STAT 530 Final Project

Ehsan Homaee, Junseok Yang, Siyun Shen and Sam Belsky

May 2024

1 Introduction

Dopamine is a neurotransmitter primarily produced in the brain and kidneys. In the brain, this chemical serves a wide array of functions, spanning basic motor control to executive function. Dopamine is classified as a catecholamine, which are a class of chemicals which are composed of a catchetol, or *o*-dihydroxybenzene, ring with an amine side chain. [1]

To achieve this chemical structure, dopamine biosynthesis uses the amino acid tyrosine as a dopamine precursor. The residue's phenol ring is first hydroxylated at C3' position, a reaction which is catalyzed by the enzyme tyrosine hydroxylase (Th). The product of this transformation, L-3,4-dihydroxyphenylalanine (L-DOPA), then undergoes a decarboxylation reaction catalyzed by L-DOPA decarboxylase (DDC) or aromatic amino acid decarboxylase (AAAD) (See Fig. 1). [1]

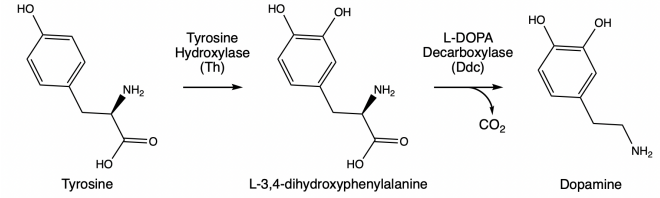

Figure 1: **Biosynthesis of dopamine from tyrosine.** Tyrosine is first converted to L-DOPA by the enzyme tyrosine hydroxylase (TH), and subsequently, L-DOPA is decarboxylated to dopamine by L-DOPA decarboxylase (Ddc).

Dopamine biosynthesis has been primarily viewed as a intracellular process, although in a recent paper, *Molecular, Spatial, and Functional Single-Cell Profiling of the Hypothalamic Preoptic Region*, it was demonstrated that of 6 Th-enriched neuronal clusters the study identified, only two were noted to also demonstrate enriched expression of Ddc, suggesting excessive production of L-DOPA in some of these cells. [2] The paper focussed on the hypothalamic preoptic (POA) region of the brain, which plays roles in mating, modulating sleep, and mediating body temperature, using dopamine in all of these processes. [3–5]. It has since been proposed that in order for dopamine biosynthesis to be completed, Th-enriched, Ddc-underexpressing (Th+/Ddc-) cells exocytose L-DOPA to Ddc-enriched, Th-underexpressing (Th-/Ddc+) cells.

Despite the validity of this hypothesis, this is not the only possible explanation for this observation. Although L-DOPA's primary neurological role is acting as a precursor to dopamine, it has been previously shown to play a few of other physiological functions. For instance, unlike dopamine, L-DOPA can pass through the blood-brain barrier through which it can reach the peripheral nervous system and be converted into dopamine. In addition to facilitating communication between the peripheral and central nervous systems, L-DOPA has been shown to act as an antioxidant and help protect brain tissue from damage by free radicals, is known to be and effective anti-inflammatory, and has inhibitory activity towards acetylcholinesterase. [6, 7] To this end, in this study, we aimed to further understand the molecular profiles of the distinct classes of dopaminergic cells (Th+/Ddc+, Th+/Ddc-, Th-/Ddc+, Th-/Ddc-), and determine if functional-spatial analysis of the two classes of partial dopaminergic cells supports the hypothesis that Th+/Ddc- cells exocytose L-DOPA to Th-/Ddc+ cells in order for dopamine biosynthesis to be completed, or if some L-DOPA production in Th+/Ddc- cells may serve to support the metabolite's understudied functions as an antioxidant, anti-inflammatory, enzyme inhibitor, and signalling molecule.

