identity.

### 5.2.7 Induced Dopaminergic Neurons Differentiation Using ATOH1 Monocycistronic Vector Via Inducible Expression

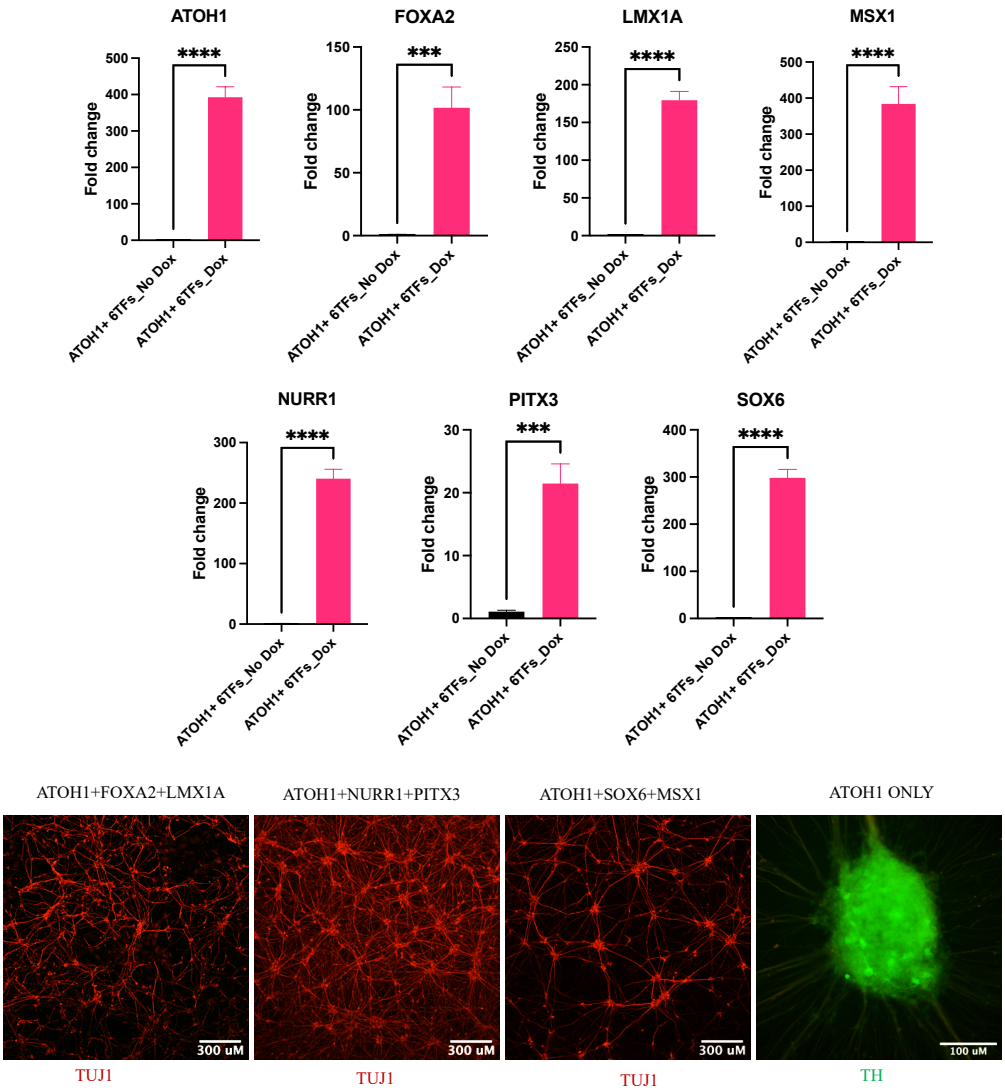



**Figure 5.8:** (A) Bar plots presenting qPCR analysis results for the expression of LMX1A, FOXA2, NURR1, PITX3, SOX6, and MSX1 in ATOH1+ 6TFs puromycin-selected iPSC pool 48 Hours post-DOX induction. The data is represented as Mean ± SEM, n = 3. Statistical analysis was performed using One-Way ANOVA with Dunnett's Multiple comparisons test. Significance levels: *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001, in comparison to the Control (iPSCs). (C) Immunostaining images illustrating TUJ1 expression in ATOH1+ FOXA2_LMX1A, ATOH1 + NURR1_PITX3, ATOH1+ SOX6_MSX1, and TH expression in ATOH1 alone iPSCs pool after 2 weeks of differentiation.

To enhance the efficiency of neuronal differentiation in the iPSCs overexpressing the proneural ATOH1 gene, a different strategy was adopted. The proneural ATOH1 was

segregated into a distinct inducible piggyBac plasmid, accompanied by a puromycin selection marker. Additionally, individual plasmids were devised for LMX1A and FOXA2 (with a neomycin marker), NURR1 and PITX3 (with a blasticidin marker), and SOX6 and MSX1 (with a hygromycin marker). Each plasmid incorporated two TFs, reflecting a total of 2 TFs per plasmid. All these plasmids were meticulously designed and obtained through the VectorBuilder company.

iPSCs were transfected with these plasmids, each harboring distinct combinations of TFs, to explore the potential for dopaminergic neuronal differentiation. The various tested conditions encompassed: 1) ATOH1 only, 2) ATOH1 + FOXA2 + LMX1A (AFL), 3) ATOH1 + NURR1 + PITX3 (ANP), 4)ATOH1 + SOX6 + MSX1 (ASM), 5) ATOH1 + FOXA2 + LMX1A + 6) NURR1 + PITX3 (AFLNP), 7)ATOH1 + FOXA2 + LMX1A + SOX6 + MSX1 (AFLSM), 8) ATOH1 + NURR1 + PITX3 + SOX6 + MSX1 (ANPSM), 9) ATOH1 + FOXA2 + LMX1A + NURR1 + PITX3 + SOX6 + MSX1 (AFLNPSM).

The expression of the constructs was verified by treating the AFLNPSM iPSCs with DOX for two days. Post DOX treatment, the cells were harvested and subjected to quantitative PCR (qPCR), revealing a marked overexpression of all the TFs.

Among the tested conditions, ATOH1 only, AFL, ASM, APN, AFLNP, and AFLNPSM conditions exhibited a high abundance of neurons. In the AFL and APN cultures, non-neuronal cells were also present, while the ASM condition demonstrated the presence of non-neuronal clumps. Interestingly, both APN and AFLNP cultures displayed the absence of non-neuronal contaminating cells. Intriguingly, neither the AFLSM nor the ANPSM condition yielded neurons. Only the ATOH1 only condition yielded TH-positive neurons

### 5.2.8 Induced Dopaminergic Neurons Differentiation Using Subcloned iPSCs With ATOH1+6TFs

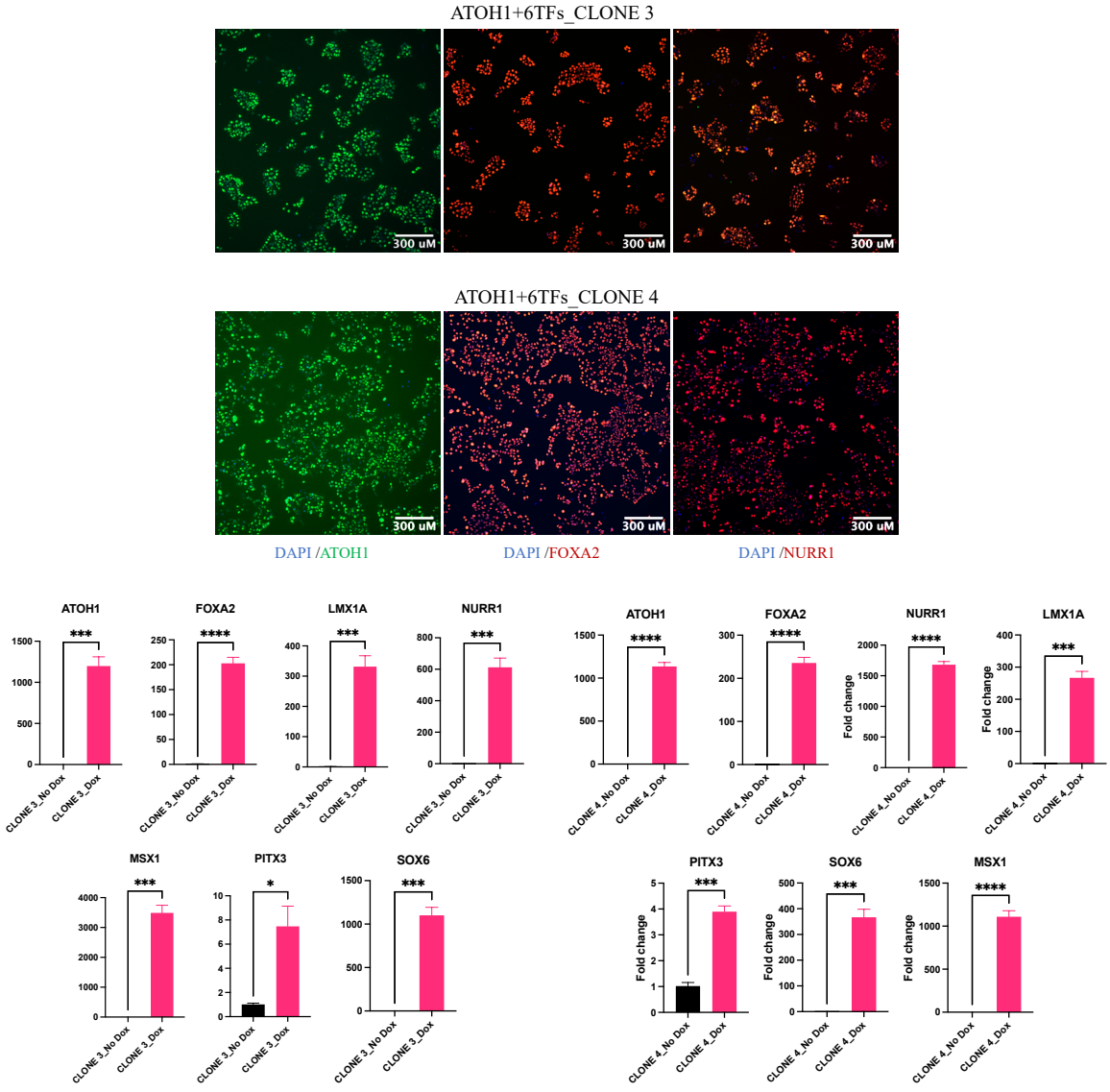



**Figure 5.9**: (A) Immunostaining images illustrating expression of ATOH1, FOXA2, LMX1A, NURR1 in ATOH1+6TFs subcloned iPSCs. 6TFs includes: FOXA2, LMX1A, SOX6, MSX1, NURR1, PITX3. (B) Bar plots presenting qPCR analysis results for the expression of LMX1A, FOXA2, NURR1, PITX3, SOX6, and MSX1 in ATOH1+ 6TFs puromycin-selected subcloned iPSC with ATOH1+6TFs 48 Hours post-DOX induction. The data is represented as Mean ± SEM, n = 3. Statistical analysis was performed using One-Way ANOVA with Dunnett's Multiple comparisons test. Significance levels: *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001, in comparison to the Control (iPSCs).

To further probe the potential of the inducible AFLNPSM iPSCs to differentiate into dopaminergic neurons, subcloning of the iPSCs was undertaken. This process led to the identification of two positive clones wherein all seven TFs were successfully integrated into the host genome, a determination validated through PCR analysis. Subsequent verification involved inducing the overexpression of the TFs via a 2-day DOX treatment, followed by qPCR analysis. This comprehensive assessment conclusively established the overexpression of all seven TFs within both identified clones. Immunostaining confirmed the uniform expression of the TFs within these cloned iPSC cell lines. Despite these affirmative findings, it is noteworthy that neither of the iPSC clones managed to yield TH-positive cells upon undergoing differentiation.

### 5.2.9 A Screen To Identify Optimal Combination Of TF For Induced DA Neurons Differentiation

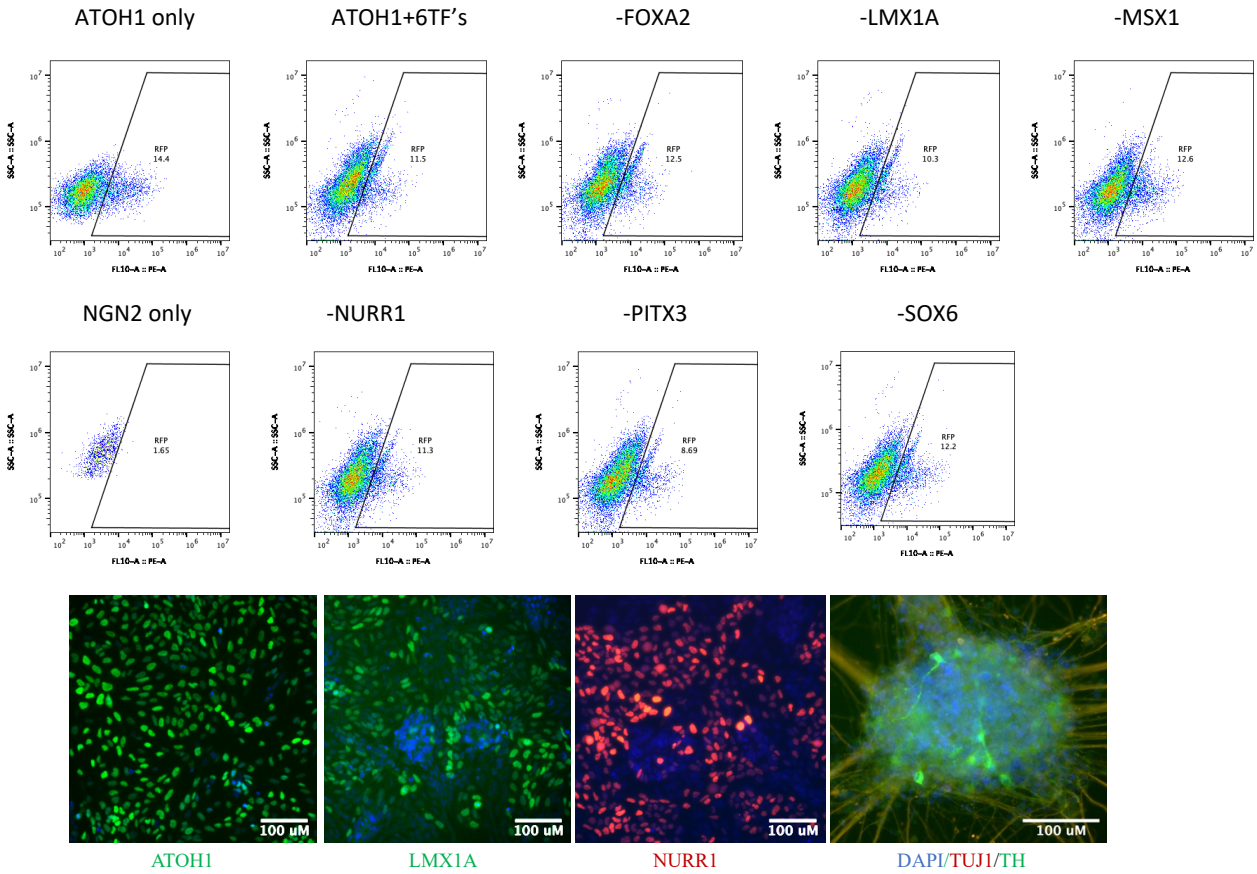



**Figure 5.10**: Flow cytometry analysis of DA iNeurons for RFP-positive cells generated using various TF combinations aimed at minimizing the number of TFs required for DA iNeuron generation. The 6TFs include FOXA2, LMX1A, NURR1, PITX3, SOX6, and MSX1. The

minus one TF condition represents the ATOH1+ 6TFs condition without the annotated TF. ATOH1 alone and NGN2 alone TFs were used as positive and negative controls, respectively. The iPSCs cell line utilized in this experiment features the TH gene tagged with RFP, allowing quantification of TH-positive cells through RFP expression measurement. (B) Immunostaining images illustrate the expression of ATOH1, LMX1A, and NURR1 48 hours post-DOX induction, alongside TH and TUJ1 expression 2 weeks into differentiation, within the puromycin-selected iPSCs pool with ATOH1+LMX1A+NURR1_PITX3.

A screen was designed to identify to reduce the number of TF's required to generate TH expressing dopaminergic neurons. This screening encompassed nine distinct conditions. Among these, the "ATOH1 + 6TFs" condition included FOXA2, LMX1A, NURR1, PITX3, SOX6, and MSX1. In other scenarios, a TF was omitted, and cells were essentially transfected with "ATOH1 + 5TFs." This study also featured NGN2-only and ATOH1-only conditions, which served as negative and positive controls, respectively. Flow cytometry was employed to gauge TH-positive cells, quantifying those displaying RFP positivity. Remarkably, each condition yielded TH-positive cells. The percentage of TH-positive cells in the "ATOH1 + 6TFs" condition was approximately 11.5%. The exclusion of SOX6, MSX1, or FOXA2 led to increased numbers of TH-positive cells, measuring 12.2%, 12.5%, and 12.6%, respectively. Conversely, the absence of LMX1A, NURR1, or PITX3 resulted in decreased numbers of TH-positive cells, at 10.3%, 11.3%, and 8.69%, respectively. These outcomes guided the selection of ATOH1, LMX1A, NURR1, and PITX3 for further exploration. Subsequent experiments lent further support to this choice, highlighting that this combination generated a comparable number of TH-positive cells when compared to the ATOH1-only condition.

A pool of inducible iPSCs overexpressing ATOH1, LMX1A, PITX3, and NURR1 was generated through the utilization of inducible piggyBac constructs, which featured antibiotic selection markers. The efficacy of TF overexpression was validated through immunostaining of DOX-treated cells, revealing widespread expression of ATOH1, NURR1, and LMX1A. Post differentiation, a uniformly neuronal culture was established, a fact corroborated by TUJ1 staining. However, it is worth noting that only a limited number of neurons exhibited TH expression.

## 5.2.10 Single Cell RNA Sequencing Analysis Of Induced DA Neurons Derived Using NGN2+6TFs And ATOH1 + 6TFs Combination

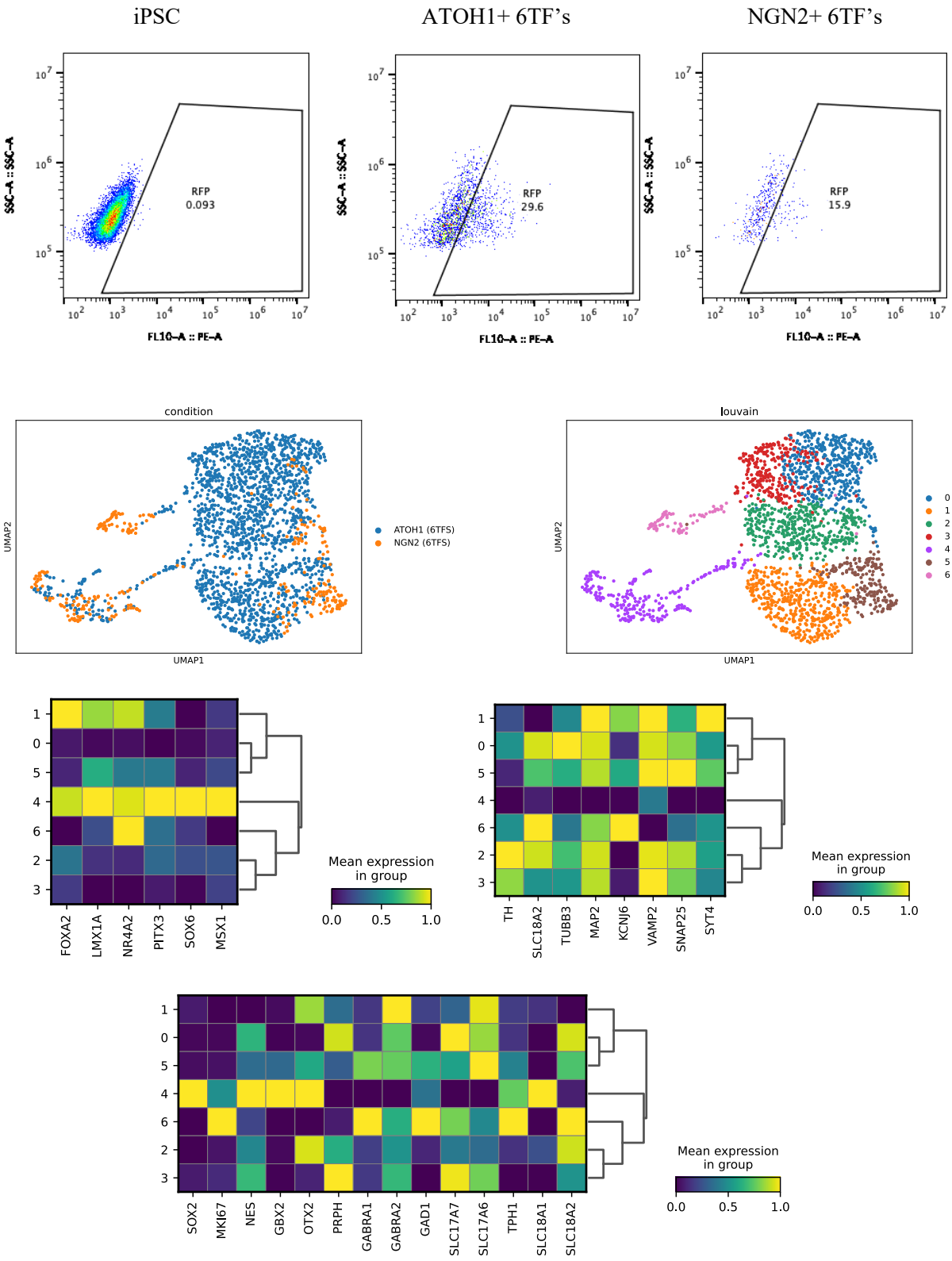


**Figure 5.11**: (A) Flow cytometry analysis of DA iNeurons for RFP-Positive cells generated using combinations of NGN2 + 6TFs and ATOH1 + 6TFs. The 6TFs include FOXA2, LMX1A, NURR1, PITX3, SOX6, and MSX1. The iPSCs cell line utilized in this experiment featured TH gene tagged with RFP, enabling the quantification of TH-positive cells through RFP expression measurement. (B) UMAP embedding of single cells originating from conditions ATOH1 + 6TFs and NGN2 + 6TFs is displayed. (C) Matrix Plots illustrates the Log2-Normalized mean expression of genes associated with various neuronal subtypes, alongside markers for different brain regions: OTX2 for VTA, GBX2 for hindbrain, and FOXG1 for forebrain. UMAP and matrix plots were generated by Nikos Patikas.

To delve further into the characterization of TH-positive iNeurons generated through overexpression of the proneural ATOH1 gene or the NGN2 gene, in conjunction with the 6TFs (FOXA2, LMX1A, NURR1, PITX3, SOX6, and MSX1) via standard piggyBac plasmids, the cells were analysed using single-cell RNA sequencing. For this experiment, the transient overexpression of the TFs was achieved, and the cells were selected with blasticidin and mitomycin selection before processing for scRNA-seq .

To qualitatively confirm the presence of TH-positive cells, immunostaining was conducted. The quantification of TH-positive cells within the cultures was performed using flow cytometry, measuring the number of cells exhibiting red fluorescence protein (RFP) positivity. In the ATOH1 + 6TFs and NGN2 + 6TFs cultures, the estimated percentages of TH-positive cells were approximately 30% and 15%, respectively. This distinction hinted at the increased efficiency of the ATOH1 + 6TFs combination in generating dopaminergic neurons, compared to the NGN2 + 6TFs approach.

With the presence of TH-positive cells confirmed, subsequently single-cell RNA sequencing was conducted. The resultant data was integrated with well-established NGN2-induced cortical neurons and human midbrain organoids datasets, enabling a thorough investigation into the cellular similarities across various cultures. Notably, human midbrain organoids have been recognized to contain *in vivo*-like dopaminergic neurons (DaN1), relevant to Parkinson's disease (PD). Upon comparative analysis, it was evident that cells stemming from the ATOH1 + 6TFs combination exhibited a striking resemblance to the midbrain organoids' DaN1 neurons.

The clustering of the unified dataset encompassing both ATOH1+6TFs and NGN2+6TFs conditions revealed six distinct clusters. Notably, clusters 4 and 5 exhibited a mixture of cells from both conditions, while cluster 6 predominantly consisted of cells from the NGN2+6TFs condition. At the gene expression level, cluster 4 demonstrated high expression of all six TFs, whereas cluster 1 solely expressed FOXA2, LMX1A, and NURR1. All clusters exhibited the expression of neuronal markers such as MAP2, TUBB3, SNAP25, and SYT4, excluding cluster 4. Interestingly, cluster 4 was characterized by the expression of progenitor markers and the absence of neuronal markers, strongly suggesting a progenitor identity. TH expression was notably observed in cluster 0, cluster 6, cluster 2, and cluster 3. Intriguingly, these clusters also displayed expression of the peripheral nervous system marker PRPH, except for cluster 6. It is noteworthy that most cells within cluster 6 emanated from the NGN2+6TFs condition. Cluster 6 cells exhibited the presence of markers typically associated with GABAergic neurons (GABRA1), glutamatergic neurons (GAD1), as well as serotonergic neurons (TPH1).

### 5.2.11 Characterization of Induced Dopaminergic Neurons Derived By Transient Overexpression Of NGN2 Alone, NGN2+6TFs, ATOH1 Alone, ATOH1+6TFS

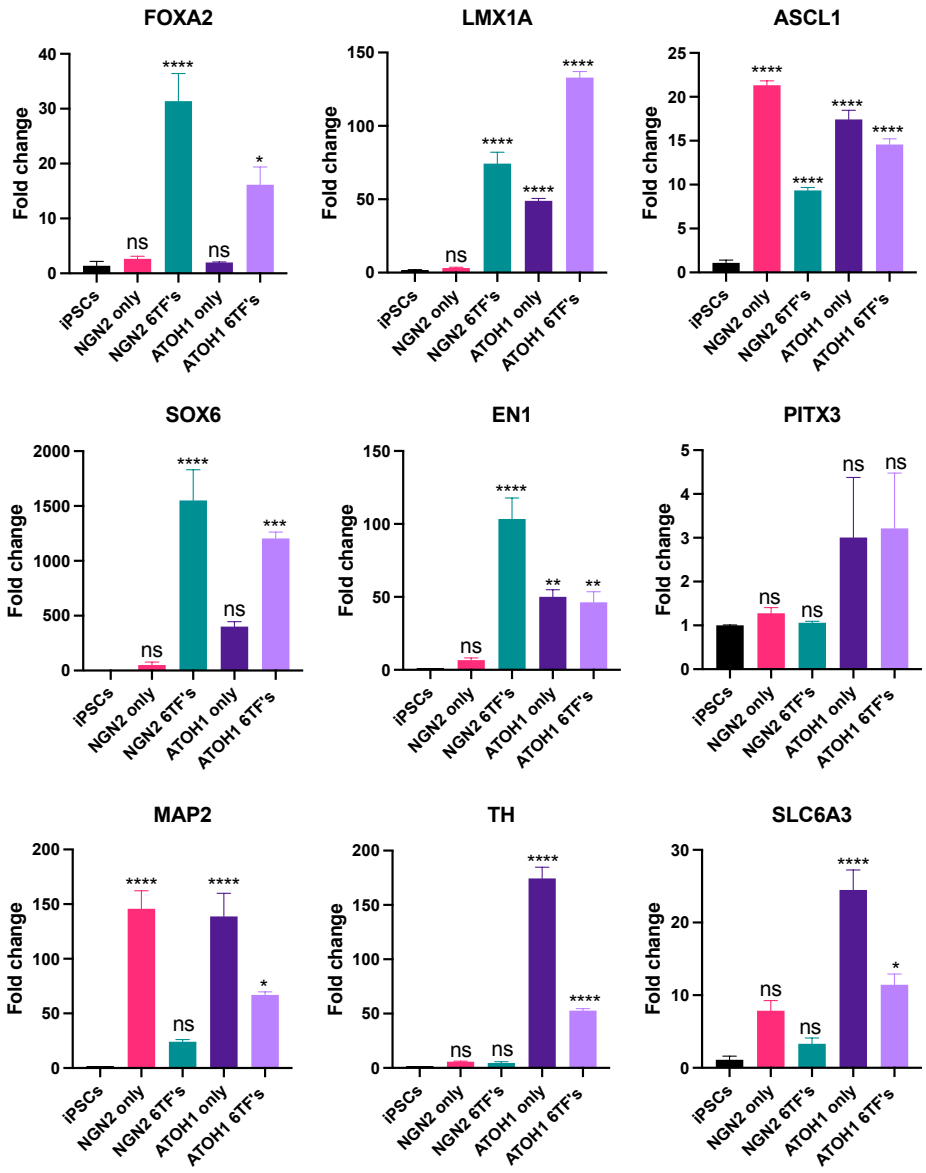



**Figure 5.11:** Bar plots presenting qPCR analysis results for the expression of dopaminergic neurons markers in iNeurons generated using various combination of TFs. 6TFs include: FOXA2, LMX1A, NURR1, PITX3, SOX6, MSX1. The data is represented as Mean ± SEM, n = 3. Statistical analysis was performed using One-Way ANOVA with Dunnett's Multiple comparisons test. Significance levels: *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001, in comparison to the Control (iPSCs).