2 Results

scRNA-seq analysis of the preoptic region

In our study, quality control measures were applied to the single-cell RNA sequencing (scRNA-seq) data [8]. We refined our dataset by applying filters based on the number of detected genes (between 1000 and 5000 per cell) and total RNA counts (capped at 30,000), effectively excluding potential outliers such as doublets or damaged cells. This filtering ensured that our analyses would only include viable, single cells, thus enhancing the accuracy of our findings. Subsequent to these steps, data clustering was performed, and the results were visualized in two dimensions using Uniform Manifold Approximation and Projection (UMAP) techniques [9]. To examine the expression patterns of the primary genes under study, namely Ddc and Th, their distribution was analyzed across various cellular clusters (refer to Fig. 2). It was observed that neither Ddc nor Th is exclusive to a single cluster; rather, their expression spans multiple cellular clusters, as depicted in Fig. 2.

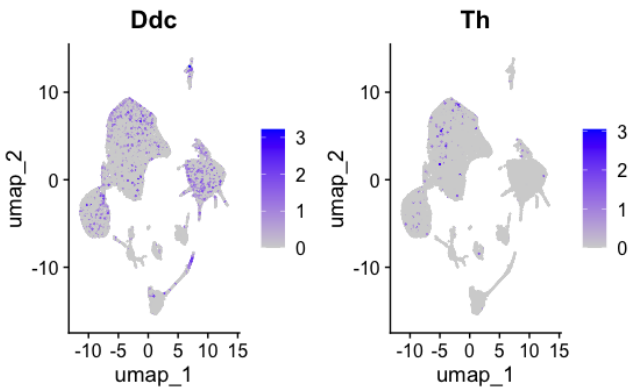

Figure 2: **Comparative Distribution of Ddc and Th Gene Expression in Neuronal Populations.** The left panel (Ddc) and right panel (Th) depict the spatial gene expression patterns visualized via Uniform Manifold Approximation and Projection (UMAP) analysis. Each dot represents an individual cell, with the color intensity reflecting the expression level of Ddc (left) and Th (right) genes, respectively, as indicated by the color scale from 0 (no expression) to 3 (high expression). The plots demonstrate that both genes are expressed in different clusters.

Subsequently, cell categorization was conducted utilizing Uniform Manifold Approximation and Projection (UMAP) based on gene expression profiles, delineating cells into four principal clusters: Ddc- Th-, Ddc+ Th+, Ddc+ Th-, and Ddc- Th+. This classification aims to evaluate whether cells sharing similar expression patterns spatially congregate within the same regions of the constructed UMAP space.

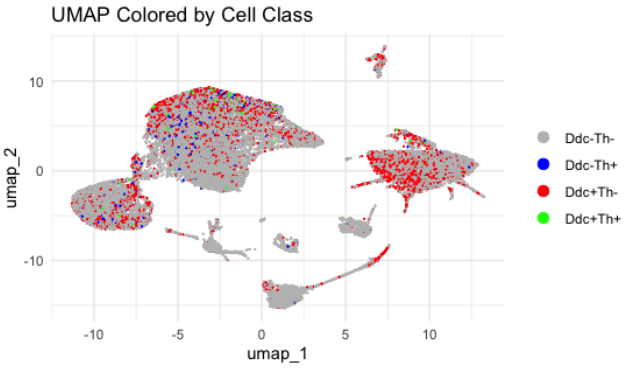

Figure 3: **MAP Visualization of Cell Class Distribution Based on Ddc and Th Gene Expression.** This UMAP plot showcases the spatial distribution of four distinct cell classes colored by their gene expression profiles: Ddc-Th- (blue), Ddc+Th+ (green), Ddc+Th- (red), and Ddc-Th+ (gray). The diverse clustering patterns observed across different areas emphasize the heterogeneity of gene expression within these populations.

As illustrated in Fig. 3, a significant majority of the cells fall within the Ddc- Th- category, while a notably smaller fraction is classified as Ddc+ Th+. The remaining cells, which are categorized as either Ddc+ Th- or Ddc- Th+, exhibit a more dispersed distribution across the UMAP space. This dispersion suggests that while Ddc and Th expression plays a critical role in cellular categorization, additional factors may influence cluster assignment within the UMAP representation, indicating the complexity of the gene expression landscape in defining cellular phenotypes.