To validate the findings from the single-cell RNA sequencing of ATOH1+ 6TFs and NGN2+ 6TFs induced Neurons, qPCR analysis was conducted on a new batch of neuronal

cultures. These cultures were generated using transient expression of ATOH1+ 6TFs and NGN2+ 6TFs combinations, with NGN2 only and ATOH1 only conditions serving as controls. The cell samples were collected on day 9 for subsequent qPCR analysis.

Both the NGN2 only and ATOH1 only conditions exhibited an approximate 150-fold increase in the expression of the neuronal marker MAP2. In contrast, the NGN2+ 6TFs and ATOH1+ 6TFs conditions showed around a 25-fold and 70-fold increase in MAP2 expression, respectively. Significant upregulation was also observed in the expression of the proneural gene ASCL1 across all conditions.

The upregulation of TH expression was noticeable in all conditions. However, distinct differences in TH expression were evident between cultures generated using ATOH1 and NGN2. Specifically, the NGN2 only and NGN2+ 6TFs conditions displayed approximately a 5-fold increase, while the ATOH1 only and ATOH1+ 6TFs conditions exhibited 174-fold and 52-fold increases, respectively.

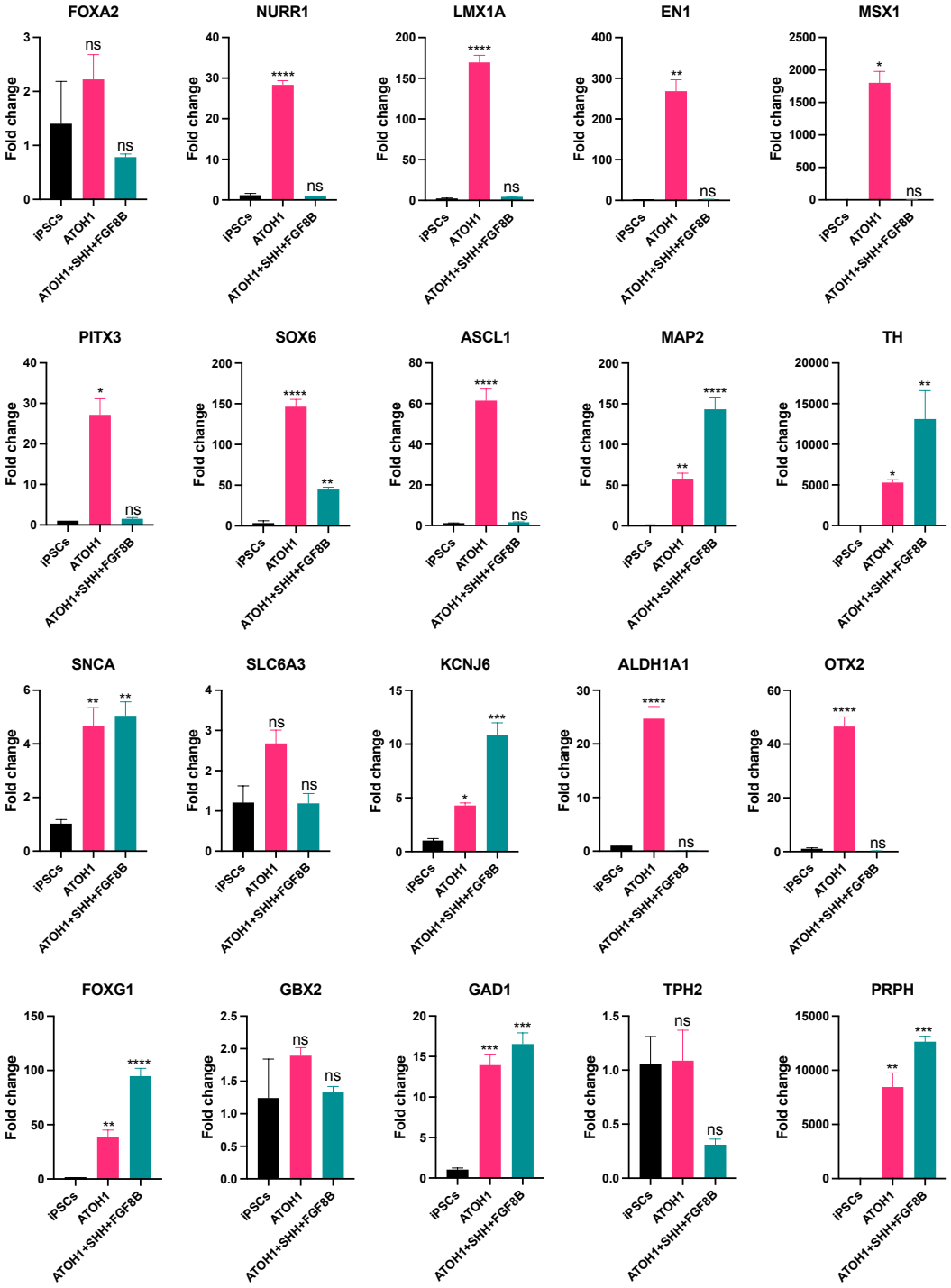
In the NGN2+ 6TFs condition, robust upregulation of dopaminergic progenitor markers FOXA2, LMX1A, and EN1 was observed (approximately 30-fold increase in FOXA2, 75-fold increase in LMX1A, and 100-fold increase in EN1). Conversely, there was a relatively modest ~25-fold increase in the neuronal marker MAP2 and a ~5-fold increase in the dopaminergic neuronal marker TH, suggesting a high proportion of undifferentiated cells.

In contrast, the ATOH1+ 6TFs condition demonstrated strong upregulation of dopaminergic progenitor markers FOXA2, LMX1A, and EN1 (approximately 15-fold increase in FOXA2, 130-fold increase in LMX1A, and 50-fold increase in EN1). Notably, there was also substantial ~70-fold upregulation in the neuronal marker MAP2, a ~50-fold increase in the dopaminergic neuronal marker TH, and a ~12-fold increase in SLC6A3 (a gene encoding the dopamine transporter), indicating the prevalence of dopaminergic neurons. Additionally, ATOH1+ 6TFs induced cells exhibited a ~5-fold upregulation of the mature dopaminergic marker KCNJ6. In the NGN2 only condition, negligible levels of FOXA2, LMX1A, PITX3, and EN1 were observed, which aligned with expectations.

Remarkably, in the ATOH1 only condition, there was upregulation of dopaminergic progenitor markers FOXA2, LMX1A, EN1, and PITX3, alongside the highest expressions of

TH and SLC6A3 (~2-fold increase in FOXA2, ~50-fold increase in EN1 and LMX1A, and ~3-fold increase in PITX3). This observation is consistent with the study by Sagal et al., 2014, wherein ATOH1-induced neurons demonstrated elevated expression of dopaminergic markers such as FOXA2, LMX1A, and NURR1. Moreover, ATOH1-only induced cells displayed ~10-fold upregulation of the mature dopaminergic marker KCNJ6.

## 5.2.12 Characterization of Induced Dopaminergic Neurons Derived Using Controlled Ectopic Expression Of ATOH1 Alone And ATOH1+ Shh_FGF8a



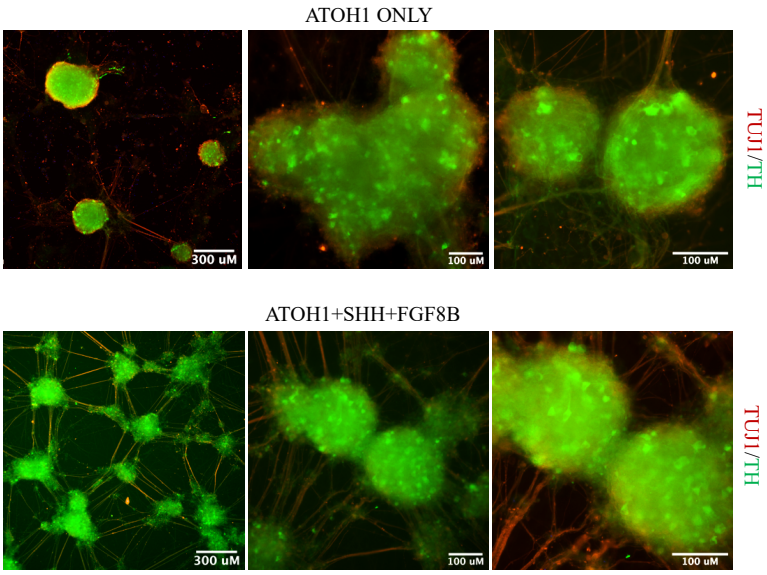



**Figure 5.12**: Immunostaining images illustrating expression of TH and TUJ1 in DA iNeurons generated using ectopic expression of ATOH1 alone and ATOH1 + Shh_FGF8b on day 35. Bar plots presenting qPCR analysis results for the expression of genes associated with various neuronal subtypes, alongside markers for different brain regions: OTX2 for VTA, GBX2 for hindbrain, and FOXG1 for forebrain, within iNeurons generated using ATOH1 alone and ATOH1+ Shh_FGF8b. 6TFs include: FOXA2, LMX1A, NURR1, PITX3, SOX6, MSX1. The data is represented as Mean ± SEM, n = 3. Statistical analysis was performed using One-Way ANOVA with Dunnett's Multiple comparisons test. Significance levels: *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001, in comparison to the Control (iPSCs).

To investigate whether the combination of Shh and fgf8b morphogens, along with ATOH1 overexpression, enhances the efficiency of dopaminergic neuron differentiation and the upregulation of various dopaminergic neuronal markers, as previously reported by Sagal et al. in 2014, we adapted their differentiation protocol. Initially, an iPSCs line was generated through transfection with a DOX-inducible ATOH1 construct carrying a puromycin selection marker. The puromycin-selected iPSCs were plated on PDL/Geltrex-coated NUNC plates on day 0. Subsequently, the cells were exposed to DOX, Shh, and fgf8b morphogens for 5 days, followed by maturation in a media containing neurotrophic factors until day 36. As a control, cells with only ATOH1 overexpression and no Shh and fgf8b treatment were utilized.

On day 37, the cells were collected for qPCR and immunostaining analyses. Immunostaining revealed substantial TH expression in both the ATOH1-only and ATOH1+Shh+fgf8b conditions.

The qPCR analysis unveiled that dopaminergic markers such as FOXA2, LMX1A, NURR1, EN1, ASCL1, SOX6, and OTX2 were exclusively upregulated in the ATOH1-only condition. In contrast, the ATOH1 + Shh + FGF8b condition did not exhibit a noticeable increase in these markers. Notably, mature dopaminergic neuronal markers including SLC6A3 and ALDH1A1 were highly expressed solely in the ATOH1-only condition. The SLC6A3 expression in the ATOH1+Shh+fgf8b condition resembled that of the iPSCs, whereas a higher expression of TH, MAP2, SNCA, and KCNJ6 was observed in the ATOH1+Shh+FGF8b condition compared to the control.

Both conditions displayed remarkably elevated levels of PRPH, a peripheral nervous system marker, indicating a peripheral nervous system identity for the neurons generated. Furthermore, these cells expressed GAD1, a GABAergic neuronal marker, and FOXG1, a forebrain marker, suggesting a certain degree of heterogeneity within the cultures. Taken together, these findings imply that the combination of ATOH1+Shh+fgf8b can indeed enhance the generation of TH-positive cells. However, the absence of dopaminergic marker expression and the high expression of peripheral nervous system markers point toward a non-central nervous system identity for the generated neurons.

## 5.3 Discussion and Future Directions

In this chapter, we explored the potential of inducing dopaminergic neuronal differentiation in human iPS cells by ectopically expressing transcription factors (TFs) associated with midbrain dopamine neuron development.

Various methods for generating induced dopamine neurons have been previously documented. Despite their promising outcomes, most of these protocols are beset by limitations, such as inefficiency in large-scale experiments, insufficient characterization of the resulting neurons, or the requirement for expensive small molecules and morphogens in addition to TF overexpression. To address these challenges, there is an urgent need to establish an efficient protocol capable of generating highly homogeneous and comprehensively

characterized substantia nigra dopaminergic neurons on a large scale within a relatively short timeframe.

In our study, we selected a group of transcription factors (TFs) including ATOH1, NGN2, FOXA2, LMX1A, NURR1, PITX3, SOX6, and MSX1 based on their known roles in midbrain dopaminergic neuron development and survival. All of these TFs, with the exception of SOX6, have been employed previously to generate induced dopaminergic neurons. However, previous attempts did not specifically target substantia nigra neurons but rather aimed to generate a broad midbrain dopaminergic neuron population.

Midbrain dopaminergic neuron subtypes are found in both the ventral tegmental area and substantia nigra pars compacta. However, it is primarily the SNpc neurons that degenerate in Parkinson's disease (PD). Studies have identified SOX6 as a critical factor in SNpc dopaminergic neuron development (Panman et al., 2014). SNpc dopaminergic neurons expressing SOX6 are more susceptible to degeneration in PD compared to CALB1-positive VTA neurons. Furthermore, reduced levels of SOX6 have been observed in PD patients (Kamath et al., 2022; Panman et al., 2014). These findings substantiate our strategy of including SOX6 in our approach to specifically generate substantia nigra neurons.

### 5.3.1 PiggyBac vs Lentivirus Vectors

We introduced these TFs in various combinations using the PiggyBac system, employing both constitutive and inducible expression constructs. Most methods for inducing neurons typically rely on lentivirus vectors, which raise safety concerns and can disrupt endogenous transcriptional processes due to the random integration of transgenes and viral components into the host genome. The primary drawback of the lentivirus system is its limited capacity for accommodating DNA of interest, requiring the generation of virus particles in packaging cells like HEK293T.

In contrast, PiggyBac vectors offer several advantages. They do not pose safety concerns related to integrating viruses, eliminate the need for pre-packaging, and have the capability to accommodate substantial amounts of DNA of interest. Within PiggyBac vectors, it’s feasible to integrate multiple transgene expression cassettes into the target cell’s genome with high efficiency using a single transposon vector (Chen et al., 2015). Conventional

transfection of PiggyBac vectors, in conjunction with a helper plasmid (transposase), results in the efficient integration of the transgene into the host genome.

### 5.3.2 ATOH1 vs ASCL1

In this section, we revisit a pivotal discovery made by Sagal et al., demonstrating that the ectopic expression of ATOH1 alone is sufficient to generate TH-positive iNeurons (Sagal et al., 2014). Initially, their research highlighted the upregulation of ATOH1 during neuronal differentiation facilitated by small molecules designed for directing the differentiation of human iPSCs. Within their study, the sole ectopic expression of ATOH1 yielded 30% TH-positive neurons, with fewer than 10% exhibiting positivity for other neuronal subtype markers, such as GAD67, serotonin, and VGluT1. Sagal et al. also revealed that this expression of ATOH1 alone could elevate various dopaminergic markers, including LMX1A, FOXA2, NURR1, DAT, and VMAT, alongside TH expression.

In our study, we conducted similar characterizations on the TH-positive neurons obtained through the sole ectopic expression of ATOH1. Our qPCR analysis unequivocally demonstrated that ATOH1 alone significantly upregulates the expression of dopaminergic neuronal markers, such as LMX1A, NURR1, ASCL1, KCNJ6, ALDH1A1, and EN1. However, we did not observe a substantial upregulation of FOXA2; this disparity may be attributable to cell line-specific differences. Additionally, we observed elevated levels of PRPH, a marker associated with the peripheral nervous system, indicating a peripheral nervous system identity. It is worth noting that Sagal et al. did not investigate the brain regional identity of the TH-positive cells. Notably, TH is not a specific marker for midbrain dopaminergic neurons, as peripheral neurons have also been shown to express TH, where it contributes to catecholamine biosynthesis (Brumovsky et al., 2017). Therefore, ensuring the accurate regional identity of induced dopaminergic neurons is crucial for their future application in research and clinical contexts.

Similar findings were reported by Ng et al., who utilized the proneural TF ASCL1 to induce TH-positive neurons from ES cells. However, ASCL1-induced TH-positive cells exhibited peripheral nervous system marker expression. Furthermore, even when combined with LMX1A and NURR1, ASCL1 was unable to completely shift their identity from peripheral to central. Powell et al. employed a very similar combination of ASCL1, LMX1B,

and NURR1 to generate induced dopaminergic neurons from human iPSCs. These TFs were introduced into the cells using a single DOX-inducible lentivirus vector. The induced dopaminergic neurons exhibited the expression of midbrain dopamine neuronal markers, demonstrated dopamine release, displayed electrophysiological properties, and had a transcriptomic profile resembling fetal midbrain dopaminergic neurons. Notably, neurons derived from human iPSCs using small molecules and morphogens have also been shown to resemble fetal neurons (16-24 post-conception weeks). Thus, the presence of these characteristics in Powell's induced neurons is consistent with existing research. It would be of interest to investigate whether Powell's induced neurons also express PRPH and to what extent they resemble *in vivo* substantia nigra neurons.

### 5.3.3 ATOH1 in Combination with SHH and FGF8b

We also explored whether ectopic expression of ATOH1, in conjunction with Shh and FGF8b, could enhance the efficiency of TH-positive neuron generation, as suggested by Sagal et al. Interestingly, TH-positive neurons generated using this combination did not exhibit higher levels of various dopaminergic markers such as LMX1A, NURR1, EN1, MSX1, ASCL1, and OTX2. However, these cells did display elevated expression of mature dopaminergic neuronal markers like TH, MAP2, KCNJ6, and ALDH1A1, indicating improved maturity compared to cells generated with ATOH1 alone. Notably, ectopic expression of ATOH1 in combination with Shh and FGF8b also did not alter the regional identity of the cells, as strong PRPH expression was observed in cultures derived using this combination. Furthermore, the culture generated with the ATOH1, SHH, and FGF8b combination exhibited high levels of FOXG1, a forebrain marker, and GAD1, a GABAergic marker, implying greater heterogeneity compared to the ATOH1 alone condition.

### 5.3.4 The Importance of Proneural TF Ectopic Expression in Induced Neurons

Our study revealed that all investigated combinations yielded TH-positive neurons, except in conditions where proneural factors such as ATOH1 or NGN2 were absent. This underscores the critical role of proneural TFs in generating induced neurons. Previous research has reported the generation of induced dopaminergic neurons from neural precursor cells through the ectopic expression of proneural factors. For example, functional dopamine (DA) neurons have been generated by ectopically overexpressing FOXA2 and NURR1, LMX1A and FOXA2, or LMX1A, LMX1B, FOXA2, and OTX2 in induced neural precursor cells (iNPCs)

(Lim et al., 2015; Pereira et al., 2014; Pfisterer et al., 2011). Similarly, Lee et al. reported the derivation of transplantable induced dopaminergic neurons through the ectopic expression of FOXA2 and NURR1 in NPCs (Lee et al., 2010). However, to the best of our knowledge, no method has been reported to generate induced neurons from pluripotent stem cells without the ectopic expression of proneural factors.

### 5.3.5 Inducible vs. Constitutive Expression Vectors

We investigated several combinations of TFs using constitutive expression PB constructs, such as ATOH1 or NGN2 + 6TFs (SOX6, MSX1, LMX1A, FOXA2, PITX3, and NURR1). These combinations did yield TH-positive neurons, with the induction of 6TFs along with ATOH1 proving to be more efficient than with NGN2. However, these results could not be replicated in cells generated using the same combinations but with inducible PB constructs. It remains unclear why the findings from constitutive expression vectors were not recapitulated using inducible expression constructs with the same combination. It is possible that the results obtained from constitutive expression experiments were somewhat misleading. In the constitutive expression experiment, there was no control over the number of TFs being expressed in a particular cell, and it was uncertain whether the resulting TH-positive neurons received all the transfected TFs. It is plausible that a majority of the TH-positive neurons generated in the ATOH1 and other TFs combination may have been induced by ATOH1 alone, with other TFs playing no significant role.

### 5.3.6 Single Cell RNA Sequencing Analysis of ATOH1/NGN2 + 6TFs

A single RNA-seq study was conducted on iNeurons generated using the ATOH1 + 6TFs and NGN2+6TFs combinations via constitutive PB vectors. This analysis revealed the presence of three distinct TH-positive populations. All three TH-positive populations exhibited strong expression of PRPH, a marker associated with the peripheral nervous system, and lacked expression of KCNJ6, a marker specific to substantia nigra A9 neurons. These findings suggest that iNeurons generated using ATOH1 or NGN2 + 6TFs lack central nervous system (CNS) identity. Given that TH-positive neurons constituted only a small fraction of the total cells in this single RNA-seq analysis, further experiments are warranted to characterize NGN2+6TFs cells, including their brain regional identity.

### 5.3.7 ATOH1 + 3TFs

A targeted TF screening was designed and executed to reduce the number of TFs required for generating TH-positive iNeurons. The results suggested that the combination of ATOH1, LMX1A, NURR1, and PITX3 held promise for further investigation. Ectopic expression of this combination via inducible PB vectors yielded TH-positive neurons, making it the sole combination in our study to achieve this result via an inducible PB construct. However, it is essential to investigate the regional identity of these neurons further. Additionally, it's worth noting that this combination lacks TFs specific to substantia nigra neurons, suggesting that it may not specifically generate substantia nigra neurons.

### 5.3.8 TFs Combination To Induce Dopaminergic Neurons Of Central Nervous System Identity

In their study, Ng et al. observed that several TF combinations, including ASCL1 alone, were ineffective in inducing central nervous system identity in TH-positive neurons. However, they successfully addressed this challenge by refining the combination of factors. They discovered that ectopic expression of midbrain-specific factors, namely ASCL1, NURR1, EN1, FOXA2, LMX1A, and PITX3, in embryonic stem cells (ESCs) via the PiggyBac system induced TH-positive dopamine neurons with a central identity. This demonstrates that even when a proneural factor alone induces peripheral identity, it can give rise to central nervous system TH-positive neurons when combined with the appropriate midbrain dopaminergic neuronal development-associated TFs.

### 5.3.9 Integration of TFs in Safe Harbour Sites Using OPTI X System

Following the identification of an effective combination of TFs for generating dopaminergic neurons from human induced pluripotent stem cells (hiPSCs), our plan was to create clonal iPSC lines with integrated TFs at genomic safe harbor sites using the Optimized Inducible Overexpression System (OPTI X) (Pawlowski et al., 2017).

The OPTI X Tet-ON system, similar to the PiggyBac Tet-ON system, consists of two key components: a constitutively expressed reverse tetracycline transactivator (rtTA) responsive to doxycycline and a Tet-responsive element (TRE)-inducible promoter controlling the expression of the transgene. However, unlike the PiggyBac system, which integrates transgenes into undesired regions of the host genome, potentially causing unwanted interference with the endogenous transcription network, all elements of OPTI X, including

selected transgenes, are precisely targeted to genomic safe harbor (GSH) sites. Specifically, the activator rtTA cassette, regulated by the CAG promoter, is directed to the ROSA26 locus, while the responder TRE promoter, responsible for transgene expression, is directed to the AAVS1 locus.

This dual-targeted approach has been proven to yield controlled, efficient, and highly inducible expression of transgenes in hiPSCs. Notably, the OPTI X system has already demonstrated successful application in forward programming hiPSCs into glutamatergic neurons, hepatocytes, myocytes, and oligodendrocytes. Therefore, employing the OPTI X system for deriving dopamine iNeurons is a preferred choice over using PiggyBac-derived cells for both *in vitro* and *in vivo* applications.

### 5.3.10 iPSCs Derived vs Fibroblast Derived induced Neurons

The use of induced pluripotent stem cells (iPSCs) offers a significant advantage in potentially providing an unlimited supply of induced dopaminergic neurons for various purposes, including research and transplantation. However, iPSC-derived induced dopaminergic neurons have been observed to closely resemble fetal midbrain dopaminergic neurons. Moreover, studies have indicated that these cells do not retain age-related phenotypes and epigenetic signatures of the donor patients (Mertens et al., 2018).

Conversely, induced dopaminergic neurons derived from fibroblasts through direct reprogramming approaches can preserve age-related changes in transcriptomics and epigenetic profiles, as well as exhibit characteristics such as DNA damage, telomere length, and mitochondrial function (Drouin-Ouellet et al., 2022). The research conducted by Drouin-Ouellet et al. demonstrated that induced dopamine neurons derived from adult human dermal fibroblasts of Parkinson's disease (PD) patients exhibited stress-induced chaperone-mediated autophagy and macro-autophagy impairment, along with the accumulation of phosphorylated alpha-synuclein. These features were notably absent in iPSC-derived induced dopaminergic neurons from the same patients. Additionally, fibroblast-derived induced neurons expressed mature isoforms of TAU(4R), a characteristic that was lacking in iPSC-derived induced neurons. Despite these promising results, scalability remains a challenge for fibroblast-derived induced neurons, as fibroblasts cannot undergo unlimited expansion and eventually become senescent after a limited number of passages.

# 6. Unravelling Cell Type Specific Response to Parkinson's Disease

## 6.1 Background

Parkinson disease is a progressive neurological disorder, characterized by the loss of dopaminergic neurons in substantia nigra pars compacta and presence of Lewy bodies which primarily comprise of aggregated alpha-synuclein proteins in the surviving neurons. Despite extensive research, the molecular and cellular mechanisms underlying this devastating disease remain elusive.

A significant stride forward is anticipated through the utilization of high-throughput single-cell/nuclei RNA sequencing (RNA-seq) technologies, which afford a distinctive avenue to systematically explore cell-type-specific responses to the disease. In recent years, RNA-seq has been employed in human PD models, encompassing *in vitro* studies and examinations of post-mortem human brain tissues. Nonetheless, acquiring a comprehensive post-mortem dataset remains challenging due to the scarcity of human brain tissue and potential RNA degradation, particularly notable among tyrosine hydroxylase (TH) positive dopaminergic neurons within PD samples, a subset more susceptible to degeneration.

Prior endeavours, like those of Agarwal et al., 2020, focused on the substantia nigra (SN), and Smajić et al., 2022, centred on the midbrain, captured fewer than 200 TH-expressing cells, predominantly neurons. Kamath et al., 2022, adopted a specific sorting strategy targeting pigmented regions of the SN, enhancing the dataset with dopaminergic neurons. This led to the capture of around 30,000 TH-expressing cells across 8 control and 6 PD samples, permitting comprehensive characterization of approximately 17,000 DA neurons within the substantia nigra. Interestingly, minute TH+ glial populations were also present in the dataset, albeit not subjected to further scrutiny.

In our study, we utilized single-cell/nucleus RNA-seq technology to unravel the cellular heterogeneity at the molecular level in the substantia nigra pars compacta of 14 control and 15 sporadic PD patients. To our knowledge, this is the first RNA-seq study conducted on control and PD SNpc at single-nuclei resolution, providing a ground-breaking insight into the multicellular response in PD. We present a comprehensive, transcriptome-level characterization of all the major brain cell types in SNpc, assessing the specificity of GWAS PD associated genes to these cell types. Our study revealed subpopulation-specific gene sets

and highlighted the differing roles of these subpopulations, providing a unique understanding of the disease mechanisms underlying PD. Importantly, we identified *TH*-positive glial subpopulations that were almost absent in PD samples, along with a shared transcriptomic signature between PD-depleted TH-expressing cell types. Our findings provide a valuable resource for future hypothesis-driven experiments and represent a significant step forward in our understanding of PD disease mechanisms.

## 6.2 Results

### 6.2.1 Single Nucleus RNA-Seq Reveals Cell Type Heterogeneity in Human SNpc

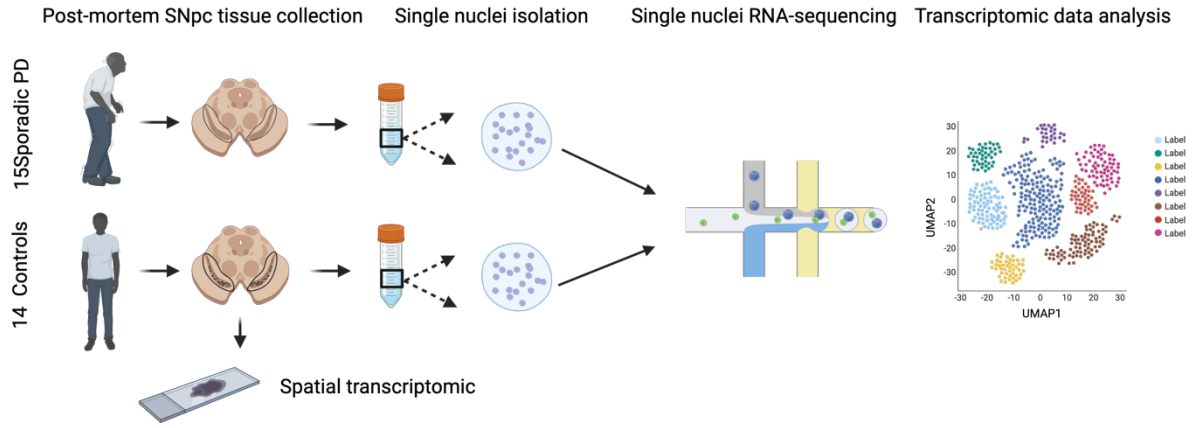



**Figure 6.1:** Schematic of the experimental design. Control and PD substantia nigra pars compacta (SNpc) nuclei were isolated from frozen post mortem brain tissues and processed with 10X genomics chromium platform. Spatial transcriptomics was performed on SNpc tissue taken from 3 Controls and 3 sporadic PD patients.

Substantia nigra pars compacta (SNpc) tissue was sampled from post-mortem human brains of 14 control individuals and 15 sporadic Parkinson's disease (PD) patients. The average age of these sporadic PD patients and control individuals was about 79 years. Both groups had similar post-mortem intervals (Control: ~18 hours, PD: ~22 hours). According to the pathology reports of these samples, all the sporadic PD patients showed the presence of Lewy bodies in the midbrain and ranged from stage 1 to stage 4 on the Braak scale for neurofibrillary tangles.

Using a 10X genomic chromium platform, single-nuclei RNA sequencing (snRNAseq) was performed on more than 80,000 high-quality nuclei from these 29 samples. A relatively similar number of nuclei were obtained from PD (~39,000) and control samples (~44,000). For

these ~83,000 nuclei, an average of ~33,971 transcripts and 1527 genes per nucleus were determined after filtering out poorly sequenced nuclei and doublets detected by the Scrublet algorithm.

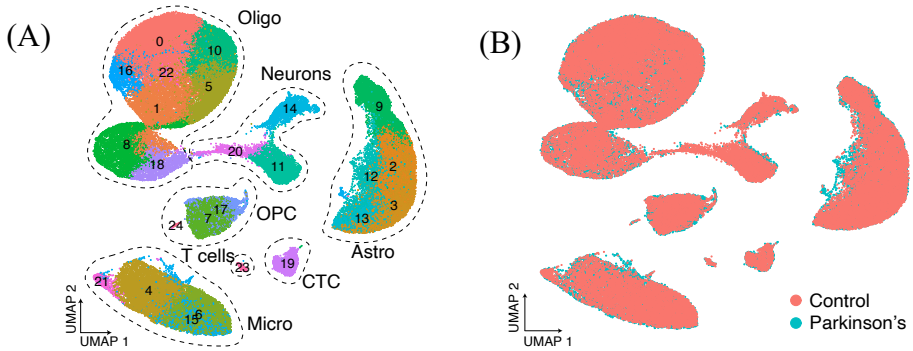
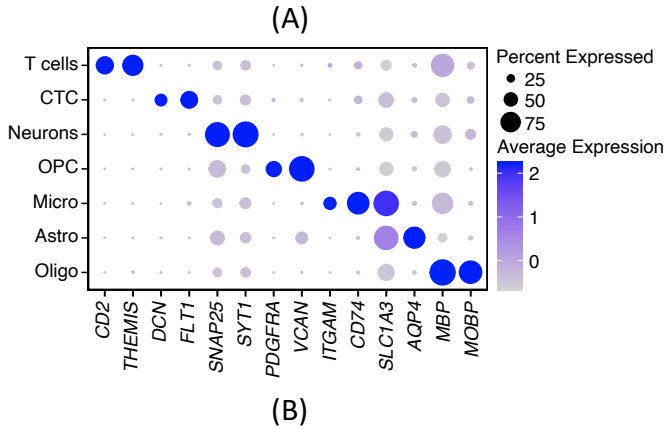



**Figure 6.2:** (A) UMAP embedding of 83,484 high quality nuclei obtained from 15 PD and 14 Control samples. Grouped clusters represent neurons, astrocytes (Astro), oligodendrocytes (Oligo), oligodendrocyte progenitor cells (OPC), microglia (Micro), T cells, and connective tissue cells (CTC), annotated based on corresponding cell marker gene expression. (B) UMAP showing representation of Control and PD samples in identified cell types. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

UMAP dimension reduction analysis of the merged control and PD datasets showed that clustering was driven by cell type identity and inter-sample variability. To normalize the inter-sample variability, an anchor-based integration approach was used for representative clustering and cell type identification. Louvain-based clustering identified 25 clusters, with both control and PD cells found to be present within all the cell clusters.



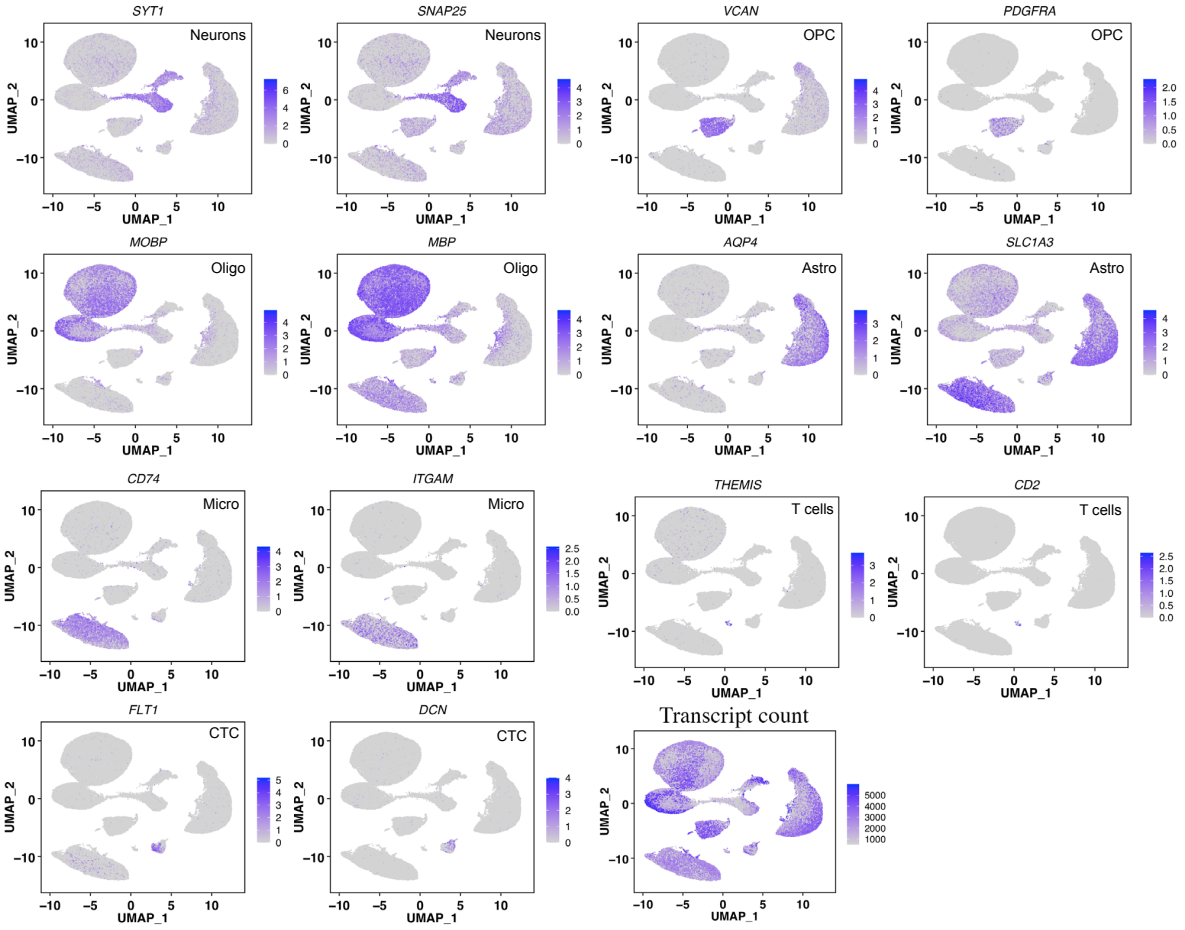



**Figure 6.3:** (A) Dot plot illustrating expression of cell markers used to identify higher level cell types in SNpc. (B) UMAP heatmaps of higher-level cell markers illustrating their specificity to the cell types. Ln-normalized SCT counts are shown. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

A panel of marker genes was utilized to characterize and annotate the cell clusters. Clusters 0, 1, 5, 8, 10, 16, 18, and 22 were assigned as oligodendrocytes based on MOBP and MBP expression. Similarly, clusters 2, 3, 9, 12, and 13 were annotated as astrocytes (AQP4 and SLC1A3), clusters 11, 14, and 20 as neurons (SYT1 and SNAP25), clusters 4, 6, 15, and 21 as microglia (CD74 and ITGAM), clusters 7, 17, and 24 as oligodendrocyte progenitor cells (OPCs; PDGFRA and VCAN), cluster 19 as connective tissue cells (CTCs; DCN and FLT1), and cluster 23 as T cells (THEMIS and CD2).

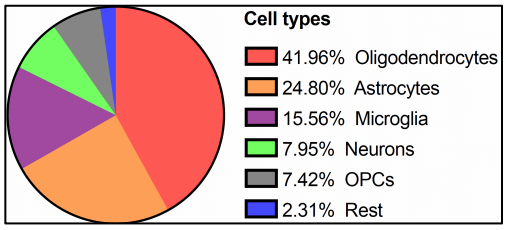



**Figure 6.4:** A pie chart showing proportion of higher-level cell types detected in the dataset.

Oligodendrocytes accounted for 41.96% (~35,032 cells) of all sampled cells. Astrocytes and microglia comprised 24.81% (~20,710 cells) and 15.57% (~12,995 cells) of the sample, respectively. OPCs made up 7.96% (~6,644 cells), neurons represented 7.42% (~6,196 cells), and the remaining cells constituted 2.28% (1,907 cells).

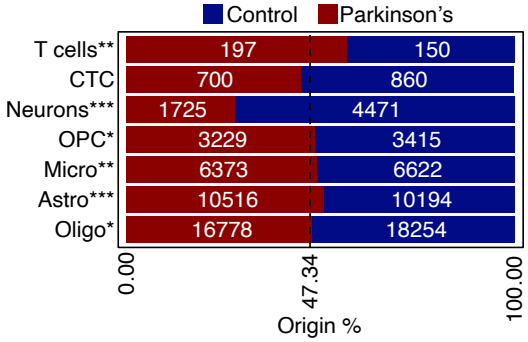



**Figure 6.5**: A comparative analysis of cell proportions between PD and Control SNpc tissues, in relation to the anticipated fraction of 47.34% cells representing the PD condition among all nuclei considered in the clustering analysis. To assess the potential deviation of cell proportions from this expected value, a binomial test was conducted. Significance levels were indicated as follows: * p-value < 0.05, ** p-value < 0.0005, *** p-value < 0.00005. Note: This plot was generated by Araks Martirosyan and Francisco Pestana.

Assessment of cell proportions highlighted that neurons are significantly depleted in PD samples (p-value < 0.00005), as expected. On the other hand, the relative proportions of oligodendrocytes (p-value < 0.05), astrocytes (p-value < 0.00005), microglia (p-value < 0.0005), OPCs (p-value < 0.05), and T cells (p-value < 0.0005) were found to be increased in PD samples.

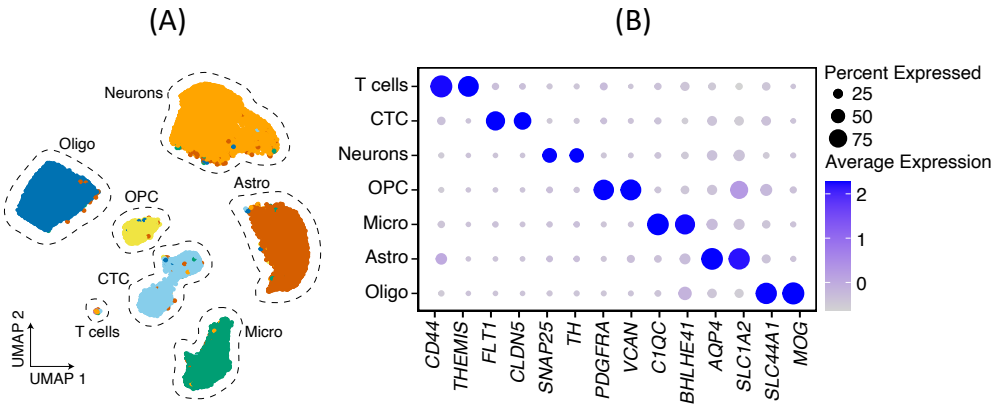



**Figure 6.6:** (A) Independent clustering of the spatial transcriptomics dataset derived from SNpc tissue samples of 3 Control individuals and 3 sporadic PD patients. (B) validation of expression of cell type marker through spatial transcriptomics analysis. Note: These plots were generated by Araks Martirosyan and Francisco Pestana.

Multiplexed spatial transcriptomics was performed to confirm the presence of cell types using post-mortem substantia nigra pars compacta tissues from 3 control individuals and 3 patients with sporadic PD. A panel of 96 cell type-specific markers was utilized to validate the presence of the cell types detected by snRNAseq. Examples of cell type-specific markers used for this analysis include CD44 and THEMIS for T cells, FLT1 and CLDN5 for CTCs, SNAP25 and TH for neurons, PDGFRA and VCAN for OPCs, C1QC and BHLHE41 for microglia, and AQP4 and SLC1A2 for astrocytes.

For cell type clustering and cell type identification, 36,806 high-quality cells were analysed and clustered using Seurat's UMAP and anchor-based integration workflows. The analysis confirmed the presence of all 7 broad cell types detected in the snRNAseq dataset.

### 6.2.2 Robust Gene Expression of PD GWAS Associated Genes In Substantia Nigra Pars Compacta Cell Types

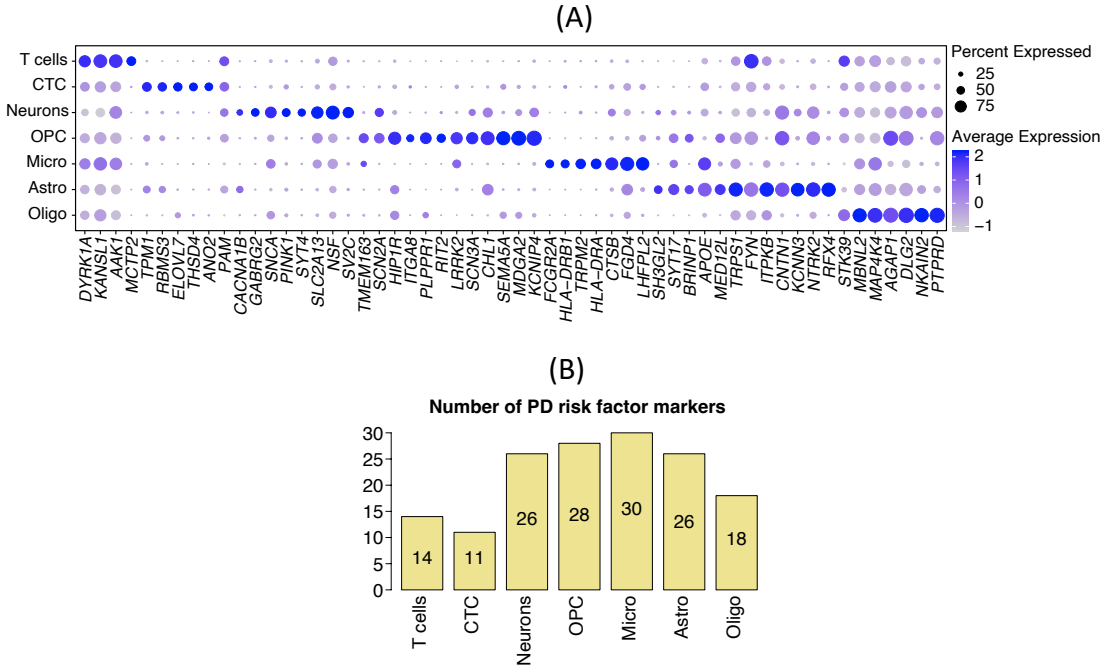



**Figure 6.7:** (A) Specificity of GWAS PD-associated genes compiled from the EMBL GWAS catalogue and Fernandes et al., 2020, across the characterized higher level cell types. Cell types indicated: CTC (connective tissue cells), OPC (oligodendrocyte progenitor cells), Micro (microglia), Astro (astrocyte), Oligo (oligodendrocyte). (B) Enumeration of GWAS PD-associated genes distinctly expressed within individual cell types. Note: These plots were generated by Araks Martirosyan and Francisco Pestana.

Genome-wide association studies (GWAS) have identified numerous susceptibility loci associated with PD, but the specificity of the genes associated with these loci in SNpc cell types has yet to be demonstrated. The expression of 153 putative PD-associated genes from GWAS was assessed in the identified cell types. The selected genes were obtained from the National Human Genome Research Institute-European Bioinformatics Institute (NHGRI-EBI) portal and Fernandes et al. (2020).

A pattern of cell-type specificity in PD GWAS-associated genes was observed across various cell types. Genes such as SNCA, PINK1, and SYT1 were specifically highly expressed in neurons. Similarly, genes like PTPRD, NKAIN2, and DLG2 were specifically highly expressed in oligodendrocytes, while RFX4, NTRK2, and KCNN3 were more prominent in

astrocytes. Microglia expressed HLA-DRA, TRPM2, and HLA-DRB1, whereas T cells showed elevated expression of AAK1, KANSL1, and DYRK1A. OPCs exhibited higher expression of CNIP4, MDGA2, and SEMA5A.

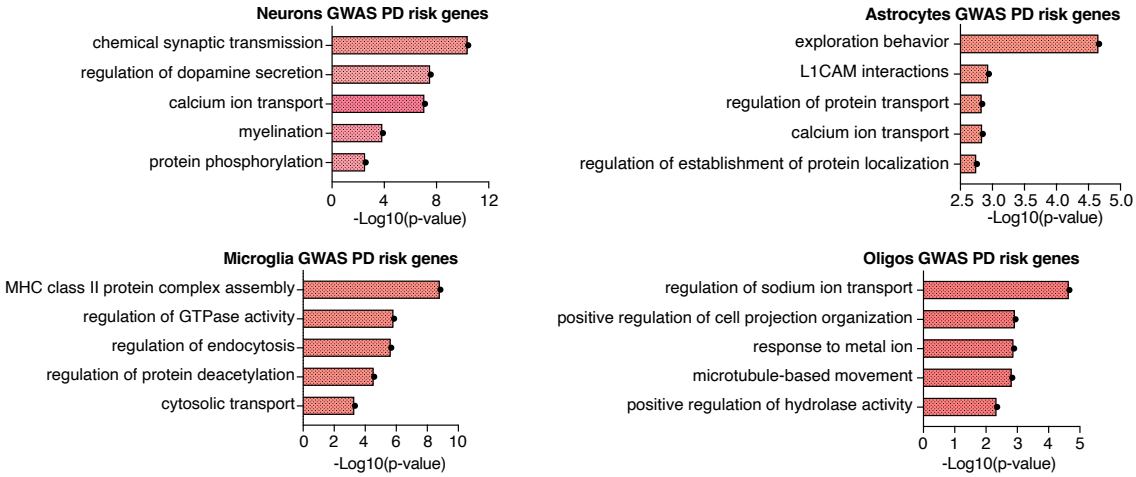



**Figure 6.8:** Pathway analysis on the GWAS PD-associated genes specifically expressed in neurons, astrocytes, microglia, and oligodendrocytes.

The pathway overrepresentation analysis on the cell-type-specific GWAS-associated gene highlighted biological processes unique to certain cell types. For example, neuronal populations exhibited an enrichment of genes involved in dopamine secretion (HTR2A, SNCA, SYT4, and PINK1), synaptic signaling (e.g., CACNA1B, CLCN3, GABRG2, HTR2A, SNCA, SYT4, PARK7, SV2C, and STX1B), myelination (e.g., ANK2, CNTN1, and SCN2A), and protein phosphorylation (CNTN1, HTR2A, SNCA, PARK7, and PINK1). Astrocyte-specific PD GWAS associated genes were found to be involved in protein transport (APOE, CD38, FYN, PAM, SREBF1, BAG3), calcium ion transport (ANK2, CACNA1B, FYN), and L1CAM interactions (ANK2, CNTN1, SH3GL2, CHL1).

Microglia-specific PD GWAS associated genes were enriched in pathways involving GTPase activity (GPR65, TBC1D5, SIPA1L2, LRRK2, FGD4, and ARHGAP27), MHC class II protein complex assembly (HLA-DQA1, HLA-DQB1, HLA-DRA, and HLA-DRB1), regulation of protein deacetylation (CAMK2D, DYRK1A, and LRRK2), and endocytosis (APOE, SNCA, TBC1D5, AAK1, and LRRK2). Oligodendrocyte-specific PD GWAS associated genes were related to the regulation of sodium ion transport (STK39, CNKSR3, and NKAIN2), regulation of hydroxylase activity (MAPT, MAP4K4, and AGAP1), microtubule-

based movement (DLG2, MAPT, and DNAH17), and cell projection organization (SCARB2, MAPT, PTPRD).

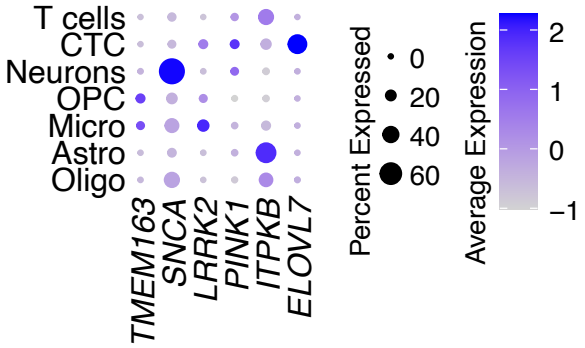



**Figure 6.9**: Validation of cell type-specific expression for PD-associated genes utilizing the spatial transcriptomics dataset: ELOVL7 predominantly expressed in connective tissue cells, ITPKB in astrocytes, TMEM163 in oligodendrocyte progenitor cells (OPCs), LRRK2 in microglia and connective tissue cells. While SNCA exhibits the highest expression in neurons, its presence extends to other cell types including oligodendrocytes, microglia, and OPCs. Note: This plot was generated by Araks Martirosyan and Francisco Pestana.

Spatial transcriptomics essentially confirmed the specificity of PD GWAS-associated genes. For example, LRRK2 exhibited specificity to microglia; SNCA to neurons; ITPKB to astrocytes; TMEM163 to OPCs and microglia; and ELOVL7 to CTCs. These observations are in line with the literature, which suggests the complex involvement of multiple cell types in PD pathology.

### 6.2.3 Neuronal Subpopulations in Aged Human Snpc And Their Response In Sporadic PD

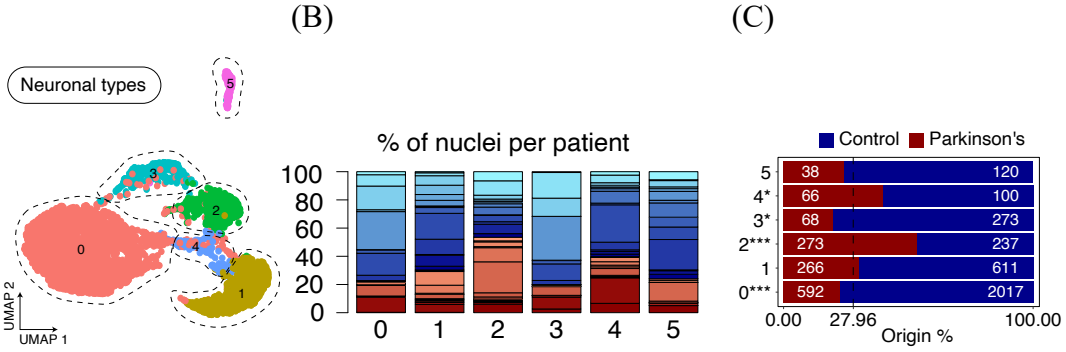



**Figure 6.10:** (A) Re-clustering of neurons as identified in Figure 1.2. (B) Distribution of nuclei proportions between neuronal types from Control (blue) and PD (red) patients, as presented in Figure 3.3A. (C) Comparative analysis of nuclei counts (expressed as a percentage) originating

from Control and PD SNpc tissue samples, contrasted with predicted abundance (calculated from 27.96% of all nuclei from PD samples: indicated by the dashed line). A binomial test was conducted to assess the presence of significant deviations in cell proportions from this value, with significance levels marked as follows: * p-value < 0.05, ** p-value < 0.0005, *** p-value < 0.00005. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

The neuronal population was re-clustered to investigate PD-affected neuronal subpopulations. Re-clustering revealed six distinct subpopulations, labelled as Neuron0 through Neuron5. All subpopulations were characterized by unique gene expression and were represented in multiple control and PD samples. The nuclei proportion analysis showed that Neuron0 and Neuron3 were significantly depleted in the PD condition, whereas Neuron2 and Neuron4 were overrepresented in the PD condition. The numbers of Neuron1 and Neuron5 were not affected at a statistically significant level in the PD condition.

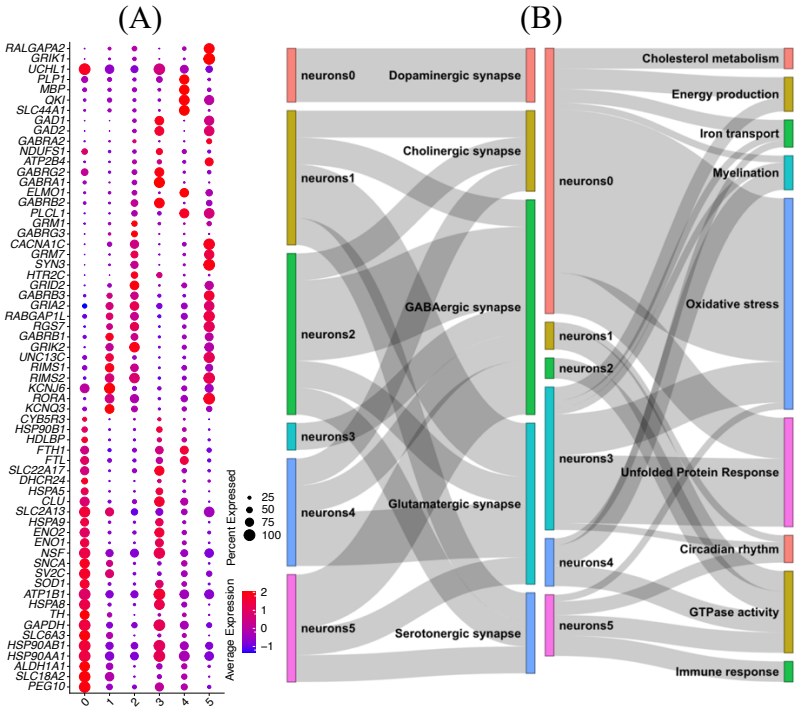



**Figure 6.11**: (A) Dot plot illustrating the expression of cell marker genes within the distinguished neuronal populations. (B) Sankey plots providing a summarized representation of pathway enrichment analysis results concerning the marker genes characterizing neuronal populations 0 through 5. **Note:** These plots were generated by Araks Martirosyan, Francisco Pestana, and Silvia Hnatova.

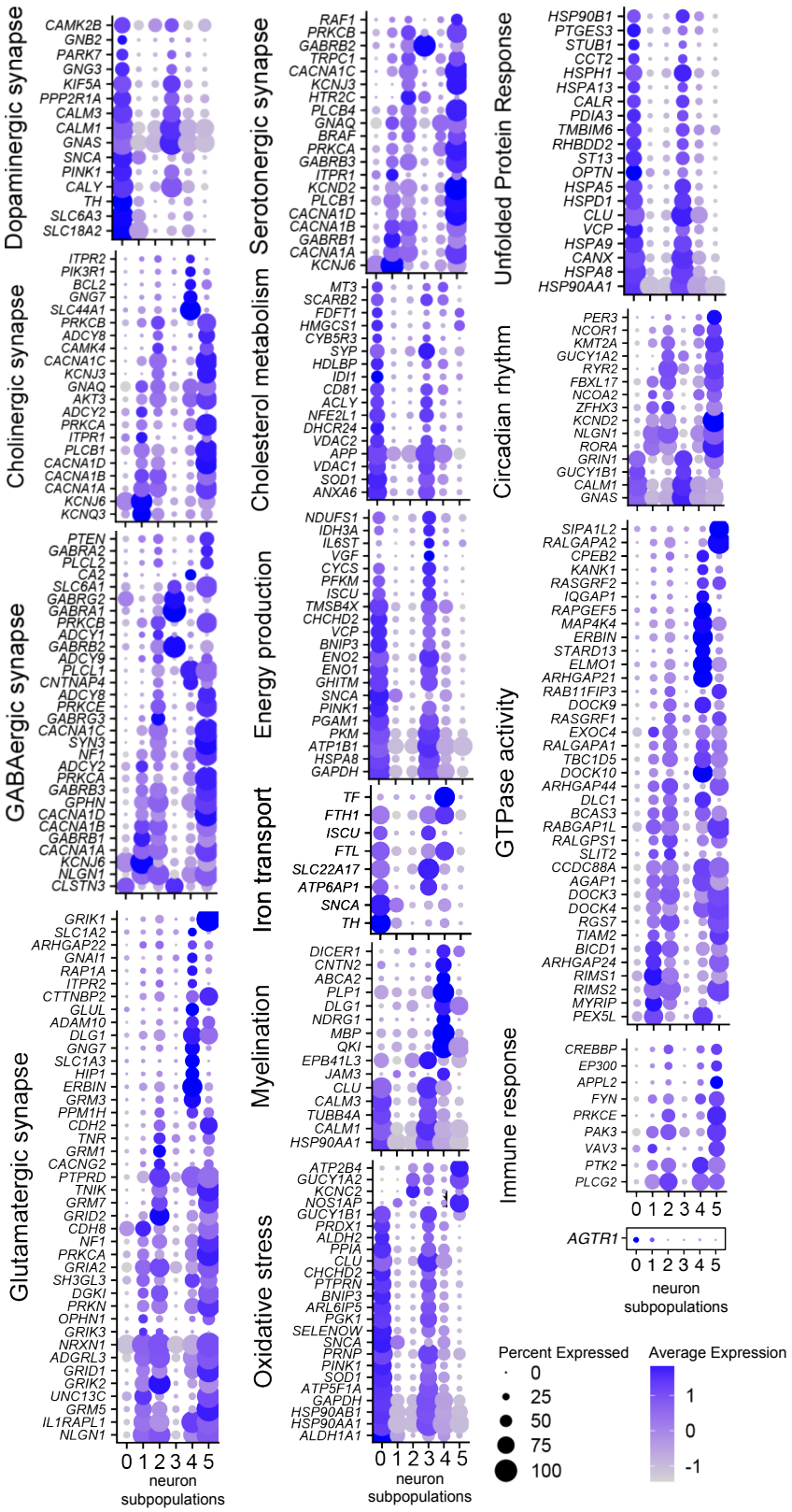



**Figure 6.12**: Cell markers for neuronal populations categorized according to enriched terms outlined in figure 1.11. Depiction of AGTR1 expression across various neuronal populations. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

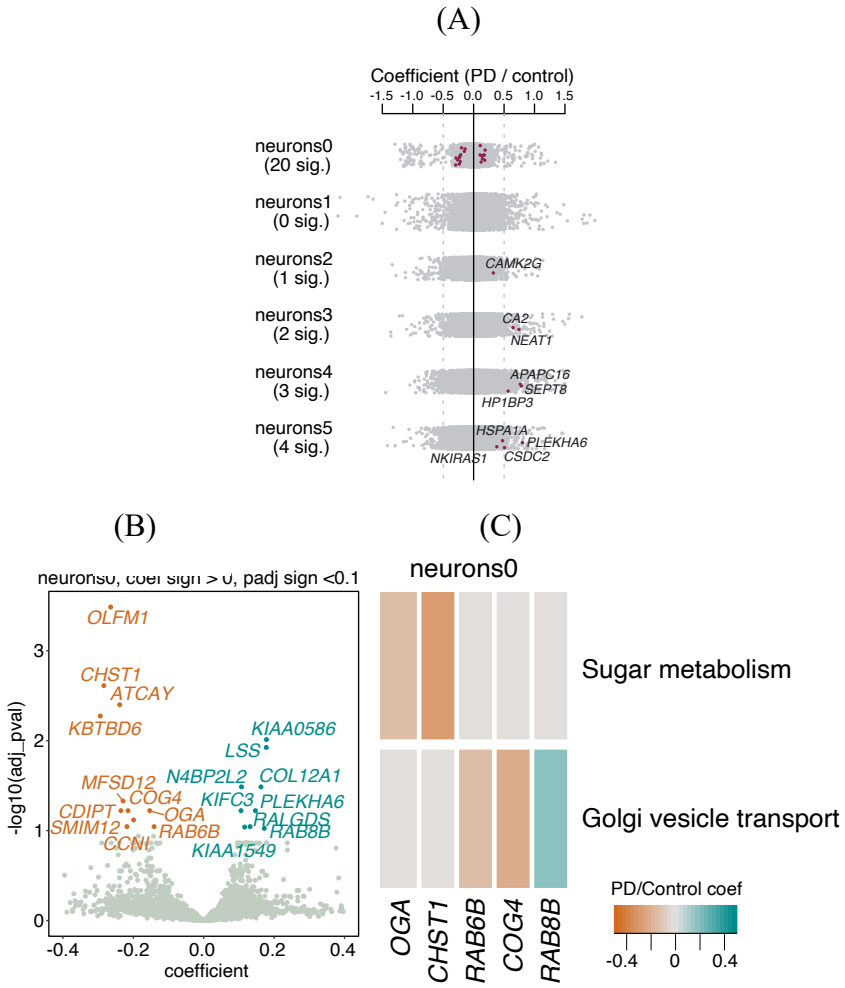



**Figure 6.13:** (A) A stripe chart presenting the count of significantly upregulated or downregulated genes within nuclei from PD donors across distinct neuronal populations, identified through the application of a linear mixed model. Red dots denote genes displaying upregulation (coefficient > 0) or downregulation (coefficient < 0) with an adjusted p-value of less than 0.1 (ANOVA test with Benjamini-Hochberg correction). (B) A volcano plot illustrating genes exhibiting upregulation (cyan) or downregulation (orange) within nuclei originating from PD donors and assigned to the Neurons0 subpopulation. (C) GO and KEGG pathways that are overrepresented by genes displaying upregulation or downregulation within nuclei originating from PD donors and associated with the Neurons0 subpopulation. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

The pathway overrepresentation analysis was performed on the full list of cell marker genes to characterize the subpopulations. Expression of key marker genes for dopaminergic neurons such as TH, SNCA, SLC6A3, KCNJ6, SLC18A2, and ALDH1A1 in Neurons0 suggested dopaminergic identity for this subpopulation. Neurons0 was found to be depleted in

PD samples (p-value < 0.0005), indicating that the primary neuronal subpopulation lost in PD samples is dopaminergic neurons, consistent with the literature. Notably, UCLHL1, a PD-associated gene, and PEG10, a DNA-binding protein-coding gene, were also highly expressed cell markers by Neurons0. This subpopulation expressed marker genes associated with PD pathology, including unfolded protein response (UPR) (e.g., HSP90AA1, HSP90AB1, and HSPA8), oxidative stress (e.g., CHCHD10, CLU, and SOD1), energy production (e.g., ATP1B1, ENO1, and ENO2), iron transport (e.g., FTL, FTH1, and SLC22A17), response to endoplasmic reticulum stress (e.g., SERINC3, FBXO2, CANX, CLU, and HSPA5), and cholesterol metabolism (e.g., DHCR24, CYB5R3, and HDLBP).

Differential gene expression (DEG) analysis between PD and control samples highlighted 20 significant DEGs in the PD Neurons0 population. KBTBD6, encoding a protein involved in proteosome-mediated ubiquitin-dependent protein catabolic process, was found to be downregulated in PD Neurons0. Pathway analysis on the PD Neurons0 DEGs suggested alterations in glucose/sugar metabolism due to the downregulation of carbohydrate sulfotransferase 1 (CHST1) and O-GlcNAcase (OGA). Dysregulation of glucose metabolism has been reported as an early event in sporadic PD. Furthermore, the dysregulation of Ras-related RAB8B and RAB6B, and Oligomeric Golgi complex 4 (COG4), suggested a dysfunction of Golgi-to-endoplasmic reticulum vesicle trafficking. These observations are consistent with recent studies showing vesicle trafficking disruption in PD resulting from alpha-synuclein aggregation and the potential of RAB GTPase to protect neurons.

Neurons3, another neuronal population, was also found to be significantly depleted in the PD sample, albeit with a lower significance value of 0.05. Neurons3 was characterized by the expression of GABAergic gene markers such as GABRB2, GAD1, and GAD2. The pathway analysis of the cell markers in Neurons3 showed that this subpopulation shared many markers associated with UPR (e.g., HSP90AA1, HSP90AB1, and HSPA8), oxidative stress (e.g., CHCHD10, CLU, and SOD1), iron transport (e.g., FTL, FTH1, and SLC22A17), and energy production (e.g., ATP1B1, ENO1, and ENO2), suggesting underlying mechanisms for the shared vulnerability to cell death in PD. Neurons3 also expressed genes associated with the dopaminergic synapse, such as GNAS, CALM1, and CALM3. DEG analysis found carbonic anhydrase II (CA2) and NEAT1 to be significantly upregulated in PD Neurons3. An increased level of CA2 in mitochondria has been reported in aging and neurodegeneration. NEAT1, encoding long coding RNA, has also been reported to be upregulated in PD.

The transcriptomic profiles of the overrepresented neuronal subpopulations, Neurons2 and Neurons4, were also explored. Neurons2 showed high expression of genes associated with GABAergic synapse (e.g., PRKCE, SYN3, GABRG3, and CNTNAP4), glutamatergic synapse (e.g., GRID2, GRIK2, and GRM7), cholinergic synapse (e.g., KCNJ3, CACNA1C, and CAMK4), and serotoninergic synapse (e.g., HTR2C, KCNJ3, CACNA1C, and CACNA1B). Neurons4 exhibited high expression of genes associated with GABAergic synapse (e.g., PLCL1, GNG7, GLUL, and GPHN), glutamatergic synapse (e.g., ERBIN, IL1RAPL1, DLG1, and PTPRD), and cholinergic synapse (e.g., GNG7, BCL2, ITPR2, and PIK3R1). The strongest expression of the serotonin neuronal marker 5-HT-2C receptor (HTR2C) was observed in Neurons2, whereas Neurons4 had the strongest expression of the cholinergic neuronal marker choline transporter-like protein 1 (SLC44A1). Both Neurons2 and Neurons4 expressed genes associated with GTPase activity. Moreover, Neurons4 also exhibited strong expression of genes involved in iron transport (e.g., FTL, FTH1, and TF) and myelination (e.g., MBP, PLP1, and DLG1). At the differential gene expression level, only CAMK2G was identified as differentially upregulated in PD Neurons2, and ANAPC16, SEPT8, and HP1BP3 were upregulated in PD Neurons4. The upregulation in calcium/calmodulin-dependent protein kinase II gamma (CAMK2G) suggests calcium signaling dysregulation in PD Neurons2.

### 6.2.4 Astrocytes Subpopulations in Aged Human SNpc and Their Response In Sporadic PD

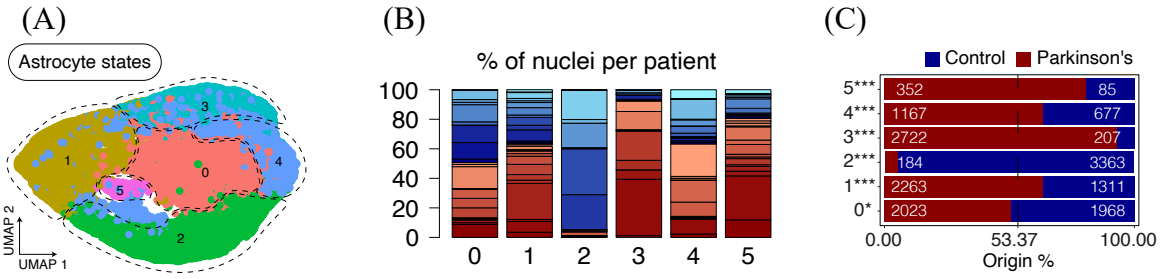



**Figure 6.14:** (A) Re-clustering of astrocytes as identified in Figure 1.2. (B) Distribution of nuclei proportions between astrocytes types from Control (blue) and PD (red) patients, as presented in Figure 3.3A. (C) Comparative analysis of nuclei counts (expressed as a percentage) originating from Control and PD SNpc tissue samples, contrasted with predicted abundance (calculated from 53.37% of all nuclei from PD samples: indicated by the dashed line). A binomial test was conducted to assess the presence of significant deviations in cell proportions from this value, with significance levels marked as follows: * p-value < 0.05, **

p-value < 0.0005, *** p-value < 0.00005. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

The astrocyte population was re-clustered to investigate PD-affected astrocytic subpopulations. Re-clustering revealed six distinct subpopulations, labeled as Astrocytes0 through Astrocytes5. All subpopulations were characterized by unique transcriptomic profiles and were represented in multiple control and PD samples. The nuclei proportion analysis showed that Astrocytes2 and Astrocytes0 were significantly depleted in the PD condition, while the remaining subpopulations were overrepresented.

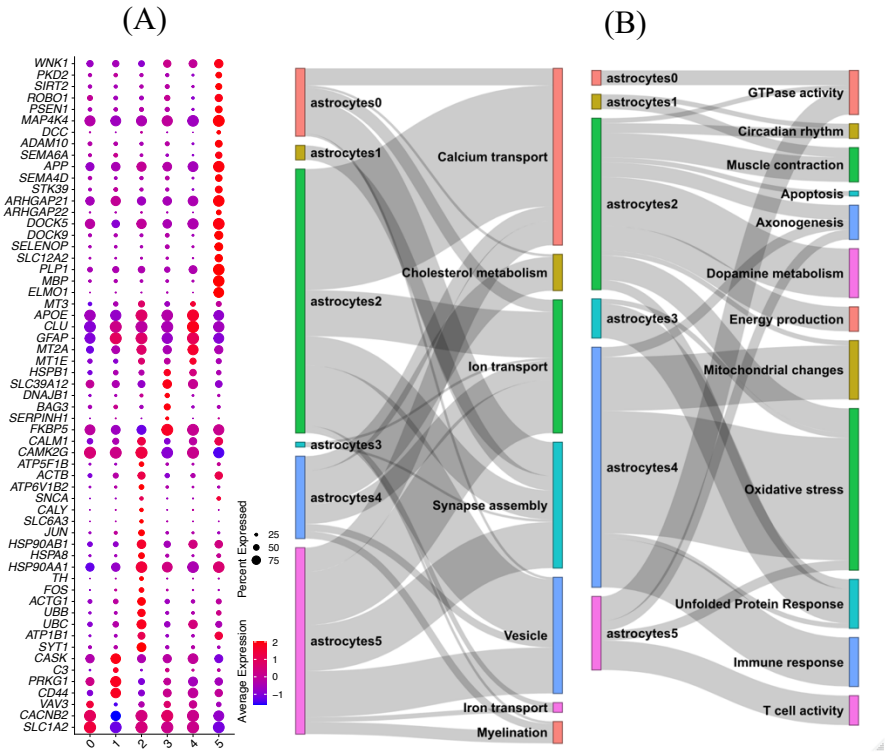



**Figure 6.15**: (A) Dot plot illustrating the expression of cell marker genes within the distinguished astrocytes populations. (B) Sankey plots providing a summarized representation of pathway enrichment analysis results concerning the marker genes characterizing astrocytes populations 0 through 5. **Note:** These plots were generated by Araks Martirosyan, Francisco Pestana, and Silvia Hnatova.

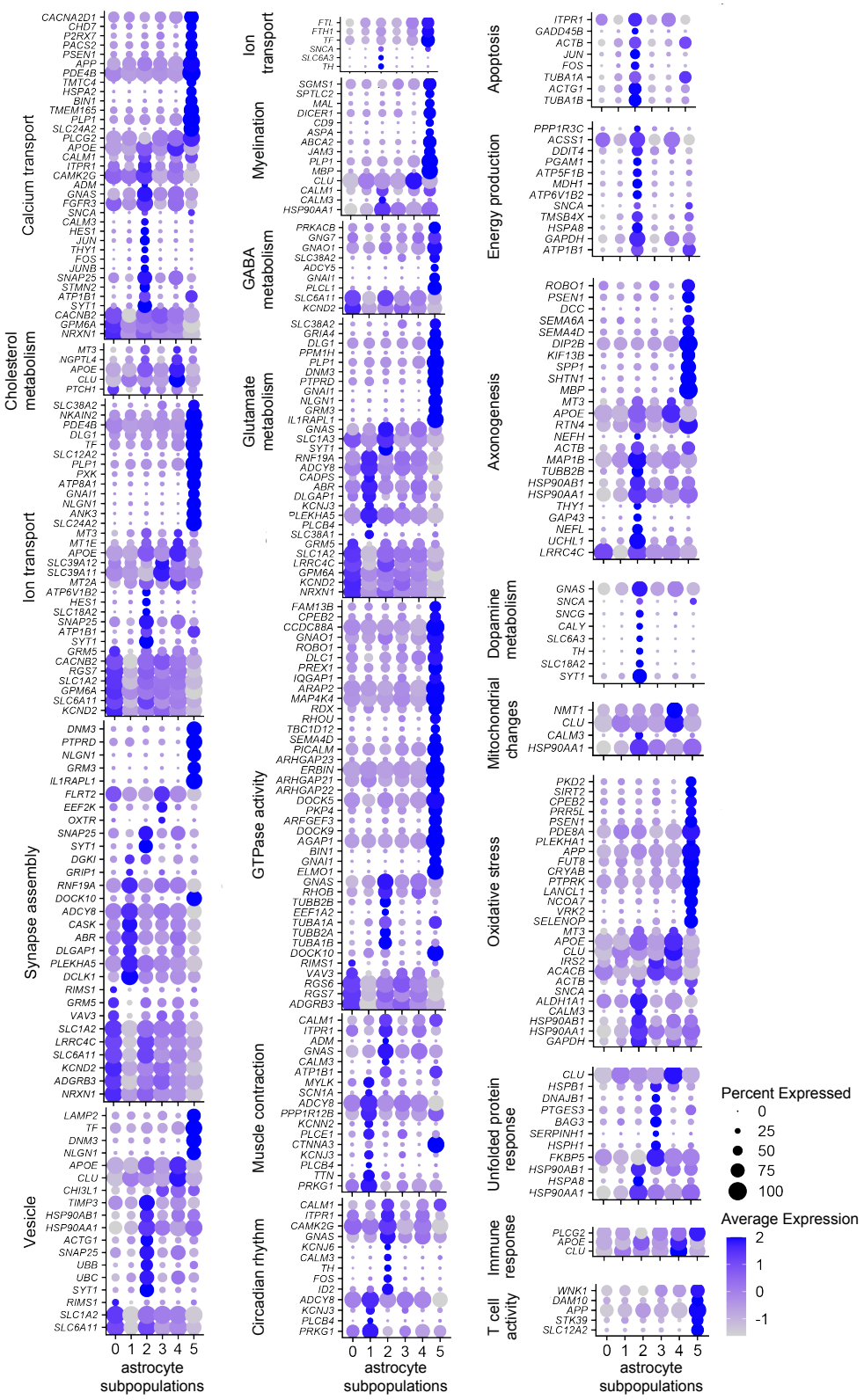



**Figure 6.16:** Cell markers for astrocytes populations categorized according to enriched terms outlined in figure 3.4. Depiction of AGTR1 expression across various astrocytes populations. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

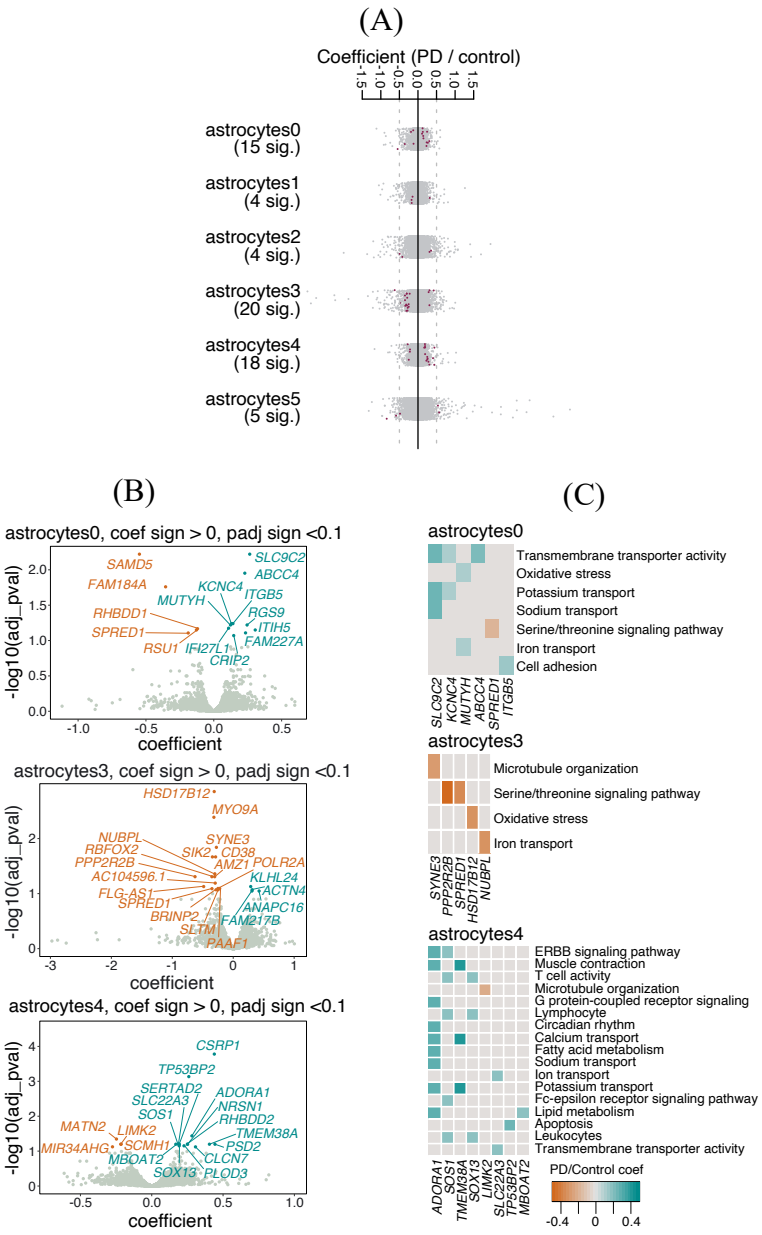



**Figure 6.17**: (A) A stripe chart presenting the count of significantly upregulated or downregulated genes within nuclei from PD donors across distinct neuronal populations, identified through the application of a linear mixed model. Red dots denote genes displaying upregulation (coefficient > 0) or downregulation (coefficient < 0) with an adjusted p-value of less than 0.1 (ANOVA test with Benjamini-Hochberg correction). (B) A volcano plot illustrating genes exhibiting upregulation (cyan) or downregulation (orange) within nuclei originating from PD donors and assigned to the Neurons0 subpopulation. (C) GO and KEGG pathways that are overrepresented by genes displaying upregulation or downregulation within nuclei originating from PD donors and associated with the Neurons0 subpopulation. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

Astrocytes2 were significantly depleted in PD samples (p-value < 0.00005) and showed the expression of genes associated with dopamine metabolism, such as TH, SLC6A3, SLC18A2, and SNCA. This finding suggests that the vulnerability of TH-expressing neurons in PD may extend to other cell types. Additionally, Astrocytes2 expressed transcripts linked to endocytic vesicle processes (e.g., UBB and UBC), synapse assembly (e.g., SYT1 and SNAP25), and calcium ion transport (e.g., CAMK2G, CALM1, and CALM3). Furthermore, Astrocytes2 exhibited genes involved in oxidative stress and UPR (e.g., CALM3, HSP90AA1, HSP90AB1, and HSPA8), as well as apoptosis (e.g., TUBA1B, ACTG1, TUBA1A, FOS, and JUN), suggesting that oxidative stress and UPR may contribute to cell death in this population. Differential gene expression analysis showed the upregulation of ABHD6 and MED14, and the downregulation of TTL4 and UST in PD Astrocytes2 population.

Astrocytes3 were mostly present in PD samples, suggesting that this may be a PD-specific astrocytic population (2722 nuclei in PD condition vs 207 nuclei in Control samples). Astrocytes showed strong enrichment for transcripts related to UPR (e.g., HSPB1, DNAJB1, BAG3, and HSPH1), suggesting reactive astrocyte identity. The expression of genes involved in fatty acid metabolism (e.g., ABHD2, ABHD3, PTGES3, and ADIPOR2) and synapse assembly (e.g., FLRT2 and EEF2K). At the differential gene expression level, PD Astrocytes3 showed the downregulation of genes involved in iron transport (NUBPL), the serine/threonine signaling pathway (PPP2RB and SPRED1), microtubule organization (HSD17B12), and the upregulation of ubiquitin ligase complex genes (KLHL24 and ANAPC16). ANAPC16 was also found to be upregulated in PD Neurons4, the neuronal population that was overrepresented in PD samples. RBFOX2, a splicing factor, was also found to be downregulated in PD Astrocytes3.

Astrocytes5, overrepresented in PD samples, expressed transcripts associated with axon development (e.g., MBP, ANK3, APLP1, and SHTN1), axon ensheathment (e.g., JAM3, MBP, NDRG1, and PLP1), neuron project development (e.g., IL1RAPL1, MBP, NLGN1, and DNM3), and synapse organization (e.g., IL1RAPL1, NLGN1, ANK3, and DNM3), indicating that this astrocytic subpopulation has an active role in neuronal maintenance and survival. Astrocytes5 cell marker genes showed enrichment for ion transport (e.g., SLC38A2, SLC12A2, and NKAIN2), calcium transport (CACNA2D1, CHD7, and P2RX7), endocytic vesicle trafficking (LAMP2, TF, and NLGN1), cholesterol metabolism (NGPTL4, APOE, and CLU),

GABA metabolism (PRKACB, GNG7, and GNAO1), and glutamate metabolism (SLC38A2, GRIA4, and DLG1), suggesting active support of this astrocytic subpopulation to neurons. Interestingly, Astrocytes5 showed strong enrichment for oxidative stress (e.g., CRYAB, SIRT2, and CPEB2), T cell activity (e.g., WNK1, DAM10, and STK39), and GTPase activity (e.g., ARHGAP21, ARHGAP25, and ARHGAP23). At the differential gene expression level, PTOV1, a gene reported to promote cell proliferation, was upregulated, and CYP7B1, a gene involved in lipid homeostasis, was found to be downregulated in PD Astrocytes5.

Astrocytes1 were characterized by the expression of the well-known reactive astrocytic markers CD44 and CD3, perhaps suggesting another reactive astrocyte state. Genes involved in synapse assembly (e.g., RNF19A, ADCY8, and DCLK1) and glutamate metabolism (ABR, DLGAP1, and PLCB4) were highly expressed in Astrocytes1. Differential gene expression analysis suggested dysregulation of tumour necrosis factor (TNF)-mediated signalling due to the downregulation of CCDC3 and the upregulation of NOL3 in PD Astrocytes1.

Astrocytes4 were characterized by the high expression of GFAP, a canonical marker of astrogliosis, APOE, CLU, and the expression of several metallothionein genes (e.g., MT3, MT1E, and MT2A). Pathway analysis on the cell markers of Astrocytes4 showed enrichments for mitochondrial changes, UPR, oxidative stress, and immune response. At the differential gene expression level, PD astrocytes showed upregulation in several genes involved in lipid metabolism (MBOAT2 and ADORA1), calcium transport (TMEM38A and ADORA1), T cell activity (SOS1 and SOX13), and downregulation in LIMK2, involved in microtubule organization.

Astrocytes0 was another subpopulation depleted in PD samples, although at a lower statistical significance level than Astrocytes2 ($p$-value $< 0.05$ in Astrocytes0, $p$-value $< 0.00005$ in Astrocytes2). Astrocytes0 showed the expression of genes involved in synapse assembly (e.g., SLC1A2, LRRC4C, and NRXN1), vesicle transport (e.g., RIMS1 and SLC6A11), and GTPase activity (e.g., RGS6, RGS7, and ADGRB3). Differential gene expression analysis showed the upregulation of MUTYH, a gene associated with oxidative stress and iron transport, and several genes associated with transmembrane transporter activity (SPRED1, KCNC4, and SLC9C2).

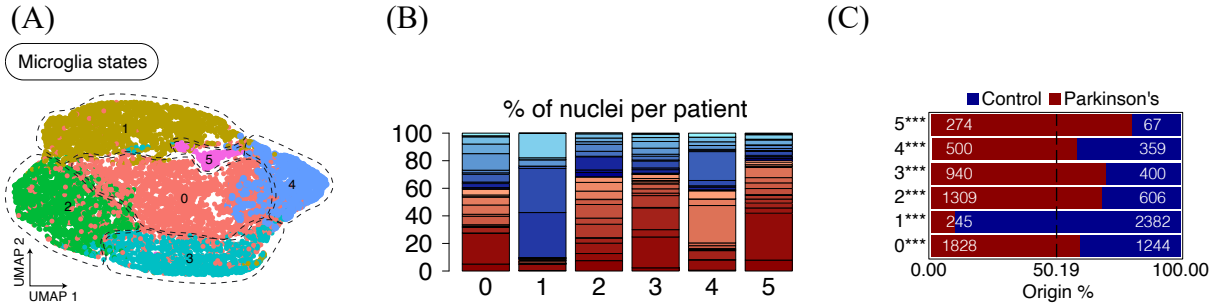



**Figure 6.18**: (A) Re-clustering of microglia as identified in Figure 1.2. (B) Distribution of nuclei proportions between microglia types from Control (blue) and PD (red) patients, as presented in Figure 3.3A. (C) Comparative analysis of nuclei counts (expressed as a percentage) originating from Control and PD SNpc tissue samples, contrasted with predicted abundance (calculated from 50.19% of all nuclei from PD samples: indicated by the dashed line). A binomial test was conducted to assess the presence of significant deviations in cell proportions from this value, with significance levels marked as follows: * p-value < 0.05, ** p-value < 0.0005, *** p-value < 0.00005. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

The re-clustering of the microglial population revealed six distinct subpopulations, designated as Microglia0 to Microglia5. All of these subpopulations were detected in multiple control and PD samples. With the exception of Microglia1, all subpopulations were overrepresented in PD samples.

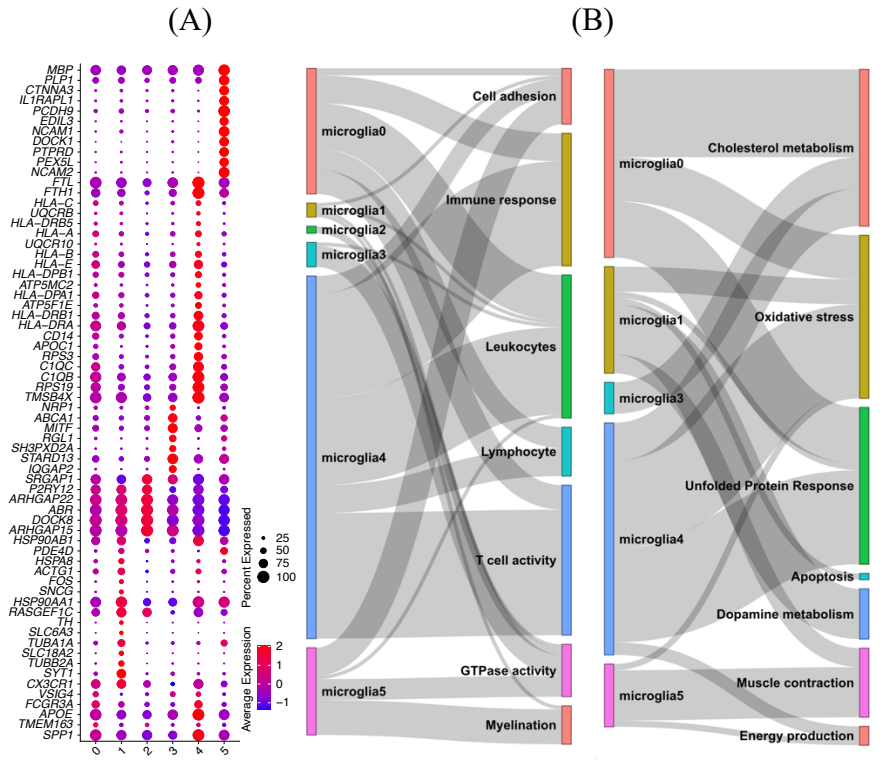



**Figure 6.19:** (A) Dot plot illustrating the expression of cell marker genes within the distinguished microglia subpopulations. (B) Sankey plots providing a summarized representation of pathway enrichment analysis results concerning the marker genes characterizing microglia subpopulations 0 through 5. **Note:** These plots were generated by Araks Martirosyan, Francisco Pestana, and Silvia Hnatova.

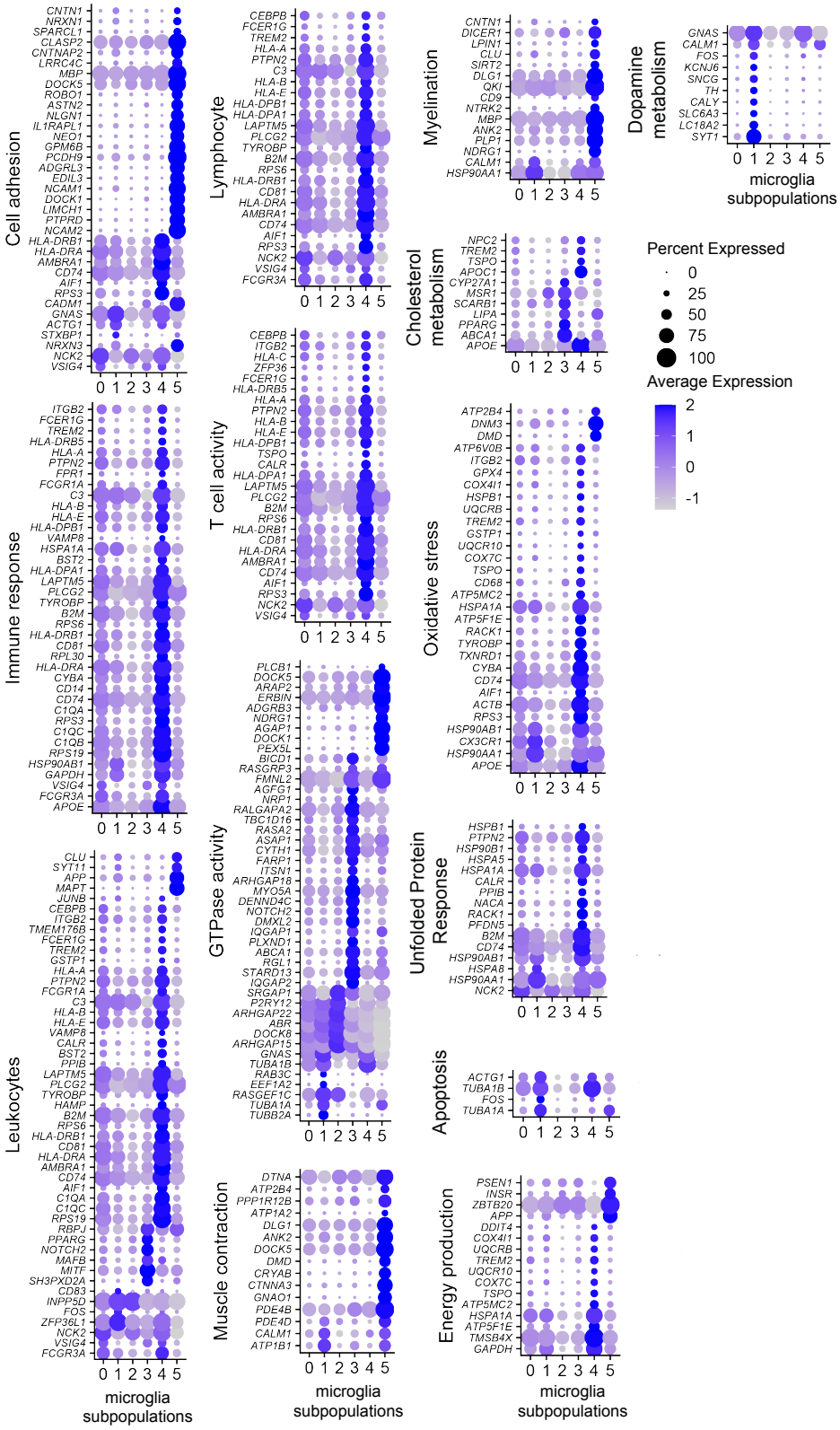



**Figure 6.20:** Cell markers for microglia populations categorized according to enriched terms outlined in figure 3.4. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

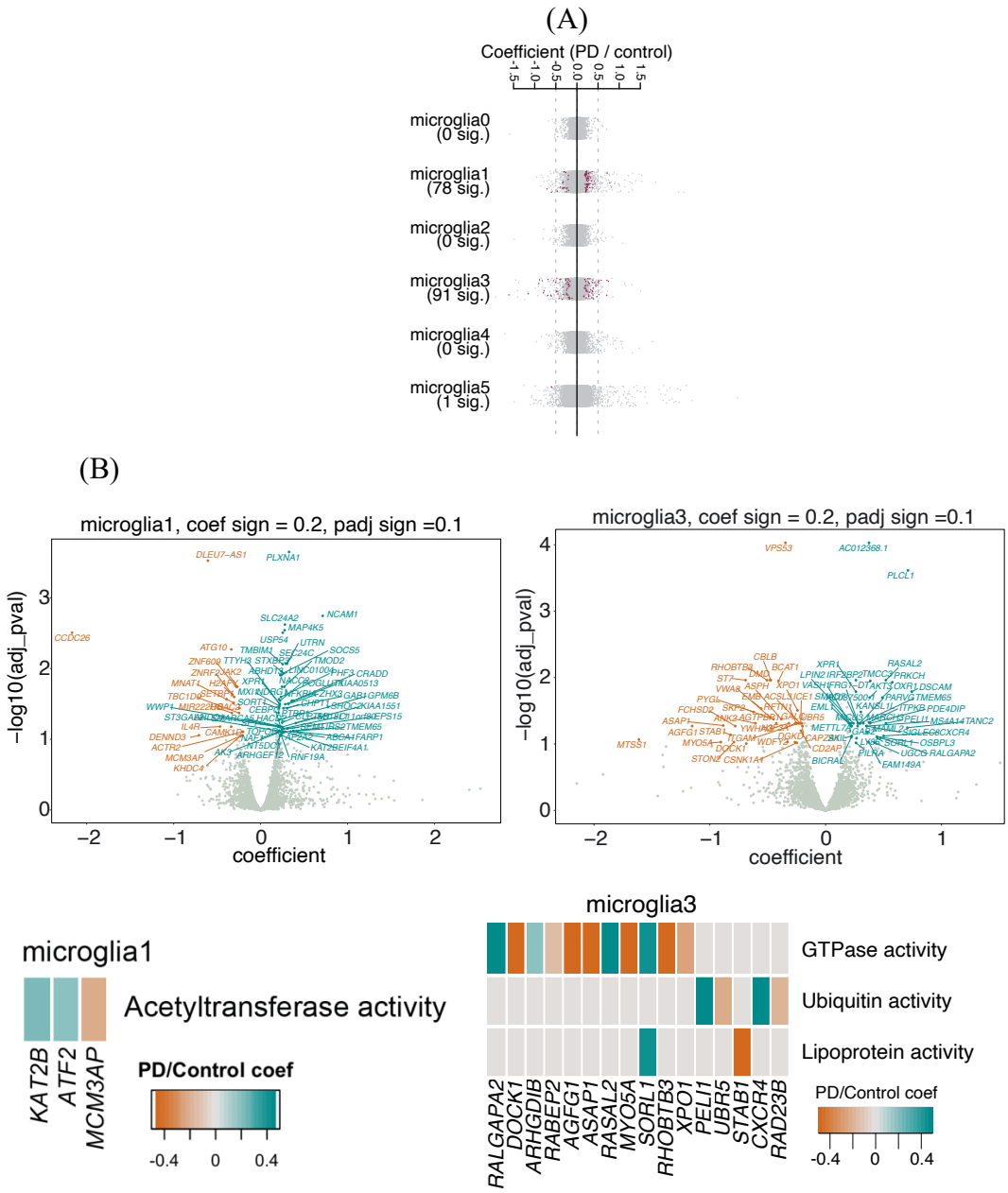



**Figure 6.21:** (A) A stripe chart presenting the count of significantly upregulated or downregulated genes within nuclei from PD donors across distinct microglia subpopulations, identified through the application of a linear mixed model. Red dots denote genes displaying upregulation (coefficient > 0) or downregulation (coefficient < 0) with an adjusted p-value of less than 0.1 (ANOVA test with Benjamini-Hochberg correction). (B) A volcano plot illustrating genes exhibiting upregulation (cyan) or downregulation (orange) within nuclei originating from PD donors and assigned to the microglia subpopulations. (C) GO and KEGG pathways that are overrepresented by genes displaying upregulation or downregulation within nuclei originating from PD donors and associated with the microglia subpopulations. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

Microglia1 were significantly depleted in sporadic PD samples (P value < 0.00005). Similar to Astrocytes2, Microglia1 exhibited the expression of genes associated with dopamine metabolism, such as TH, SLC6A3, and SLC18A2. This microglial subpopulation also highly expressed CX3CR1, a marker of activated microglia. Pathway analysis of the cell marker genes indicated enrichment for oxidative stress and UPR (e.g., HSP90AB1, HSP90AA1, HSPA8, and HSPA1A), as well as apoptosis-related genes (e.g., ACTG1, TUBA1B, FOS, and TUBA1A). Several differentially expressed genes were observed in PD Microglia1. Genes associated with acetyltransferase activity (KAT2B, ATF2, and MCM3AP) were found to be dysregulated in PD Microglia1. Acetyltransferase activity is reported to be involved in the inflammatory response in microglia. Furthermore, genes associated with cholesterol metabolism (SORT1 and ABCA1) were upregulated in PD Microglia1.

Microglia2 were characterized by the expression of P2RY12, a P2Y receptor reported to be involved in microglia motility and migration towards ATP-releasing damaged cells, the initial step of the neuroinflammation process. Microglia2 also highly expressed ARHGAP22 and ARHGAP15, linked to the microglia's activated state during aging, and DOCK8, a gene associated with neuroinflammation. Overall, these results suggest that Microglia2 may represent an inflammatory activated state of the microglia.

Microglia4 were characterized by high expression of SPP1, APOE, FTL, and FTH1. Microglia4 appeared to represent an activated state of the microglia, expressing genes involved in the human leukocyte antigen (HLA) system (e.g., HLA-DPA1, HLA-DPB1, HLA-DRA, and HLA-DRB1), the complement cascade (e.g., C1QC, C1QB, and C1QA), oxidative stress response (e.g., RPS3, TREM2, GSTP1, and HSPB1), and unfolded protein response (e.g., HSP90B1, HSPA5, PTPN2, and HSPA1A). Microglia0 also expressed genes involved in oxidative stress, UPR, and immune response, but to a lesser extent than microglia4. Moreover, both Microglia4 and Microglia0 expressed genes associated with cholesterol metabolism. Taken together, these results indicate that the immune response in Microglia2 and Microglia4 may be driven by oxidative stress and UPR. Interestingly, differential gene expression was not observed in the Microglia0, Microglia2, and Microglia4 subpopulations.

Microglia3 were characterized by the expression of STARD13, MITF, and ABCA1. Pathway analysis of the cell marker genes in this population highlighted enrichments for

GTPase activity (e.g., MYO5A, ITSN1, ASAP1, and FARP1) and cholesterol metabolism (e.g., MSR1, LIPA, ABCA1, and SCARB1). At the differential gene expression level, PD Microglia3 showed dysregulation of genes involved in GTPase activity (e.g., RALGAPA2, DOCK1, AGFG1, and ASAP1), lipoprotein activity (SORL1 and STAB1), and ubiquitin activity (UBR5, PELI1, RAD23B, and CXCR4).

### 6.2.6 Oligodendrocyte Subpopulations in Aged Human SNpc and Their Response In Sporadic PD

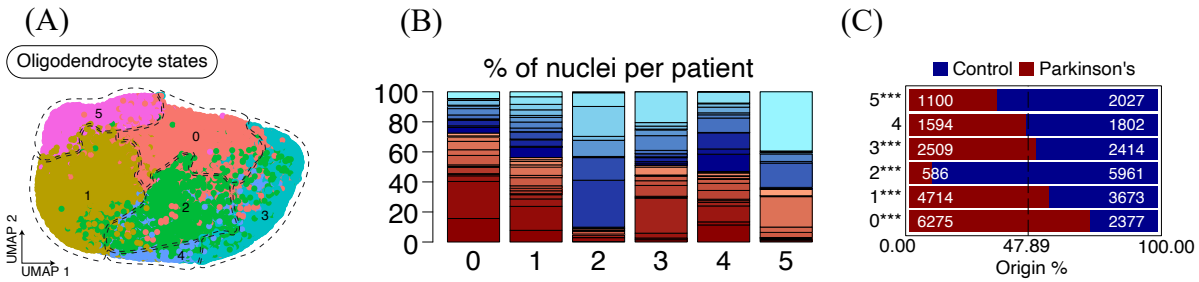



**Figure 6.22:** (A) Re-clustering of oligodendrocytes as identified in Figure 1.2. (B) Distribution of nuclei proportions between oligodendrocytes types from Control (blue) and PD (red) patients, as presented in Figure 3.3A. (C) Comparative analysis of nuclei counts (expressed as a percentage) originating from Control and PD SNpc tissue samples, contrasted with predicted abundance (calculated from 53.37% of all nuclei from PD samples: indicated by the dashed line). A binomial test was conducted to assess the presence of significant deviations in cell proportions from this value, with significance levels marked as follows: * p-value < 0.05, ** p-value < 0.0005, *** p-value < 0.00005. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

The re-clustering of the oligodendrocyte population also revealed six distinct subpopulations, designated as Oligos0 to Oligos5. All subpopulations were present in multiple Control and PD samples. Nuclei proportion analysis indicated that Oligos2 and Oligos5 (P value < 0.00005) were significantly depleted in PD samples, while Oligos0, Oligos1, and Oligos3 were overrepresented (P value < 0.00005). Oligos4 did not exhibit a statistically significant increase or decrease in number in PD samples.

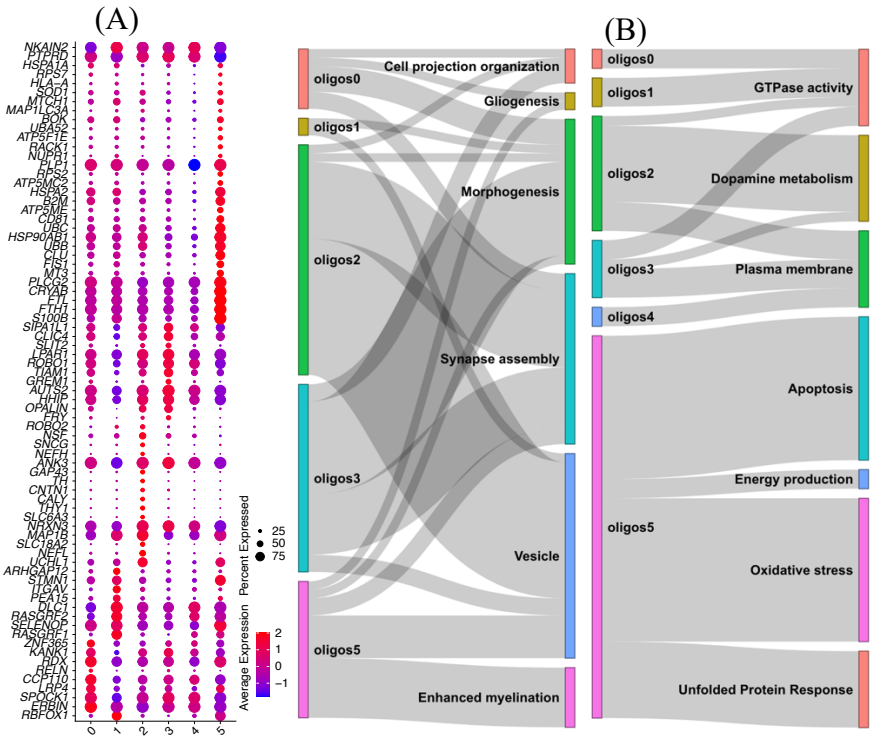



**Figure 6.23**: (A) Dot plot illustrating the expression of cell marker genes within the distinguished oligodendrocyte populations. (B) Sankey plots providing a summarized representation of pathway enrichment analysis results concerning the marker genes characterizing oligodendrocyte populations 0 through 5. **Note:** These plots were generated by Araks Martirosyan, Francisco Pestana, and Silvia Hnatova.

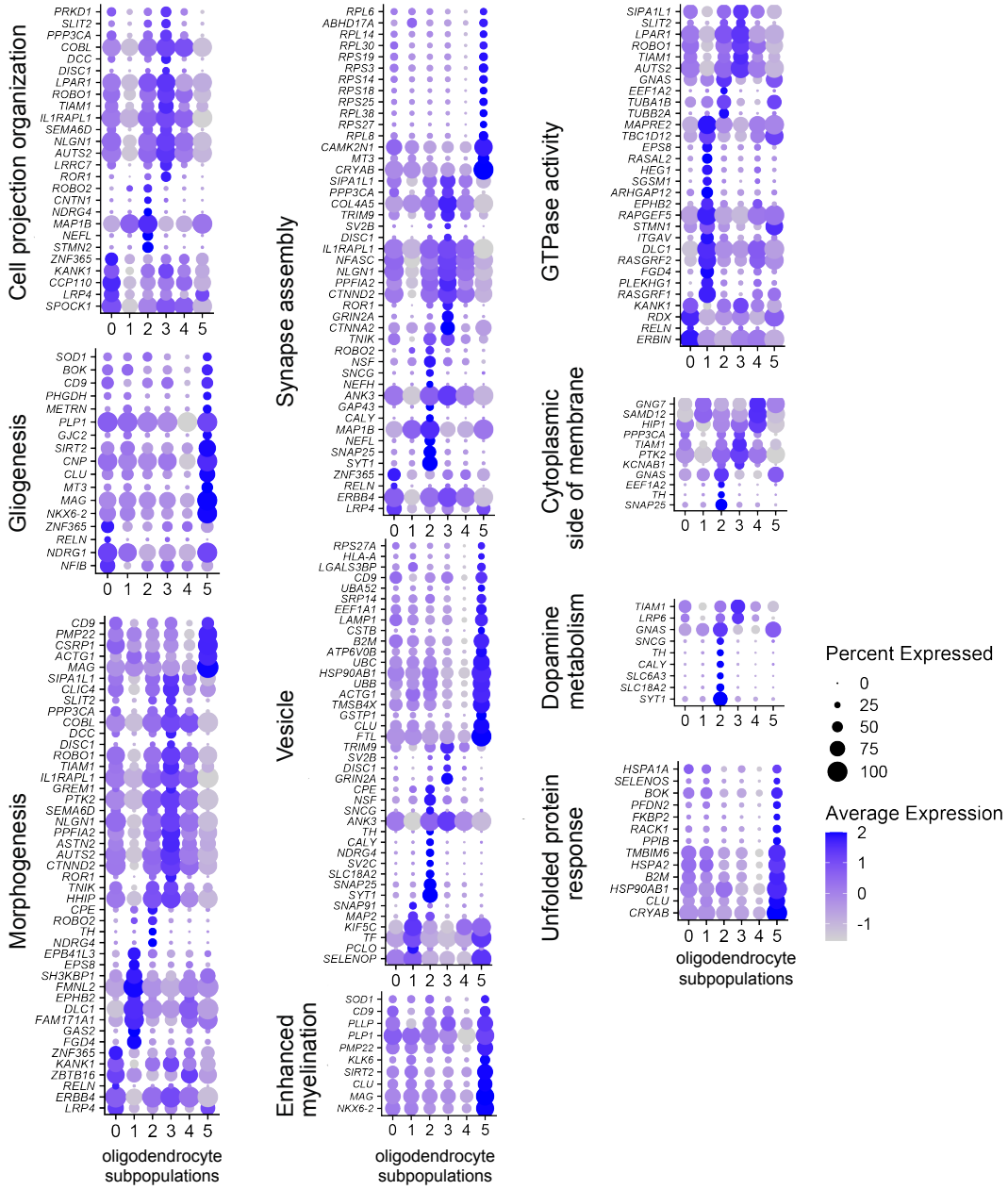



**Figure 6.24**: Cell markers for oligodendrocyte populations categorized according to enriched terms outlined in figure 3.4. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

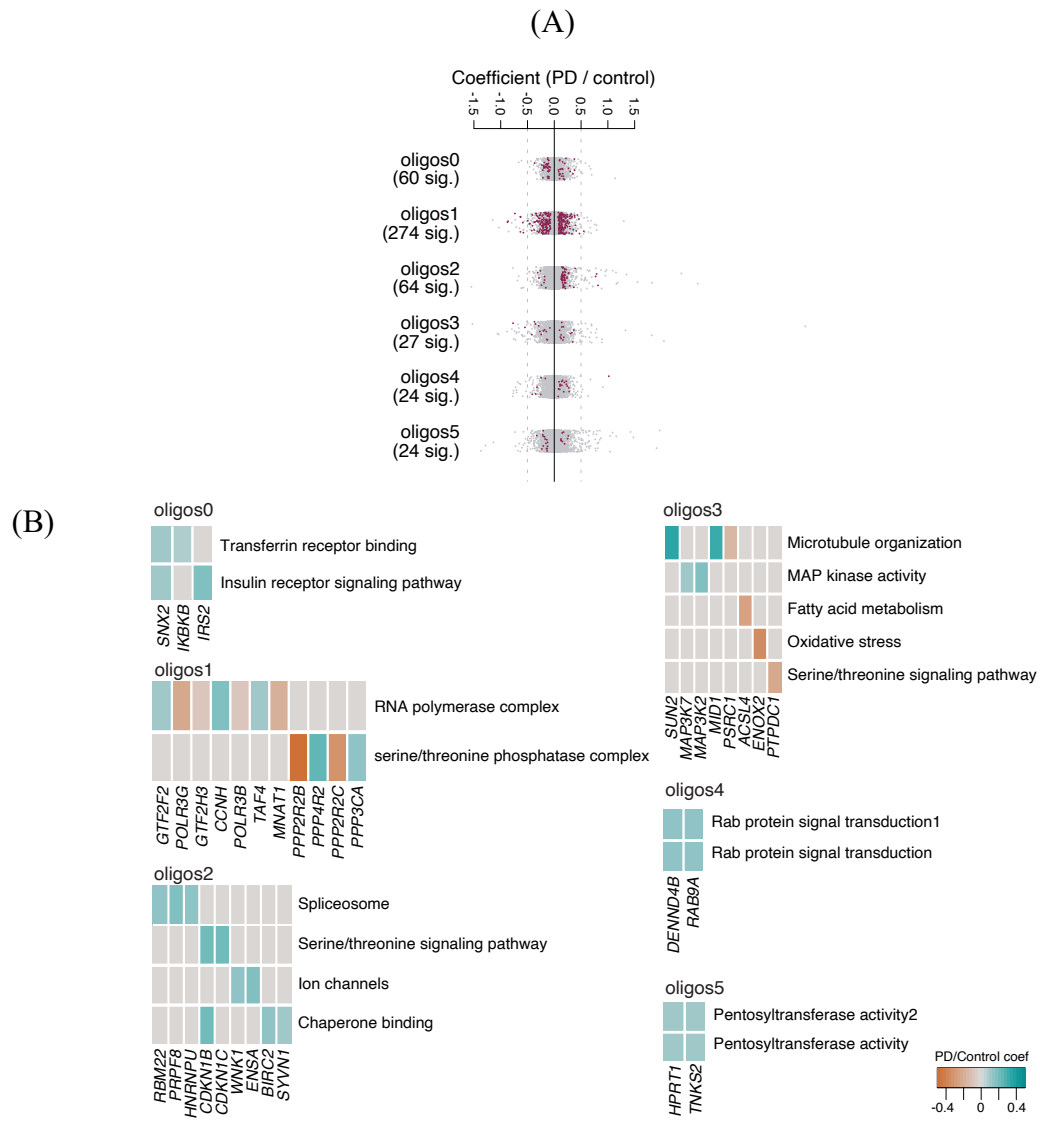



**Figure 6.25**: (A) A stripe chart presenting the count of significantly upregulated or downregulated genes within nuclei from PD donors across distinct oligodendrocytes populations, identified through the application of a linear mixed model. Red dots denote genes displaying upregulation (coefficient > 0) or downregulation (coefficient < 0) with an adjusted p-value of less than 0.1 (ANOVA test with Benjamini-Hochberg correction). (B) GO and KEGG pathways that are overrepresented by genes displaying upregulation or downregulation within nuclei originating from PD donors and associated with the oligodendrocytes subpopulation. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

Similar to Microglia1 and Astrocytes2, Oligos2 also expressed genes associated with dopamine metabolism (TH, SLC6A3, and SLC18A2) and was found to be depleted in PD samples. Oligos2 exhibited a strong expression of genes linked to synapse assembly (e.g., SNAP25, SYT1, and NEFL), vesicle trafficking (e.g., TRIM9, SV2B, and GRIN2A), cell

projection development (e.g., ROBO2, CNTN1, and NDRG4), and ion transport (e.g., CNTN1 and ANK3). This subpopulation also highly expressed UCHL1, a gene associated with PD. Interestingly, unlike Microglia1 and Astrocytes2, pathway analysis on cell marker genes did not reveal enrichment for transcripts associated with oxidative stress and unfolded protein response. Several genes were found to be differentially expressed in this population. Among them, Oligos2 showed upregulation of genes involved in the spliceosome (PRPF8, RBM22, and HNRNPU), the serine/threonine signaling pathway (CDKN1B and CDKN1C), ion channels (ENSA and WNK1), and chaperone binding (BIRC2, SYVN1, and CDKN1B). Furthermore, calmodulin-dependent protein kinase II gamma (CAMK2G) and translocase of outer mitochondrial membrane 40 (TOMM40) genes were observed to be upregulated in PD Oligos2. Dysregulation of CAMK2G is in line with studies showing alterations in calmodulin-dependent protein kinase signaling in PD.

Oligos5 were found to be significantly depleted in PD samples. Oligos5 were characterized by high expression of CRYAB, a small heat shock protein (sHSP) reported to be involved in several protein aggregation-related neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS), PD, and Alzheimer's disease. Oligos5 also highly expressed iron storage genes (FTH1 and FTL), as well as S100, a gene linked to glial stress response. Pathway overrepresentation analysis on cell marker genes demonstrated enrichment for gliogenesis (e.g., NKX6-2, MAG, CNP, and SIRT2), synapse assembly (e.g., ABHD17A, CAMK2N1, and RPL8), synaptic vesicle trafficking (e.g., B2M, LAMP1, SRP14, and UBA52), and myelination (e.g., PLLP, PLP1, MAG, and PMP22). Interestingly, Oligos5 also exhibited positive expression for transcripts associated with UPR (e.g., TMBIM6, RACK1, HSPA2, and HSPA1A), oxidative stress (CRYAB, SELENOP, MAP1LC3A, and MT3), and apoptosis (e.g., ACTB, TUBA1A, ACTG1, and TUBA1B), suggesting that UPR and oxidative stress might contribute to the decline of Oligos5 in PD conditions. At the differential gene expression level, PD Oligos5 displayed upregulation of TRF1-interacting ankyrin-related ADP-ribose polymerase 2 (TNKS2) and hypoxanthine guanine phosphoribosyltransferase (HPRT1), both associated with pentosyltransferase activity.

Oligos1 were characterized by high expression of RBFOX1, a risk gene for several neurodevelopmental conditions. Oligos1's cell marker genes exhibited strong enrichment for GTPase activity (e.g., RASGRF1, RASGRF2, RAPGEF5, and ARHGAP2). However, genes associated with myelination, synapse assembly, and vesicle trafficking were not strongly

expressed in this population. At the differential gene expression level, PD Oligos1 displayed dysregulation in genes associated with serine/threonine phosphatase activity (e.g., PPP4R2, PPP2R2B, PPP3CA, and PPP2R2C), protein ubiquitination (e.g., FBXW8, ANAPC16, KLHL21, and SKP2), lysosomal activity (e.g., DMXL1, SORT1, and ARSB), histone acetyltransferase complex function (e.g., JADE2, EP300, CREBBP, and TAF4), and RNA polymerase complex function (e.g., CCNH, POLR3G, GTF2F2, and GTF2H3).

Oligos3 were found to be overrepresented in PD samples and characterized by the expression of the myelinating oligodendrocyte marker OPALIN. Pathway analysis on cell marker genes exhibited strong enrichment for genes involved in synapse assembly (e.g., CTNNA2, GRIN2A, COL4A5, and ROR1), cell projection organization (e.g., NLGN1, TIAM1, ROR1, and SEMA6D), and GTPase activity (AUTS2, ROBO1, LPAR1, and SIPA1L1). Differential gene expression analysis highlighted several dysregulated genes associated with microtubule organization (SUN2, MID1, and PSRC1), MAP kinase activity (MAP3K7 and MAP3K2), fatty acid metabolism (ACSL4), oxidative stress (ENOX2), and the serine/threonine signaling pathway (PTPDC1).

### 6.2.7 TH Among Several Common Genes Expressed in Depleted Cell Types In Sporadic PD Samples

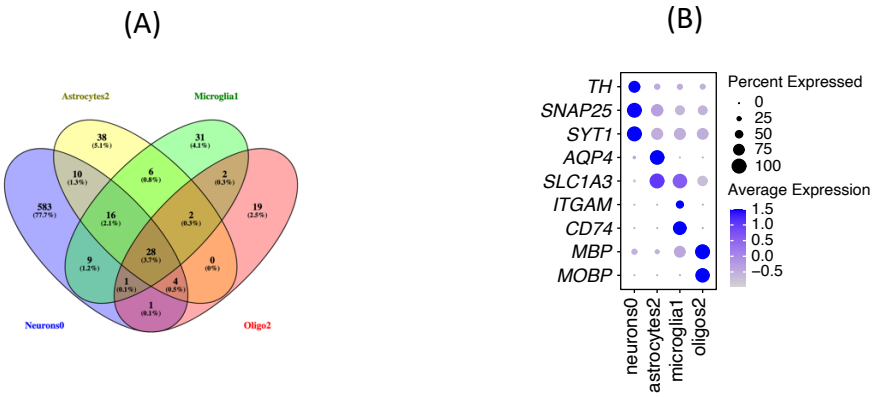



**Figure 6.26:** (A) A comparison of TH expression between dopaminergic Neurons0 and TH-expressing glial cell types (Astrocytes2, Microglia1, and Oligos2). Higher-level cell markers were utilized to validate the identity of these subpopulations. (B) Venn diagram showing overlap of cell markers defining dopaminergic Neurons0 and TH expressing glial cell types. **Note:** Dot plot was generated by Araks Martirosyan and Francisco Pestana.

In addition to the significant depletion of TH-positive neurons in PD samples, TH-positive astrocytes, oligodendrocytes, and microglia were also identified as being depleted in PD samples. To investigate potential shared molecular mechanisms of action, the transcriptomic profiles of these four TH-positive subpopulations were compared. Firstly, the TH expression levels were compared among these subpopulations, revealing that TH was expressed at relatively lower levels in the glial subpopulations compared to Neurons0. Furthermore, a lower percentage of glial cells exhibited TH expression when compared to Neurons0. The confirmation of the identity of these TH-positive subpopulations was reaffirmed using cell type-specific markers. For instance, Astrocytes2 co-expressed TH and the astrocyte marker AQP4, Microglia1 co-expressed TH along with the microglia markers ITGAM and CD74, Oligos2 co-expressed TH alongside the oligodendrocyte markers MBP and MOBP, and Neurons0 co-expressed TH along with the neuronal markers SYT1 and SNAP25.

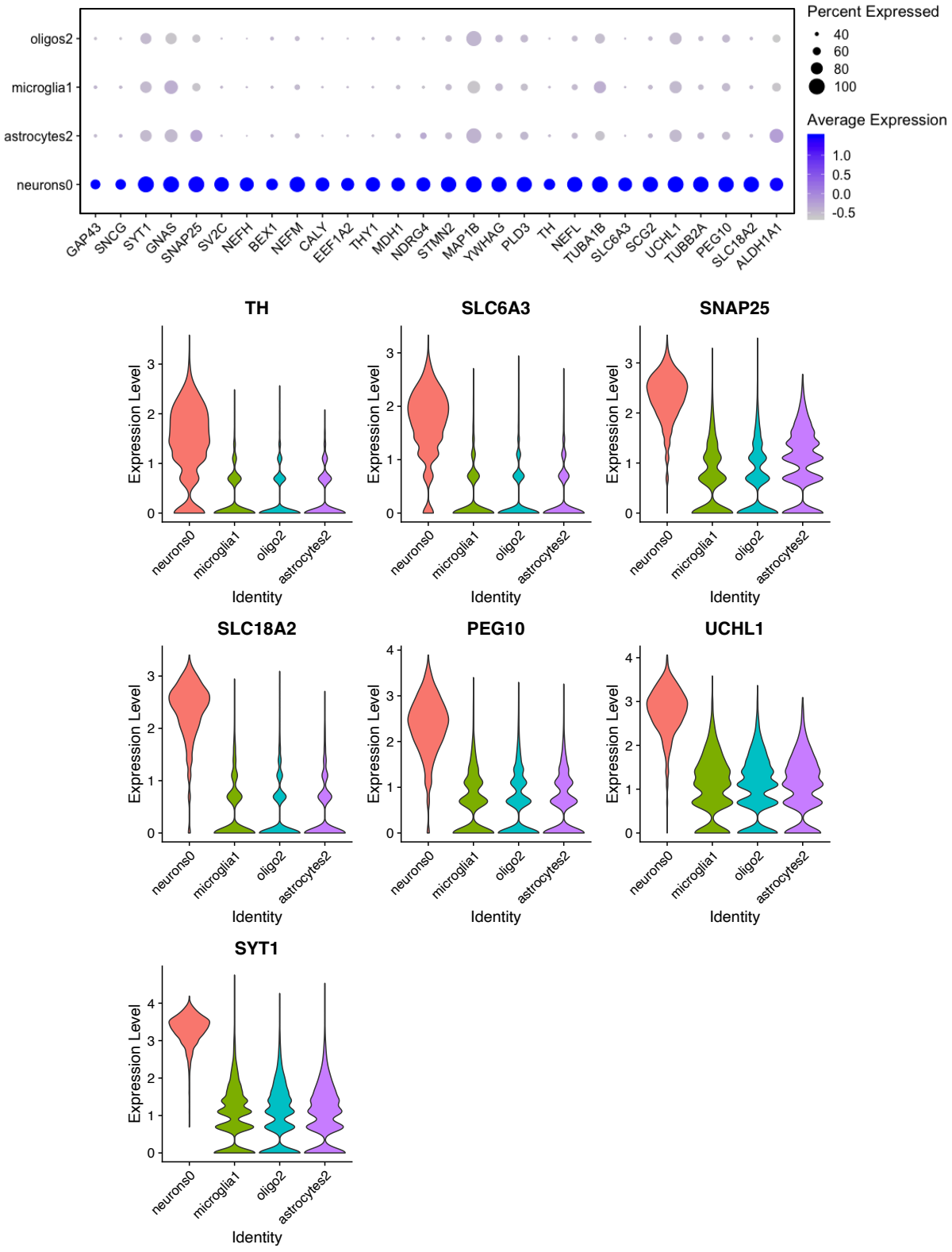



**Figure 6.27:** Gene expression levels of cell markers shared between dopaminergic Neurons0 and TH expressing glial cell types. **Note:** These plots were generated by Silvia Hnatova and William Plumbly.

A comparative analysis of the cell marker genes from each of the TH-positive subpopulations revealed 28 genes, including PD-associated genes such as GNAS, TUBB2A, TUBA1B, and UCHL1, that were shared among the PD-depleted subpopulations. Similar to the TH expression levels, all of these shared genes were expressed in a low percentage of cells and at a low level compared to Neurons0.

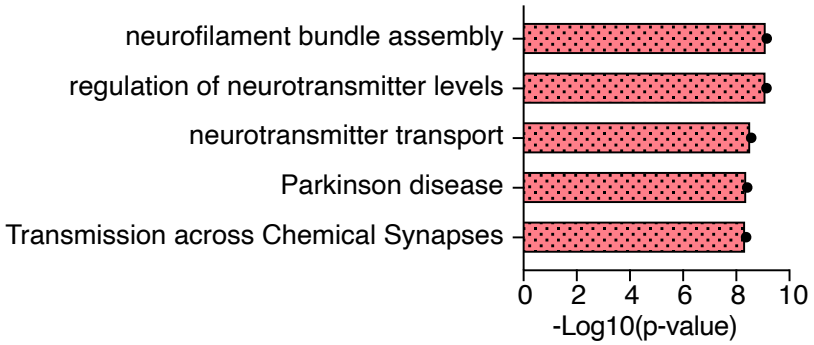



**Figure 6.28:** GO term pathway analysis on the shared cell markers.

Pathway analysis of the shared genes revealed enrichment in Gene Ontology (GO) terms, including dopaminergic neurogenesis (e.g., TH, SLC6A3, SLC18A2, ALDH1A1), neurofilament assembly (e.g., NEFM, NEFH, and NEFL), and regulation of neurotransmitter levels and transport (e.g., SNCG, SLC18A2, SLC6A3, and SNAP25).

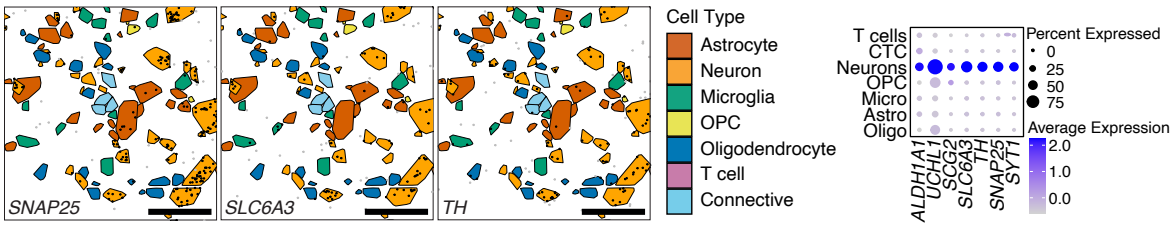



**Figure 6.29:** (A) Representative spatial transcriptomic images showing presence of TH, SNAP25 and SLC6A3 transcripts (dots) in the higher-order cell types. OPC, oligodendrocyte progenitor cells. Scale bar, 50 µm. (B) Gene expression level confirmation of shared cell markers between the higher-order cell types using spatial transcriptomics. CTC, connective tissue cells, Micro, microglia; Astro, astrocytes; Oligo, oligodendrocytes. **Note:** These plots were generated by Araks Martirosyan and Francisco Pestana.

TH expression was confirmed in all major cell types detected in the SNpc tissues obtained from both PD and Control patients using the Molecular Cartography spatial transcriptomics platform developed by Resolve Biosciences. The spatial transcriptomic results

were consistent with the single nuclei RNA sequencing data, as the shared markers such as TH, SLC6A3, UCHL1, SYT1, and ALDH1A1 were expressed at relatively low levels in glial cells compared to neurons.

## 6.3 Discussion and Future Directions

In this study, we have presented extensive single-nucleus RNA sequencing (snRNA-seq) datasets from 14 control individuals and 15 sporadic PD patients' substantia nigra pars compacta (SNpc) brain tissues. This study represents a significant advancement toward the development of new *in vitro* and *in vivo* models of PD.

### 6.3.1 Cell Types Composition in SNpc

The dataset encompasses transcriptomic profiles of more than 80,000 high-quality nuclei, enabling the identification of all major cell types within the human brain. Our analysis revealed the presence of oligodendrocytes (~41%), astrocytes (~25%), microglia (~15%), neurons (~7%), OPCs (~8%), and T cells (<2%) in our dataset. Comparing cell proportions between control and PD samples unveiled a decrease in neurons and an increase in glial cells in the PD condition.

Wang et al.'s studies reveal a depletion of dopaminergic neurons in PD conditions, while other non-neuronal and glial cells show minimal changes. In their dataset, oligodendrocytes (~51%) were the most abundant cell type, followed by neurons (13.1%), microglia (9.4%), astrocytes (8.4%), with the remaining cell types constituting approximately 18% of the population in the substantia nigra. In Smajic et al.'s dataset, authors observed an increase in microglia and astrocytes but a decrease in oligodendrocytes in PD conditions. In terms of the neuronal population, they noted an increase in excitatory neurons and dopaminergic neurons, coupled with a decrease in inhibitory and GABAergic neurons in PD conditions. Given that their dataset comprised only 0.18% of the total nuclei (74 cells), they performed additional quantitative immunofluorescence analysis to confirm the depletion of TH-positive/neuromelanin-positive dopaminergic neurons in their PD samples. Oligodendrocytes (51.32%) were the most abundant cell type in Smajic's dataset as well, followed by neurons (12.82%), astrocytes (12.6%), microglia (9.41%), with the remaining cell types making up 13.85% of the total population in the midbrain brain.

Oligodendrocytes were consistently the most abundant cell type in all snRNA-seq datasets of the substantia nigra/midbrain. However, there were variations in their proportion changes in PD conditions, warranting further investigation. To our knowledge, only our dataset captured T cells, albeit in low numbers (347 cells), while other discussed studies did not detect any T cells in their samples. T cells are known to be present in the substantia nigra and have been shown to play a key role in neuroprotection in preclinical models of PD (Baird et al., 2019). The analysis of the captured T cells in our dataset may offer insights into their response to PD.

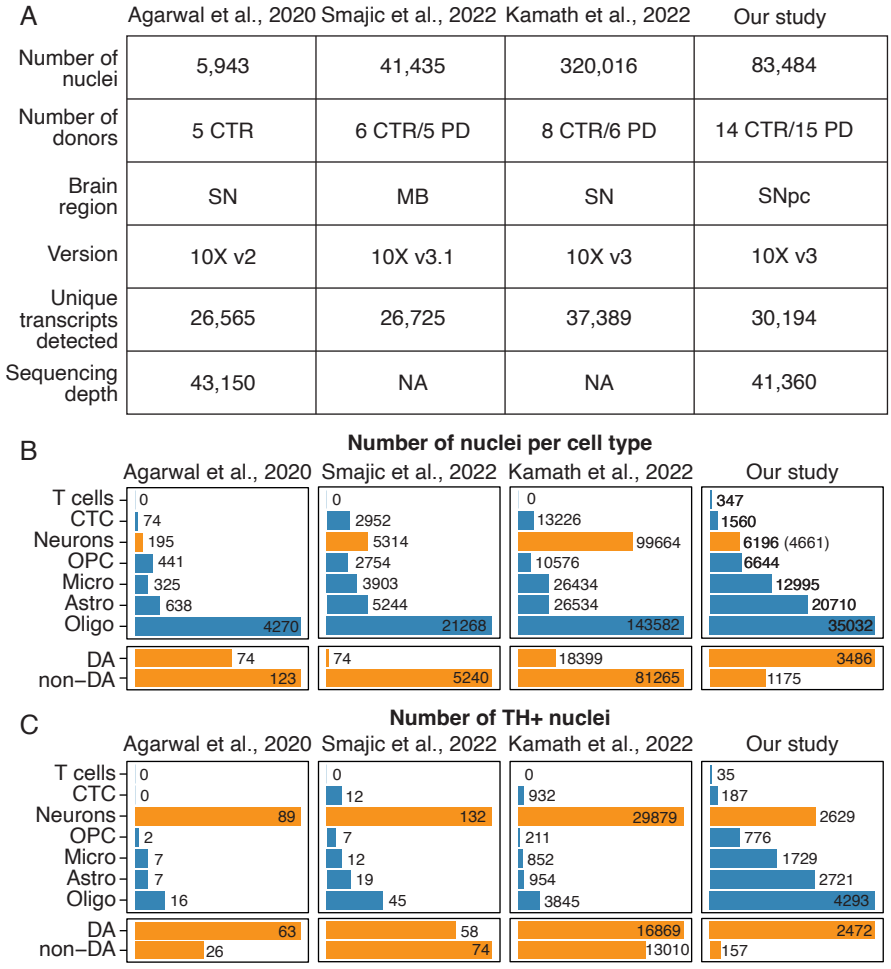



**Figure: Comparative analysis of our substantia nigra snRNA-seq dataset and published datasets on SN/Midbrain.**

Obtaining a high-quality postmortem dataset poses a challenge due to limited access to human brain tissue and difficulties associated with extracting high-quality RNA from postmortem cells, particularly dopaminergic neurons in PD samples. In comparison to

previously published snRNA-seq studies, our study is among the largest in terms of the number of donors included, featuring 15 sporadic PD patients and 14 control individuals. We have generated transcriptomic profiles of more than 80,000 high-quality nuclei, capturing more than 2,500 TH-expressing neurons, with the majority classified as DA neurons in our study. In contrast, Smajic et al.'s study included 6 control and 5 PD patients, profiling ~41,000 nuclei from post-mortem human midbrain tissue. Notably, they employed whole midbrain tissue and did not specifically target the substantia nigra region. They captured only 132 TH-expressing neurons, of which only 58 were classified as DA neurons. Agarwal et al.'s study included substantia nigra tissue from 5 control individuals, profiling ~6,000 nuclei. In their dataset, 89 TH-expressing neurons were detected, with only 63 classified as DA neurons. Kamath et al.'s study included 8 control and 6 PD patient samples, and they enriched dopaminergic neurons by sorting Nurr1-positive cells from the neuromelanin-pigmented region of the substantia nigra. Consequently, their dataset included more than 30,000 TH-expressing neurons, with 17,000 cells classified as DA neurons. These studies underscore the challenges associated with generating high-quality snRNA-seq datasets comprising a substantial number of dopaminergic neurons.

### 6.3.2 Expression Of PD GWAS Associated Genes In SNpc Cell Types

Genome-wide association studies (GWAS) have identified numerous susceptibility loci associated with PD, but the specificity of the genes associated with these loci in SNpc cell types has yet to be demonstrated. To this end, we also explored the expression of 153 putative PD GWAS associated genes in detected cell types in our SNpc dataset. This analysis showed the specificity of PD GWAS associated gene expression to cell types, suggesting involvement of multiple cell types in PD pathology. Confirmation of GWAS PD risk gene expression specificity to disease-relevant cell types provides a robust platform for future studies of PD risk factors using iPSCs-derived models.

### 6.3.3 Dopaminergic Neurons0 Degeneration in PD Condition

Dopaminergic Neurons0 were found to be depleted in PD samples, consistent with existing literature. This specific neuronal subpopulation expressed ubiquitin carboxy-terminal hydrolase L1 (UCHL1), a gene implicated in PD. For instance, an UCHL1 I93M mutant mouse model exhibited degeneration of dopaminergic neurons in the substantia nigra, accompanied by a reduction in dopamine content in the striatum (Setsuie et al., 2007). Furthermore, cells

carrying the I93M mutant UCHL1 displayed dysfunction in chaperon-mediated autophagy and elevated α-synuclein levels (Kabuta et al., 2008). Notably, UCHL1 is one of the components found in Lewy bodies, alongside α-synuclein, and its S-nitrosylation has been shown to induce structural instability, thereby promoting α-synuclein aggregation (Kumar et al., 2017).

Paternally expressed 10 (PEG10), a DNA-binding protein-encoding gene, was another notably highly expressed cell marker. Reports have indicated that the PEG10-encoded protein exhibits increased binding affinity for SNCA-associated enhancer alleles, potentially contributing to SNCA expression and, consequently, PD risk (McClymont et al., 2018). This subpopulation displayed an enrichment of genes associated with PD pathology, including the unfolded protein response (UPR), oxidative stress, energy production, iron transport, response to endoplasmic reticulum stress, and cholesterol metabolism.

Differential gene expression analysis revealed the downregulation of OGA and CHST1 genes, which are associated with sugar/glucose metabolism, as well as the downregulation of RAB6B and COG4 genes, which are linked to Golgi-to-endoplasmic reticulum vesicle trafficking in PD dopaminergic Neurons0. Abnormal glucose metabolism has been reported as an early event in sporadic PD and is considered a non-motor symptom of PD (Dai et al., 2023). Studies have demonstrated that glycosylation in the endoplasmic reticulum (ER) and Golgi is a crucial quality control procedure for regulating protein folding status (Stanley et al., 2011; Bell et al., 2020). OGA is involved in the O-GlcNAcylation process, which has been shown to regulate dopaminergic neuron function and survival in PD. Increases in O-GlcNAcylation through genetic and pharmacological methods have been demonstrated to alleviate neurodegeneration, synaptic damage, and motor symptoms in an *in vivo* model of PD (Lee et al., 2020). Dysregulation of OGA is also associated with insulin resistance. A recent study has shown that insulin resistance can promote PD through aberrant expression of α-synuclein, deregulation of polo-like kinase 2 signaling, and mitochondrial dysfunction (Hong et al., 2020).

### 6.3.3 GABA Neurons3 Degeneration in PD Condition

While the primary focus in Parkinson's disease (PD) has traditionally centered on the loss of dopaminergic neurons and the depletion of dopamine levels, it is increasingly evident that other neurotransmitters, such as GABAergic, glutamate, and serotonin, could play pivotal roles in both motor and non-motor symptoms. In our analysis, we discovered a significant

depletion of GABAergic neurons in PD samples. Like dopaminergic neurons, this particular subpopulation exhibited an enrichment of genes associated with the unfolded protein response (UPR), oxidative stress, energy production, and iron transport. Interestingly, this subpopulation also expressed several genes involved in the dopaminergic synapse.

Recent research has shed light on the fact that within the striatum, axons of dopaminergic neurons release not only dopamine but also GABA. This finding implies that the degeneration of dopaminergic neurons could result in a reduction of GABA within the basal ganglia circuits (Murueta-Goyena et al., 2019). This observation is significant because GABAergic networks play a crucial role in regulating processes that depend on calcium, such as mitochondrial function and oxidative stress. A reduction in GABA inhibition may potentially lead to the accumulation of abnormal levels of intracellular calcium, which, in turn, could trigger neurodegenerative mechanisms.

In line with this concept, previous studies have demonstrated that pharmacological GABA agonists can protect dopaminergic neurons and alleviate motor symptoms in animal models of PD (Lozovaya et al., 2018; Hajj et al., 2015). Moreover, a recent study has shown that PD motor symptoms can be rescued by directly reprogramming astrocytes into GABAergic neurons *in vivo* in the 6-OHDA animal model of PD. Additionally, deep brain stimulation (DBS), a surgical treatment for PD, involves the stimulation of the subthalamic nucleus (STN) or globus pallidus internus (GPi) (Zhang et al., 2021). These specific brain regions house the GABAergic neurotransmitter system, and DBS is believed to modulate their activity to ameliorate PD symptoms (Lanciego et al., 2012; Jakobs et al., 2019).

We also observed an upregulation of NEAT1(nuclear-enriched abundant transcript 1) in PD Neurons3 compared to Control samples. NEAT1, which encodes a long non-coding RNA, has consistently shown upregulation in PD across several reports. However, it remains unclear whether this upregulation of NEAT1 serves as a protective or detrimental mechanism in PD. A recent study has highlighted elevated NEAT1 levels in the peripheral blood cells of individuals with PD (Boros et al., 2020). Furthermore, there is compelling evidence indicating a high expression of NEAT1 in the substantia nigra of PD patients, and its potential involvement in regulating mitochondrial stability (Boros et al., 2021). While various reports propose that NEAT1 overexpression might induce both apoptosis and autophagy (Dong et al., 2021), some studies conversely suggest that increased NEAT1 expression is associated with

neuroprotection against oxidative stress and the modulation of neuroinflammation (Simchovitz et al., 2019; Pan et al., 2022). Given the conflicting findings that have emerged from multiple studies, our data may offer some much-needed clarity regarding the role of NEAT1 in PD.

### 6.3.4 UPR And Oxidative Stress in Astrocytes

Astrocytes are one of the major cell types in the brain and are believed to play a significant role in the pathogenesis of Parkinson's disease (PD). In our dataset, we have identified two subpopulations (Astrocytes 2 & 0) that are depleted and four subpopulations (Astrocytes 1, 3, 4, & 5) that are enriched in PD samples. In Astrocytes 3 and Astrocytes 4, we observed the unfolded protein response (UPR). It has been demonstrated that UPR in astrocytes can induce a reactive state that leads to non-cell autonomous neurodegeneration (Smith et al., 2020). Astrocytes 4 also exhibited an enrichment of transcripts related to oxidative stress and mitochondrial changes, in addition to UPR. Furthermore, Astrocytes 4 displayed high expression of GFAP, a canonical marker of astrogliosis. On the other hand, Astrocytes 1 expressed markers associated with reactive astrocytes, such as CD44 and CD3, but did not exhibit UPR or oxidative stress signatures. Astrocytes 5 demonstrated an enrichment of genes associated with axon development, ensheathment, and synapse organization, suggesting their involvement in neuronal maintenance and survival. This underscores the heterogeneity of astrocytes in the substantia nigra pars compacta.

Oxidative stress in astrocytes is considered to play a critical role in the pathogenesis of sporadic and familial PD (Rizor et al., 2019). Several reports have implicated mitochondria-induced oxidative stress in astrocytes in neurodegenerative diseases (Chen et al., 2020; Michael et al., 2011). In neurotoxin models of PD, such as MPTP, rotenone, and 6-OHDA models, astrocytes have been shown to produce reactive oxygen species (ROS), which may originate from impaired mitochondrial function and dysregulated astrocyte antioxidant defense machinery (Rizor et al., 2019). In contrast, the enhancement of the astrocyte antioxidant defense system by overexpressing DJ-1 appears to protect dopaminergic neurons in the substantia nigra of the rotenone-induced PD model (Miranda et al., 2018).

### 6.3.5 Microglia States in Substantia Nigra Pars Compacta

We conducted an in-depth exploration of microglia subpopulations within our dataset, examining their corresponding molecular profiles in both Control and PD samples. Our

findings revealed a notable degree of heterogeneity among microglia subpopulations. Notably, all microglia subpopulations exhibited overrepresentation in PD samples, except for TH-enriched Microglia1.

Within this diverse landscape, we identified a distinct microglia subpopulation, Microglia2, characterized by high expression of the well-known microglial marker P2RY12, along with several markers associated with microglia activation during aging (ARHGAP22 and ARHGAP15), and neuroinflammation (DOCK8) (Namekata et al., 2019; Zoller et al., 2018). However, Microglia2 did not exhibit an enrichment of transcripts associated with an immune response. This observation suggests that Microglia2 may represent an early stage of microglia activation.

On the contrary, two other microglia subpopulations, Microglia4 and Microglia0, displayed an enrichment of transcripts related to the UPR, and oxidative stress. Notably, Microglia4 exhibited a higher expression of genes associated with immune responses when compared to Microglia0. This distinction implies that Microglia0 might represent a state that lies along a continuum with Microglia4, potentially arising due to UPR and oxidative stress conditions. Furthermore, Microglia4 exhibited the highest expression levels of genes responsible for human leukocyte antigen (HLA) genes, and components of the complement cascade (a major inflammatory mechanism). It is worth noting that elevated expression of HLA-DRA in microglia has previously been reported in aged ventral tier substantia nigra tissue using the MPTP animal model of PD (Collier et al., 2017). Additionally, there have been reports suggesting that alpha-synuclein aggregates can induce the activation of the complement pathway and subsequent complement-mediated neurotoxicity (Gregersen et al., 2021). Studies have also shown that Complement activation takes place on Lewy bodies and neuromelanin containing neurons in substantia nigra of PD brains (Loeffler et al., 2006).

### 6.3.6 “Dysfunctional” Oligodendrocyte Subpopulations In PD

Like other cell types, we observed a significant level of heterogeneity within the oligodendrocyte populations. We identified three oligodendrocyte subpopulations, namely Oligos3, Oligos0, and Oligos1, that were found to be overrepresented in PD samples. Notably, Oligos3 exhibited a significant enrichment of OPALIN, a marker associated with myelinating oligodendrocytes, whereas Oligos0 and Oligos1 did not show this enrichment. Furthermore,

Oligos3 displayed a strong enrichment in processes such as morphogenesis, synapse assembly, and vesicle trafficking, which were not observed in Oligos0 and Oligos1.Oligos1 exhibits a high expression of RBFOX1, a gene encoding an mRNA splicing factor. It is worth mentioning that dysregulation of RBFOX1 has been linked to various neurodevelopmental conditions, including autism, intellectual disability, and epilepsy (Hamada et al., 2016). Furthermore, it is interesting to note that the upregulation of RBFOX1 and dysregulation of splicing have been previously reported in PD patient iPSC-derived dopaminergic neurons (Lin et al., 2016). The presence of seemingly dysfunctional oligodendrocytes in such abundance in PD samples may indeed be specific to PD and could potentially contribute to the distinctive pathology associated with this disease.

### 6.3.7 UPR And Oxidative Stress in Oligodendrocyte Subpopulation

Oligos5 were notably reduced in PD samples. These Oligos5 were characterized by their elevated expression of CRYAB, a small heat shock protein (sHSP) known to be associated with various protein aggregation-related neurodegenerative diseases, including amyotrophic lateral sclerosis (ALS), PD, and Alzheimer's disease. Additionally, Oligos5 exhibited high expression levels of genes related to iron storage (FTH1 and FTL), along with S100, a gene linked to the glial stress response. Interestingly, Oligos5 also expressed transcripts associated with UPR, oxidative stress, and apoptosis, suggesting that UPR and oxidative stress might contribute to the decline of Oligos5 in PD condition.

Previously, oligodendrocytes have been shown to be damaged in the MPTP animal model of PD (Takagi et al., 2007). Polymorphism in MOBP (Myelin-associated oligodendrocyte basic protein) is associated with PD (Soikas et al., 2021). Notably, MOBP protein is found in Lewy bodies only in PD and LBD disease but not in other neurodegenerative diseases such as Alzheimer’s disease, multiple system atrophy, and Pick’s disease (Zhang et al., 2022).

The results obtained in our study, along with some of these previous findings, confirm that oligodendrocytes are associated with PD pathology and justify further investigation.

### 6.3.8 Shared Vulnerability of *TH*-Enriched Glial Subpopulations in Substantia Nigra

A subset of astrocytes, microglia, and oligodendrocytes (Astrocytes2, Microglia1, and Oligos2) was found to be significantly depleted in PD samples. These particular glial

subpopulations display an abundance of markers associated with dopaminergic neurons, including *TH, ALDH1A1, SLC1A3, and SLC18A2*. Some of these findings were corroborated using spatial transcriptomics, which included the detection of *TH* expression in glial cells. The precise role of *TH* expression in these glial subpopulations remains unclear. However, prior research has suggested that inducing *TH* expression in astrocytes may have the potential to prevent parkinsonism in rats (Segovia et al., 1998; Tornatore et al., 1996). Additionally, reports have indicated that microglia express functional dopamine receptors (Matt et al., 2020; Mastroeni et al., 2008), and there is evidence of DAT, MOA-B, and COMT expression in astrocytes, implying their ability to uptake and metabolize dopamine (Matt et al., 2020).

Notably, both Astrocytes2 and Microglia1 express genes associated with apoptosis, oxidative stress, and the unfolded protein response (UPR), whereas Oligos2 does not display this expression pattern. This suggests that the reduction in the numbers of Astrocytes2 and Microglia1 in PD may be a result of UPR and oxidative stress. Considering that dopamine oxidation generates reactive quinones and reactive oxygen species, which can lead to cellular damage and death (Takahashi et al., 2021), it is plausible that these *TH*-enriched PD glial subpopulations are susceptible to dopamine-induced oxidative stress.

Furthermore, dopamine has been reported to function as an inhibitor of the NLRP3 inflammasome in primary human microglia (Pike et al., 2022). Consequently, the reduction in dopamine levels within the brain, arising from the loss of dopaminergic neurons and *TH*-expressing glial cells, may contribute to activated microglia-mediated neuroinflammation and subsequent neurodegeneration.

The diminished dopaminergic neurons and glial subpopulations also displayed an enrichment of *ALDH1A1*. Studies have demonstrated that *ALDH1A1*-positive SNpc dopaminergic neurons preferentially degenerate in PD (Liu et al., 2014). *ALDH1A1* is involved in the conversion of DOPAL into DOPAC, which is subsequently broken down to form HVA in dopaminergic neurons. Additionally, *ALDH1A1* has been suggested to mediate the synthesis of GABA in midbrain dopaminergic neurons (Kim et al., 2015). Although *ALDH1A1* has been reported to protect dopaminergic neurons from toxic dopamine alpha-synuclein metabolites, it remains unclear why *ALDH1A1*-positive SNpc dopaminergic neurons degenerate in PD. It is possible that a similar mechanism may underlie the depletion of TH-positive glial subpopulations.

### 6.3.9 Limitation of The Study

The study has several important limitations. Firstly, the key findings must undergo validation in an independent PD and Control cohort. Furthermore, it lacks functional validation using techniques beyond snRNA-seq and spatial transcriptomics.

Another notable limitation is the utilization of snRNA-seq, which does not capture the complete transcriptome of a cell; instead, it only captures nuclear RNA transcripts, with many existing as pre-mRNA. Generally, single-cell/single-nucleus RNA-seq technology exhibits low transcript capture efficiency, estimated at approximately 30% of the total transcriptome. Consequently, single-nucleus RNA sequencing is not expected to offer a comprehensive transcriptomic profile of the studied samples, resulting in a significant number of transcript dropouts.

Additionally, the quality of RNA extracted from post-mortem tissue is typically subpar, leading to the removal of a substantial number of nuclei during the data quality control and filtering process. In our study, out of approximately 173,000 nuclei obtained, more than half were excluded from the dataset, including a considerable number of TH-positive nuclei.

Nevertheless, we validated several findings from this study, such as TH expression in glial cells, using spatial transcriptomics. Furthermore, a meta-analysis confirmed TH expression in glial cells in other single-nuclei RNA-seq datasets of the substantia nigra/midbrain. However, TH expression in glial cells has not been previously discussed. These findings, including others, should also be validated at the protein level. Regardless of these limitations, our dataset serves as a valuable resource for future hypothesis-driven experiments.

# 7. Methods and Material

## 7.1 iPSC Culture

The KOLF2 and KUCG3 hiPSC lines were sourced from the Wellcome Sanger Institute. The KOLF2 TH RFP cell line was engineered in the Emmanouil Metzakopian laboratory at the UK Dementia Research Institute (UK DRI), with a Tyrosine Hydroxylase (TH) gene tagged with Red Fluorescence Protein (RFP) reporter, as described in Snowden et al. (2020). The SNCA iPSC lines, originally named AST23-4KO, AST23, AST23-2KO, were sourced from Dr. Martin Levesque's lab in Canada. The SNCA lines were generated by Dr. Tilo Kunath from the University of Edinburgh, as described in Chen et al. (2019). The BOB WT iPSC line was generously donated by Ludovic Vallier's lab at the University of Cambridge, and the BOB NGN2 iPSC line was a kind donation from Mark Kotter's lab, also at the University of Cambridge.

The hiPSC cell lines were cultured feeder-free on vitronectin-coated plates with TeSR E8 media. The cells were passaged as clumps using 0.5 mM EDTA solution every 3-4 days.

## 7.2 Plate Coating

Vitronectin coating was performed by diluting the vitronectin stock 1:100 in PBS, and 1 ml per well was added to the 6-well plate. The plates were then incubated at 37°C for 1 hour before use. The vitronectin-coated plates were stored in the refrigerator for up to 1 week.

PDL coating was carried out using a 25µg/ml PDL solution diluted in PBS. The plates were incubated at room temperature for 1 hour and washed three times with PBS. Subsequently, the plates were air-dried at room temperature and stored in the refrigerator for up to a year.

To perform Geltrex coating, the Geltrex stock was defrosted on ice overnight. The Geltrex stock was then diluted 1:100 in cold DMEM/F12, and 0.5 ml per well was added to the 24-well plate. The plates were incubated at 37°C for 1 hour before use. The Geltrex-coated plates were stored in the refrigerator for up to 1 week.

## 7.3 Antibiotic Selection

Transfected iPSCs were selected for antibiotic resistance using media supplemented with 1µg of puromycin, 2.5µg of blasticidin, 50µg/ml of hygromycin, or 250µg/ml of Geneticin (G418), based on the construct design.

## 7.4 iPSC Subcloning

hiPSCs were dissociated into individual cells using Accutase. The dissociated cells were counted and their viability was assessed using the Countess II. Subsequently, 2,000 viable cells were plated in StemFlex media supplemented with 10µM ROCK inhibitor on a 10 $cm^2$ vitronectin-coated dish. After approximately 10-14 days, single colonies were selected and transferred into vitronectin-coated 96-well plates for genotyping and banking purposes.

## 7.5 HEK293T culture

HEK293T cells were grown in DMEM/F12 supplemented with 10% FBS and 1% penicillin and streptomycin. The cells were passaged when they reached 75-80% confluency at a ratio of 1:10 to 1:20 for general maintenance.

## 7.6 DNA Extraction

The cells were washed with PBS and then incubated in lysis buffer (consisting of 1 mg/ml Proteinase K, 10mM Tris, 10mM KCl, 10mM $MgCl_2$, 0.45% NP40, 0.45% Tween20, and $H_2O$) at 60°C for 4 hours in a shaking incubator. The enzyme was heat-inactivated at 94°C for 10 minutes. After cooling the samples to room temperature for a few minutes, they were stored at -20°C until required for PCR.

## 7.7 PCR

PCR reactions were prepared by mixing 10µL of KAPA 2X genotyping mix, 0.5µL of forward primer (10µM), 0.5µL of reverse primer (10µM), 1µL of DNA lysate, and $H_2O$. The PCR thermocycling conditions were as follows: polymerase activation at 95°C; Stage 1: 8 cycles at 95°C for 30 seconds, 66-60°C (1°C gradient) for 30 seconds, and 72°C for 2 minutes; Stage 2: 35 cycles at 95°C for 30 seconds, 60°C for 30 seconds, and 72°C for 2 minutes; Final

extension at 72°C for 2 minutes. PCR products were run on a 1.5% agarose gel at 100V for 1 hour, and the gel image was captured using the iBright CL1000.

## 7.7 Human Midbrain Organoid Derivation- Jo et al. 2016 Protocol

Two types of basal media were used, each with specific small molecules, to derive midbrain organoids. Basal media 1 was composed of Neurobasal: DMEM/F12 (1:1) supplemented with 1X B27 without vitamin A, 1X N2, 1µg/mL Heparin, 1X Glutamax, 1X non-essential amino acid (NEAA), and 50µM β-mercaptoethanol. Basal media 2 included all the components of basal media 1, except for DMEM/F12 and heparin.

To initiate the process, hiPSCs were incubated with Accutase at 37°C for 10 minutes to dissociate them into single cells. The single cells were then counted, and viability was assessed using trypan blue on Countess II. The cells were diluted in neural induction media and plated into non-tissue culture-treated V-bottom plates at a density of 10,000 cells per well to form embryoid bodies (EBs).

The neural induction media consisted of basal media 1 supplemented with 0.8µM CHIR99021, 200ng/ml noggin, 10µM SB431542, and 10µM ROCK inhibitor (Y-27632 dihydrochloride). The media was changed the following day, replacing the neural induction media with the one lacking the ROCK inhibitor. On day 4, the cells were patterned toward the midbrain floor plate by adding 100 ng/mL FGF8a and 100ng/ml SHH-C25II to the neural induction media for 3 days. On day 7, the patterned EBs were embedded in 30 µL of reduced growth factor (GFR) Matrigel and cultured in tissue induction media. The tissue induction media consisted of basal media 2 supplemented with 100ng/mL FGF8a, 0.1µM SAG, 2.5µg/mL insulin, 200ng/mL laminin, and 1X penicillin/streptomycin. The following day, the Matrigel-embedded EBs were transferred to an ultra-low attachment 24-well plate containing neuronal maturation media and placed on an orbital shaker at 100 rpm to enhance nutrients and gasses exchang. The neuronal maturation media comprised basal media 2 supplemented with 100ng/mL GDNF, 100ng/mL BDNF, 200µM Ascorbic Acid, 125µM db-cAMP, and 1X penicillin/streptomycin. The media was changed every 3 days using the neuronal maturation media.

## 7.8 Human Midbrain Organoid Derivation- Ansari Protocol

For this protocol, the preparation of the basal media and plating of the cells into non-tissue culture treated cells were executed in the same manner as described above in the modified Jo et.al., 2016 method section. For chimeric organoids, SNCA lines were mixed with 50% KOLF2 cells. The neural induction media consisted of basal media 1 supplemented with 100nM LDN-193189, 10µM SB431542, and 10µM ROCK inhibitor (Y-27632 dihydrochloride). On day 1 and 2, the cells were grown in neural induction media consisting of basal media 1 supplemented with 100nM LDN-193189, 10µM SB431542, 100ng/mL FGF8a, 2µM purmorphamine, and 0.1µM SAG. On day 3 and 4, the cells were grown in neural induction media consisting of basal media 1 supplemented with 3µM CHIR99021, 100nM LDN-193189, 10µM SB431542, 100ng/mL FGF8a, 2µM purmorphamine, and 0.1µM SAG. On day 5 and 6, the cells were grown in neural induction media consisting of basal media 1 supplemented with 3µM CHIR99021, 100nM LDN-193189, 100ng/mL FGF8a, 2µM purmorphamine, and 0.1µM SAG. On day 7, the patterned EBs were embedded in 30 µL of reduced growth factor (GFR) Matrigel and cultured in tissue induction media. The tissue induction media consisted of basal media 2 supplemented with 3µM CHIR99021, 100nM LDN-193189, 2.5µg/mL insulin, 200ng/mL laminin, and 1X penicillin/streptomycin. Organoid development steps were performed as described in the modified Jo et.al.,2016 protocol section.

## 7.9 Oxidative Stress Treatment

To induce oxidative stress, the human midbrain organoids were grown in maturation media lacking antioxidants (B27 minus antioxidant, mercaptoethanol, and ascorbic acid) and were exposed to 0.125µM rotenone overnight.

## 7.10 L DOPA Treatment

Midbrain organoids were treated with 10uM L-DOPA for 15 days. Whole organoid and bright field images were captured using a mobile camera or an EVOS microscope at 10X magnification.

## 7.11Mitomycin C Treatment

Mitomycin C stocks were prepared by reconstituting Mitomycin C powder in sterile water. The cells were treated with 5µg/mL mitomycin C in maturation media and incubated at

37°C for 2 hours. After incubation, the media containing mitomycin C was removed, and the cells were washed three times with PBS. The media was fully replaced the next day to remove dead cells and debris.

## 7.12 L-MOSES Treatment

L-Moses was reconstituted in sterile dH2O according to the manufacturer's instructions. Midbrain organoids were treated with 12.5µM L-Moses on day 42 for seven days, and on day 49, the organoids were harvested for the downstream processing.

## 7.13 Cortical Induced Neurons Differentiation

BOB NGN2 iPSCs were dissociated into single cells using accutase and plated on PDL- and Geltrex-coated plates at a density of 25,000 cells per cm2 in E8 media supplemented with 10uM ROCK inhibitor. On day 1 and day 2, cells were grown in DMEM/F12 media supplemented with 1X N2, 1X GlutaMAX, 1X non-essential amino acids, 50µM β-mercaptoethanol, 1X penicillin and streptomycin, and 1µg/ml doxycycline. From day 3, cells were cultured in DMEM/F12 media supplemented with 1X B27 without vitamin A, 1X GlutaMAX, 1X non-essential amino acids, 50 µM β-mercaptoethanol, 1X penicillin and streptomycin, 1µg/ml doxycycline, 10µg/ml NT3, and 10 ng/ml BDNF. The media was fully replaced for the first six days, and then half fed every alternate day starting from day 7.

## 7.14 Immunohistochemistry

The midbrain organoids were fixed in 4% paraformaldehyde (PFA) at 4°C overnight, followed by three PBS washes. The fixed organoids were then incubated in 30% sucrose at 4°C overnight. To prepare the 30% sucrose solution, 30g of sucrose was reconstituted in 100ml of PBS and filtered using a 0.22µm strainer to remove contaminants. The organoids were then frozen in a plastic mold containing OCT compound on dry ice-cold isopropanol. Tissue blocks were sectioned on a cryostat to produce 12µm thick sections, which were placed on super frost glass slides for immunohistochemistry.

For the immunohistochemistry, the sections were washed once with PBS and then incubated in a 1% BSA/0.25% Triton X-100 blocking/permeabilization solution for 1 hour at room temperature. The blocking/permeabilization solution was replaced with the desired primary antibody solution prepared in a 1% BSA solution and incubated overnight at 4°C. The

sections were washed three times with PBS and then incubated in a secondary antibody solution for 1 hour at room temperature. The sections were washed four times with PBS, including a wash with PBS and DAPI. Finally, the stained sections were mounted with coverslips using a mounting medium for imaging. The images were acquired on EVOS microscope at 4X/10/20X magnification.

The following antibodies were used: TUJ1/βtubulin (1:250) (Millipore), TH (1:250) (Millipore), FOXA2(1:250) (Santa Cruz), LMX1A (1:100) (Milllipore), Ki67(1:200) (Abcam), SOX2(1:200) (Abcam), donkey anti-rabbit 488(1:1000) (Abcam), donkey anti-rabbit 647(1:1000) (Abcam), donkey anti-mouse 488(1:1000) (Abcam) and donkey anti-mouse 647(1:1000) (Abcam).

### 7.15 RNA Extraction

RNA extraction was performed using the Macherey-Nagel RNA purification kit following the manufacturer's protocol. The midbrain organoids were collected in Eppendorf tubes and washed once with PBS. The PBS-washed organoids were homogenized and lysed in lysis buffer by pipetting up and down until the tissue was dissolved. For the monolayer cultures, the cells were also washed with PBS, and lysis buffer was added to them. The lysing of the monolayer cells was performed in the same plate. The lysate was collected in an Eppendorf tube for further processing as per the manufacturer's instructions or stored at -80°C.

### 7.16 RNA Quantification

The RNA was quantified on the Qubit fluorometer according to the manufacturer's instructions (Invitrogen). The Qubit working solution was prepared by diluting the Qubit reagent 1:200 in Qubit buffer. 1 µL of the RNA sample was added to 199 µL of the Qubit working solution and vortexed for 2-3 seconds. The sample was then incubated for 2 minutes at room temperature. The sample was read on the Qubit fluorometer, and the RNA concentration was determined using standards.

### 7.17 RNA Quality

The RNA was quality checked on the TapeStation. The reagents were equilibrated at room temperature for 30 minutes and vortexed before use. Samples were prepared by mixing 1µL of the RNA sample with 5µL of RNA sample buffer. The RNA samples were denatured at

72°C for 3 minutes, followed by incubation on ice for 2 minutes. The samples were loaded into the TapeStation instrument, and the RNA quality check results were recorded in the form of RNA Integrity Number (RIN). RNA samples with a RIN number of 8 or above were used for further analysis.

## 7.18 cDNA Synthesis

The cDNA synthesis was performed using either the QuantaBio or QuantiTect Reverse Transcription kit, following the manufacturer's protocol. Up to 1µg of RNA was used for cDNA synthesis in 20µL reactions per sample. The cDNA samples were then diluted to a concentration of 5ng/µL in nuclease-free water for quantitative qPCR.

## 7.19 Quantitative qPCR

Quantitative qPCR reaction samples were prepared by mixing 2.5µL of Luna Universal probe qPCR master mix, 0.25µL of TaqMan probe, 1µL of cDNA (5 ng), and 1.25µL of nuclease-free water. The reactions were run in 5µL volumes using the comparative CT Fast thermocycling protocol on an ABI QuantStudio 5 instrument. The qPCR thermocycling conditions were as follows: stage 1 - initial denaturation at 95°C for 60 seconds (1 cycle), and stage 2 - denaturation at 95°C for 15 seconds and extension at 60°C for 30 seconds (40 cycles).

## 7.20 Protein extraction

The midbrain organoids were collected in Eppendorf tubes and washed once with PBS. The midbrain organoids were resuspended in ice-cold protein extraction lysis buffer (50 mM Tris, pH 7.5, 0.5 M NaCl, 1% NP40, 1% sodium deoxycholate (DOC), 0.1% SDS, 2 mM EDTA, COMPLETE protease inhibitors (1 tablet in 10 mL), distilled H2O) and triturated up and down using a pipette, and vortexed on and off until the tissue was completely homogenized. The homogenized lysates were kept on ice for at least 10 minutes. The samples were spun down >20,000g for 10 minutes at 4°C, and the supernatant was collected.

## 7.21 Protein Quantification

Protein quantification was performed using the Pierce BCA protein assay kit as per the manufacturer's instructions. The kit standards were diluted in the protein extraction buffer. The working solution was prepared by mixing 50 parts of reagent A and 1 part of reagent B. 10µL of protein sample was added to the 200µL of working solution in a 96-well plate. The plate was

covered in silver foil and incubated at 37°C for 30 minutes. Post-incubation, the plate was cooled down to room temperature, and absorbance was read at 562nm using XXXX. Protein concentration in samples was determined by plotting a standard curve.

## 7.22 Western Blotting

20µg protein per sample was mixed with 4X Laemmeli buffer supplemented with a reducing agent and heated at 95°C for 5-10 minutes for protein denaturation. Denatured protein samples were loaded into the wells of the 4-20% Mini-Protean TGX precast gel, and electrophoresis was performed at 100V for 1 hour. The BIO-RAD trans-Blot-Turbo transfer system was used to transfer separated proteins onto a 0.2 µm PVDF membrane. The protein transfer was performed using Bio-Rad's mixed molecular weight protocol (1.3A, 25V, 7 minutes). The membrane was then blocked in 5% milk in PBST (0.1% Tween) solution for 1 hour at room temperature. The blocking solution was removed and replaced with the primary antibody solution and incubated overnight at 4°C. The membrane was washed 3 times with TBST for 5 minutes each wash at room temperature. The membrane was incubated with the secondary antibody solution for 1 hour at room temperature. The membrane was washed 3 times with TBST for 5 minutes each wash at room temperature. The protein blots were developed using the SuperSignal™ West Pico PLUS Chemiluminescent Substrate kit as per the manufacturer's instructions and captured using iBright CL1000. For all the incubation steps, the membrane was placed on a shaker. The following antibodies were used: GAPDH (1:5000), α-synuclein (1:2500), TH (1:2500), LMX1A (1:1000), goat anti-mouse HRP (1:5000), and goat anti-rabbit (1:5000). The densitometry of the blots was performed on ImageJ. The ratio of arbitrary values of the protein of interest and GAPDH was plotted for visualization.

## 7.23 Sample Preparation For scRNA-Seq

To dissociate the midbrain organoids into single cells, the organoids were first washed with PBS and then incubated at 37°C in 0.25% trypsin for 15 minutes. Following the incubation, the organoids were pipetted up and down to break up the tissue. Accutase was added on top, followed by an additional 15 minutes incubation at 37°C. The clumps were further broken down by pipetting up and down. If clumps were still present after the second round of incubation, cells were incubated at 37°C for another 10-15 minutes. Trypsin/Accutase was neutralized with maturation media, and the cell suspension was collected in a Falcon tube. The cells were washed three times with PBS supplemented with 0.04% BSA to remove debris/dead

cells, counted, and checked for viability on the countess 3FL. For each sample, 10,000 live cells in 43µL PBS supplemented with 0.04% BSA solution were prepared for the scRNA-seq procedure. The samples were loaded onto the Chromium controller (10X Genomics) for single-cell partitioning and barcoding. The single-cell cDNA synthesis, amplification, and sequencing libraries were produced as per the manufacturer's protocol (10X Genomics 3' RNA sequencing). The libraries were sequenced on an Illumina HiSeq 4000 using the following parameters: Read 1: 26 cycles, i7: 8 cycles, i5: 0 cycles; Read 2: 98 cycles to produce 75 bp paired-end reads.

### 7.24 SNCA Triplication Screening

SNCA triplication genotyping was performed as described in Devine et al., 2011. Genomic DNA extraction was carried out as previously described. The dosage of SNCA was then determined using SYBR Green qPCR with primers specific to SNCA exon 1 and exon 2, and β-globin (HBB) as the diploid control gene on the ABI QuantStudio 5 instrument. The qPCR thermocycling conditions were as follows: stage 1 - initial denaturation at 95°C for 60 seconds (1 cycle), and stage 2 - denaturation at 95°C for 15 seconds and extension at 60°C for 30 seconds (40 cycles); Melt curve: 95°C for 1 second, 60°C for 20 seconds, 95°C for 1 second. The primers used for this screening are listed below:

| | |
|---|---|
| SNCA exon 1_F | AAAGGCCAAGGAGGGAGTT |
| SNCA exon 1_R | ATCCTAACCCATCACTCATGAAC |
| SNCA exon 4_F | CCTGTGGATCCTGACAATGA |
| SNCA exon 4_R | TGCAAGTTGTCCACGTAATGA |
| HBB_F | TTGGACCCAGAGGTTCTTTG |
| HBB_R | GAGCCAGGCCATCACTAAAG |

## 7.25 Plasmids Designs

The forced expression PiggyBac plasmids were designed and acquired through VectorBuilder. For the standard PiggyBac plasmids, EF1alpha promoter was used to drive the expression of gene of interest. The standard PiggyBac plasmid had 1 transcription factor (TF), blasticidin resistance gene for selection and Blue Fluorescence Protein (BFP) as a reporter. For inducible PiggyBac plasmids, the activator(rtTA) expression was driven by EF1alpha promoter and TFs were driven by the responder (TRE) promoter.

## 7.26 Plasmid Amplification

The plasmids were received in bacteria as glycerol-frozen stocks. For each plasmid, a bacterial culture was streaked on ampicillin agar 10cm² dish and incubated at 37°C overnight. A bacterial colony was picked and cultured in 5ml of LB media supplemented with 100µg/ml ampicillin overnight at 37°C. The next day, the culture was further expanded to a 500ml culture at a 1:500 dilution using LB media with ampicillin and incubated overnight at 37°C. The culture was collected and spun down at 6000g for 15 minutes to pellet the cells. The cell pellets were stored at -20°C or carried forward for the plasmid extraction process.

## 7.27 Bacterial Transformations

Frozen NEB10 bacterial culture was defrosted on ice. 1uL of plasmid DNA (100ng) was mixed with 25uL of NEB10 bacterial culture and gently mixed by flicking the bottom of the tube. The mixture was then incubated on ice for 20-30 minutes. Heat shock was performed by placing the transformation tube in a water bath at 42°C for 30 seconds. The tube was then placed back on ice for 5 minutes. Next, 250uL of NEB bacterial outgrowth media was added to the tube and incubated at 37°C on a shaker for 45 minutes. Afterward, 25uL of the culture was spread onto an ampicillin agar plate and incubated overnight at 37°C. The following day, a colony was picked and expanded as described above.

## 7.28 Plasmid Extraction

plasmid extraction was performed using Qiagen Maxi kit according to the manufacturer's protocol. The quantification of the extracted plasmids was done using Qubit (Invitrogen).

## 7.29 Induced Dopaminergic Neurons Differentiation

For the dopamine neuron differentiation using standard PiggyBac plasmids, the dissociated hiPSCs were seeded into Geltrex (Thermo Fisher) coated 24 well plate (Corning) at 50,000 cells per well in StemFlex media (Thermo Fisher) with 10µM ROCK inhibitor. Next day, transfections were performed using Lipofectamine stem reagent (Thermo Fisher) as per manufacturer's protocol. A total of 500ng of forced expression plasmid DNA and 50ng of transposase plasmid DNA were transfected per well. On day 1, the media was replaced with fresh StemFlex media. On day 2 & 3, the transfected cells were cultured with maturation media which was made up of basal media 2 supplemented with 1X pen/strep, 100ng/mL BDNF, 100 ng/mL GDNF, 125µM db-cAMP and 200µM Ascorbic Acid. On day 4 & 5, the cells were selected with 2µg/ml blasticidin (Thermo Fisher) in maturation media. The cells were media changed with maturation media every day until collected for immunocytochemistry. On day 7, the cells were treated with 5ug/ml mitomycin C at $37^0$C for 2 hours to remove proliferative cells.

The hiPSCs were differentiated into dopamine neurons using inducible PiggyBac plasmids with slight changes to the above transfection protocol. On day 1, the transfected cells were treated with 1µg/ml doxycycline in StemFlex media. The doxycycline treatment was kept on from day 1 until the end of the experiment.

## 7.30 Induced Dopaminergic Neurons Differentiation- Sagal et al. 2014 protocol

ATOH1 subcloned iPSCs were plated at a density of $8 \times 10^4$ cells per $cm^2$ on Geltrex-coated plates, and they were cultured in E8 medium supplemented with the ROCK inhibitor (Y-27632, 10µM). The induction of Atoh1 was initiated by exposing the cells to doxycycline (0.5 µg/ml) from day 1 through day 5. From day 1 to day 3, medium was changed daily, transitioning gradually from Essential 6 medium to DMEM/F-12 medium supplemented with N2 (N2 medium). The cells were maintained in N2 medium until day 7. At this point, cells were dissociated using Accutase and then replated at a density of $3 \times 10^5$ cells per $cm^2$ on dishes pre-coated with poly-d-lysine (1µg/ml) and laminin (1µg/ml). Neurobasal medium supplemented with B27, brain-derived neurotrophic factor (20ng/ml), glial cell line-derived neurotrophic factor (20 ng/ml), transforming growth factor type β3 (1ng/ml), ascorbic acid (0.2 mM), dibutyryl cAMP (0.5 mM), and the γ-secretase inhibitor DAPT (10µM) were employed for cell

culture. During days 8 to 36, half of the cell culture medium was replenished every 3 to 4 days. For the Atoh1+ SHH_FGF8b condition, in addition to doxycycline (0.5µg/ml), the cells were treated with SHH C25II (100 ng/ml) and FGF8b (100 ng/ml) from day 1 to day 5.

### 7.31 Induced Dopaminergic Neurons scRNA-seq sample preparation

The culture was washed with PBS and then incubated with Accutase at 37°C for 30-45 minutes until the culture consisted mostly of single cells. The cell suspension was pipetted up and down to dissociate the clumps every 15 minutes using a 1ml pipette. The culture was washed with PBS three times to remove dead cells and debris. The cells were viability checked using trypan blue and counted on Countess II. Finally, 10,000 live single cells were resuspended in a 0.04% BSA solution and proceeded for scRNA-seq as described above.

### 7.32 Induced Dopaminergic Neurons Immunocytochemistry

The cells were fixed in 4% paraformaldehyde at room temperature for 15mins. The fixed cells were blocked in 5% donkey serum or 1%BSA-0.1% Triton X-100 solution for 1 hour at room temperature. The primary, secondary and remaining steps were performed as described in 'Midbrain organoids immunohistochemistry' section.

### 7.34 Cloning

Gradient PCRs were performed using Q5 high-fidelity 2X master mix to obtain desired regions from the donor plasmids. The thermocycling conditions were the same as described in the 'PCR' section. The amplicons were run on a 1.5% agarose gel to select clean PCR products with no non-specific bands. The selected PCR products were purified using the AMPure XP PCR clean-up kit.

In parallel, restriction digestion was performed on the recipient plasmid by mixing the desired enzymes, enzyme-specific buffer, 1µg of plasmid DNA, and water for at least 1 hour at 37°C. The mixture was then run on a 1.5% agarose gel to confirm digestion. The digested products were purified using the AMPure XP PCR clean-up kit.

Both the donor fragment and the linearized/digested recipient plasmid DNA were quantified using a nanodrop instrument.

Gibson assembly was performed in a 20µL reaction by mixing 10µL of 2X Gibson assembly master mix, $H_2O$, and up to 200ng of purified donor fragment and digested recipient plasmid. The assembled reaction was transformed into NEB10-beta competent E. coli, as

described in the 'Bacterial transformation' section. The bacterial colonies were picked and screened for positive clones using PCR to confirm the expected assembled region.
The positive colonies were expanded as described in the 'Plasmid amplification' section and verified by Sanger sequencing. The efficiency of the plasmid was tested by transfecting HEK293T cells using Lipofectamine LTX with PLUS reagent and checking the expression of the cloned gene by immunocytochemistry or qPCR.
**Note:** Sandeep Rajan helped with this procedure.

## 7.35 Human Midbrain Organoids Single Cell RNA sequencing Data Analysis

### 7.35.1 Processing and Filtering of Single Cell RNA-seq

Cellranger v3.1.0 (G. X. Y. Zheng et al. 2017) was used to process the raw FastQ files. Doublets were detected using souporcell or scrublet(Wolock, Lopez, and Klein 2019). The doublets and cells with less than 1500 genes were excluded from the analysis.
**Note:** This step was performed by Nikos Patikas.

### 7.35.2 Dimensionality Reduction

The scVI module from scvi-tools (Xu et al.2021) and scanpy (Wolf, Angerer, and Theis 2018) were used to perform dimensionality reduction.
**Note:** This step was performed by Nikos Patikas.

### 7.35.3 Cell Type Annotation

Cell type annotation were generated using healthy KOLF2, SNCA-Corrected, SNCA-Triplication organoids cultures. The KOLF2 iPSC cell line was derived from a healthy male donor, whereas, SNCA triplication and its isogenic control originated from a female PD patient from the lowa kindred (Devine et al. 2011). The clustering was perform using the scVI latent space and scanpy's louvain function at resolution of 0.5. The clusters were annotated using a panel of known marker genes associated with neural stem cells, and DA neurons specification, development, and maturity. Clusters 0, 1, 4, and 12 were merged due to the similarities in dopaminergic neuronal marker gene expression profiles.
**Note:** This step was performed by Nikos Patikas.

scvi differential_expression() was used to obtain top 200 cell marker gene list for each cell type. EPI cell type was annotated by analysing top 25 marker genes of this population in

scfind using the mouse brain atlas (J. T. H. Lee et al. 2021; Zeisel et al. 2018). The same clustering and cell type annotation were used for the rotenone and chimera samples. The label transfer workflows from the scANVI model (Xu et al. 2021) was used to process the rotenone and chimera samples.

**Note:** This step was performed by Nikos Patikas.

### 7.35.4 2D And *In Vivo* Single-Cell Gene Expression Integration

The data sets integration was performed using scvi model's dimensionality reduction workflow. The cell type label prediction was done using scANVI map our midbrain organoid neuronal cell types to *in vivo* neuronal nuclei and vice versa.

**Note:** This step was performed by Nikos Patikas.

### 7.35.5 Differential Expression and Gene-Set Enrichment Analysis

For the gene set enrichment analysis (GSEA), top 200 differentially expressed genes (DEGs) were used. The DEGs were sorted based on either by Bayes Factors or by p-adj for switchde output. The Bayes factors values were log normalized and had a substantial significance threshold (>3.2) (Kass and Raftery 1995). The pathway analysis on these genes were performed using metascape (Y. Zhou et al. 2019). The most interestingly enriched pathways were shown in the figures.

scanpy's rank_genes_groups using the wilcoxon test were used for chimeras biological analysis. The top 200 genes were selected based on adjusted p-value< 0.05. For the non-cell autonomous effects, we used the DC1 dimension and correlated with all the normalized gene expression.

**Note:** This step was performed by Nikos Patikas.

### 7.35.6 Rotenone Treatment Analysis

MASC (Xu et al. 2021) was used to assess the effect of rotenone on cell type numbers. Pseudotime scores of rotenone treated and untreated samples were generated using an in-silico rescue method which relied on a Bayes Factor for each DAn1 cell's c latentrepresentation L(c). The significant determinants along the pseudotime trajectory were identified using the switchde package (Campbell 2022). The activation coefficient K were sorted according to their p-value. For pseudo-time heatmaps, z-scores of the scaled expression were plotted.

**Note:** This step was performed by Nikos Patikas.

### 7.35.7 Chimera Genotype Calling

To assign identities in our chimera samples, a modified workflow of souporcell (Heaton et al. 2020) were used as described(https://github.com/wheaton5/souporcell), was adapted to include some extra controls. A genetic variant reference for KOLF2 and SNCA cells were generated for analysis. The accuracy of the algorithm was validated assigning non-chimera cells to the correct donor.

**Note:** This step was performed by Nikos Patikas.

### 7.35.8 Visualization

Organoid characterization gene expression dot plots were inspired by Fiorenzano et al. (Fiorenzano et al. 2021) and used scaled expression data. Organoids integration with *in vitro* and *in vitro* datasets gene expression dot plots used normalized log-expression.

**Note:** This step was performed by Nikos Patikas.

### 7.35.9 Wang et al. SNpc Neuronal Nuclei Data Set

The *in vivo* substantia nigra single nuclei dataset was sourced from Wang et.al., 2022. The datasets of only healthy donors(n=9) were used in our comparative analysis. The clustering was performed same as organoid cultures using scvi workflow except donors as covariates to control for inter donor variability. Louvain () based clustering highlighted high neuronal marker SYT1 expressing cluster which was used for the integration with the organoid datasets.

**Note:** This step was performed by Nikos Patikas.

### 7.35.10 Fiorenzano et al. Ventral Midbrain Organoids

The Fiorenzano et al.2021 day 60 organoids scRNA-seq data set was sourced, and dimensionality reduction performed same as our organoids cultures using scvi workflow. However, the cell type annotations were kept same as Fiorenzano et al.2021 for the comparisons.

**Note:** This step was performed by Nikos Patikas.

## 7.36 Post-Mortem Substantia Nigra Single Nucleus RNA Sequencing Data Analysis

### 7.36.1 Selection of Donors

For this study, fresh frozen post-mortem brain tissue of 29 donors (15 sporadic PD and 14 Controls) from the Oregon brain bank (human tissue repository for brain research), were selected. The donors were of both sexes (females and male), and predominantly over 60 years of age. The detailed clinical reports were received along with the tissue samples. These reports provided information on the sex, age, post-mortem interval (PMI) before the tissue was collected, donor's clinical diagnosis, AD-related changes in cortex (presence of neuritic plaques and neurofibrillary tangles), presence of Lewy bodies in the midbrain, neocortical (frontal cortex), limbic (amygdala) regions. The Ethics Committee of UZ Leuven/KU Leuven approved the study under the S-number S64182.

### 7.36.2 Isolation Of Nuclei From Frozen Post-Mortem Brain Tissue And Single Nuclei RNA-Seq

First, brain tissue samples were cut at 5uM thickness using a cryostat, and haematoxylin and eosin (H&E) staining performed to identify substantia pars compacta (SNpc) region in PD and control tissue. The SNpc tissue were taken out as 2mm punch biopsy, and nuclei isolation was performed. The single-cell cDNA synthesis, amplification, and sequencing libraries were produced as per the manufacturer's protocol (10X Genomics; Single Cell 3 Reagent Kits v.3). The libraries were sequenced on an NovaSeq6000 system using the following parameters: 1% PhiX, paired-end sequencing with 10X v3 parameters (28-8-0-91 cycles).

**Note:** This step was performed by Dietmar Thal and colleagues.

### 7.36.3 Pre-Processing Single Nuclei RNA-Seq Data

Gene counts were derived through the alignment of reads to the human hg38 genome (GRCh38/Ensemble 93 pre-built by 10x) utilizing CellRanger software (v.3.0.2) from 10x Genomics. To accommodate unspliced nuclear transcripts, reads that mapped to pre-mRNA were tabulated, following the recommendation of CellRanger v.3.0.2. The study adhered to the default CellRanger parameters throughout the process.

**Note:** This step was performed by Araks Martirosyan.

### 7.36.4 Quality Control

The initial dataset encompassed 173,197 nuclei, each providing insight into the transcriptome profile of 30,194 distinct transcripts. During the initial filtration phase, certain

criteria were established to eliminate nuclei with insufficient RNA content (< 500 transcripts), as well as to exclude doublets characterized by excessive RNA content (> 6000 transcripts). Additionally, MALAT1 was deliberately removed from the list of transcripts due to its pronounced enrichment, which had the potential to influence clustering.

Subsequently, to ensure that only singular nuclei remained in the dataset, the Scrublet algorithm was applied to the dataset of each donor, employing default parameters to further identify and remove doublets. In total, 2,496 doublets were successfully identified and subsequently eliminated from the dataset.

Finally, attention was directed towards libraries displaying a relatively high content of mitochondrial RNA. A preliminary round of clustering was executed using the Seurat v3.1.1 R package, with mitochondrial gene detection capped at 1%, 2%, and 5% for each library. This analysis revealed a distinct inclination to group cells together based on elevated levels of mitochondrial transcripts, which was particularly evident with the 5% threshold. This effect became less pronounced with the 1% and 2% thresholds. To retain a maximal number of cells while circumventing the bias introduced by mitochondrial transcripts driving clustering, libraries where 2% or more of the detected transcripts were potentially of mitochondrial origin were excluded. This step resulted in a final dataset of 83,484 nuclei expressing 30,194 unique transcripts, ready for subsequent analysis.
**Note:** This step was performed by Araks Martirosyan.

### 7.36.5 Integration of The Datasets And Clustering

Data normalization and clustering procedures were executed using the Seurat package. Initially, each dataset underwent SCT normalization via the SCTransform() function. Subsequently, the top 6000 most variable features were selected for downstream integration through employment of the SelectIntegrationFeatures() function. Principal Component Analysis (PCA) was independently conducted on each dataset using the RunPCA() function, considering the top 50 principal components.

Integration anchors were identified by utilizing the FindIntegrationAnchors() method, with sample s.0096 serving as the reference. The integration of datasets was accomplished by employing the IntegrateData() function across the top 20 dimensions. Further PCA analysis

was performed on the integrated dataset, and the resultant top 10 principal components (PC) were employed for UMAP analysis, facilitated by the RunUMAP() function.

Finally, the leading 10 principal components (PC) were harnessed to construct a k-nearest-neighbors graph, achieved through the FindNeighbors() function, employing a resolution of 0.6 to delineate nuclei into clusters. Subsets displaying positivity for recognized markers specific to distinct cell types were grouped together, thereby defining overarching cell categories. Reclustering of neuronal and glial sub-populations was conducted based on the top 4 principal components (PC) using SCT normalized data, employing a resolution of 0.2.
**Note:** This step was performed by Araks Martirosyan.

### 7.36.6 Marker Identification

Marker genes were identified for each subcluster via a differential gene expression analysis between nuclei within the specific subcluster against all remaining nuclei in the population. This analysis employed a Wilcoxon rank-sum test, requiring an FDR-corrected p-value of $\leq 0.05$, in conjunction with a log2 ratio (mean gene expression across cells in the sub-cluster / mean gene expression across cells in other sub-clusters) of at least 0.25. Furthermore, transcripts present in a minimum of 25% of the nuclei within a given subcluster were taken into consideration. **Note:** This analysis was performed by Araks Martirosyan.

### 7.36.7 Datasets of PD GWAS Associated Genes

The list of PD GWAS associated of genes was sourced from Fernandes et al., 2020, along with the GWAS EMBL catalogue (release version 23.11.21). **Note:** This step was performed by Araks Martirosyan and Rizwan Ansari.

### 7.36.8 Differential Gene-Expression Analysis

The evaluation of differential expression in snRNAseq data was conducted employing a linear mixed model that incorporated the origin individual as a covariate. This analysis was executed utilizing the R packages. Subsequently, a one-way ANOVA test accompanied by Benjamini-Hochberg correction was conducted to ascertain the statistical significance of the identified up or downregulation of genes. This statistical procedure was carried out using the R package car. Note: This analysis was performed by Araks Martirosyan.

### 7.36.9 Pathway Analysis

Gene-enrichment and functional annotation analyses were conducted on subpopulation-specific genes as well as genes that were either upregulated or downregulated in PD condition. These analyses were performed using clusterProfiler version 4.4.4 or Metascape.org. Note: Pathway analysis using clusterProfiler was performed by Araks Martirosyan.

### 7.36.10 Spatial Transcriptomics - Sample Preparation

Tissue sections from the Substantia Nigra pars compacta (SNpc) were prepared from the brains of three Control donors and three PD donors. The initial step involved placing the samples onto a sectioning block, which were then embedded using FSC 22 frozen section media. The Leica CM3050 cryostat was utilized for the subsequent sectioning process. During this phase, the chamber temperature was maintained at -18°C, while the object temperature was held at -16°C. Samples were expertly sectioned to a thickness of 10 µm and delicately placed onto pre-chilled coverslips. Post-sectioning, the samples were carefully preserved at a temperature of -80°C, eventually being transported to Resolve GmbH on dry ice for further processing.

**Note:** This step was performed by Francisco Pestana.

### 7.36.11 Spatial Transcriptomics - Molecular Cartography

Spatial transcriptomics was conducted using the Molecular Cartography™ platform provided by Resolve Biosciences. Tissue sections were subjected to hybridization with customized probes targeting 96 distinct transcripts. After hybridization, a thorough washing procedure was employed to eliminate excess probes. A two-step colour development protocol was then applied to label the samples fluorescently. During the decolorization step, the fluorescent signal was removed after imaging. This color development, imaging, and decolorization cycle was iteratively performed to establish a unique combinatorial code for each target transcript. Images featuring DAPI staining and associated coordinates for every transcript within the image were furnished. The coordinate files served as input for object segmentation utilizing the Baysor tool.

**Note:** This step was performed by Francisco Pestana.

### 7.36.12 Spatial Transcriptomics – Object Segmentation

Segmenting individual molecules into distinct objects involved a two-step process using Baysor version 0.5.2. Initially, molecular data (x, y, z coordinates) was segmented using the following parameters: cell radius (45 pixels), cell radius standard deviation (25%), and a minimum molecule count per cell (15 molecules). To address potential under-segmentation of large objects, the process was repeated, employing a radius of 90 pixels. In this second step, only molecules flagged as noise during the initial segmentation and objects surpassing 80% of the calculated area (0.8 * π * 45px^2) were considered as input. Proper segmentation was confirmed if the number of resultant objects was fewer than the input large objects; otherwise, the large objects from the initial segmentation were deemed properly segmented. Successfully segmented objects were combined to generate individual objects containing transcript counts, in addition to metadata about single molecules and cells.
Note: This step was performed by Francisco Pestana.

### 7.36.13 Spatial Transcriptomics - Quality Control

After eliminating objects with undefined area or infinite density, a total of 75,416 cells remained. Objects were retained based on criteria such as having over 10 molecules, expressing more than 2 unique genes, a molecule-to-gene ratio higher than 1, elongation below 10, density below 0.04, average confidence above 0.7, and an area larger than 500 pixels. These thresholds were selected through visual examination of histograms for each variable's distribution. No maximum area threshold was set due to potential variation in cell sizes. Consequently, 41,198 segmented objects were obtained. An initial quality check for clustering was performed by subjecting the filtered objects to clustering utilizing Seurat v4.3.0, followed by Harmony batch effect correction. Specific parameter values were utilized as outlined.
**Note:** This step was performed by Francisco Pestana.

### 7.36.14 Spatial Transcriptomics - Cell Type Classification and Clustering

Cell type identification employed Seurat's unimodal UMAP projection workflow, facilitating projection of the spatial dataset onto a reference UMAP structure derived from the snRNA-seq dataset described in the manuscript. For the reference dataset, downsampling to 26,907 cells was conducted, sampling 5,000 cells from each cell type except for 1,560 connective tissue cells and 347 T-cells. Anchors connecting the query and reference datasets were identified using specific functions and reductions. After mapping, a secondary filtering

step was executed, removing cells not belonging to the majority population within each spatial cluster, yielding 36,806 high-quality cells. Final clustering and batch effect correction followed the procedures elucidated in the quality control section.

**Note:** This step was performed by Francisco Pestana.