MERFISH measurements of the preoptic region

Analysis of the single-cell RNA sequencing (scRNA-seq) data provided evidence for the presence of partial dopaminergic neurons. The subsequent phase of this research involves mapping these cells within the Multiplexed Error-Robust Fluorescence In Situ Hybridization (MERFISH) dataset, which includes spatial information. The proximity of Th+;Ddc- cells to Th-;Ddc+ cells could support the hypothesis that the former group exocytoses L-DOPA to the latter to facilitate dopamine synthesis.

However, a challenge is posed by the absence of Ddc gene expression data in the MERFISH dataset, which only contains Th expression data. Additionally, the MERFISH dataset presented issues with missing values in the gene expression profiles. To address this, we implemented median imputation to fill these gaps, ensuring a more robust dataset for analysis. Following this preprocessing step, we employed a machine learning framework to infer the Ddc expression status using the available gene expression profiles. We developed a classifier trained on the scRNA-seq data, designed to predict the Ddc status by leveraging 156 genes shared between the scRNA-seq and MERFISH datasets. This approach enables an indirect, yet robust, assessment of Ddc expression within the spatially-resolved MERFISH dataset.

To ensure compatibility between the datasets, gene expression values were normalized within a range of 0 to 1. For the training of the classifier, we partitioned the scRNA-seq data into 80% for training and 20% for validation to assess the model's performance on unseen data. Another issue addressed was the class imbalance in the training data, with Ddc- samples outnumbering Ddc+ by nearly sixfold. This required careful consideration to maintain classifier accuracy. We evaluated multiple algorithms, including Multi-Layer Perceptrons (MLPs) [10], Support Vector Machines (SVM) [11], Decision Trees [12], XGBoost [13], Logistic Regression [14], and a composite Voting Classifier [15] utilizing both hard and soft voting strategies. Despite considering regression models to exploit potential gene regulatory networks influencing Ddc expression, classification approaches proved more effective.

Logistic Regression was ultimately selected for its superior performance, achieving an approximate accuracy of 70% on the validation set. Applying this classifier to the MERFISH dataset allowed for the determination of Ddc expression statuses, revealing counts of 559,627 Ddc+ cells and 468,221 Ddc- cells. Integration of Th expression data from the same MERFISH dataset enabled comprehensive profiling of cell populations, yielding

counts of 336,478 Ddc+ Th-, 284,668 Ddc- Th-, 223,149 Ddc+ Th+, and 183,553 Ddc- Th+ cells. This integrated approach provides a detailed mapping of cellular interactions essential for understanding dopamine synthesis mechanisms within the brain's spatial architecture.

Now equipped with the Ddc and Th status of cells, we utilized the MERFISH dataset's spatial data to visualize the distribution of cells and examine their proximity within actual tissue sections. To validate the hypothesis that TH+; Ddc- cells exocytose L-DOPA to TH-; Ddc+ cells for dopamine synthesis, we analyzed their spatial distribution using the provided coordinates. This analysis aimed to determine if these cells are closely positioned, supporting the proposed intercellular interaction among partial dopaminergic neurons. Figure 4 displays only two slices from a total of 182 distinct slices, clearly showing that the two types of partial dopaminergic cells are quite well mixed with each other. This spatial analysis is crucial for understanding the functional architecture of dopaminergic signaling within the tissue.

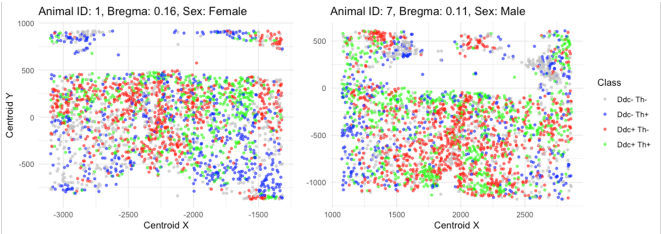

Figure 4: **Spatial Distribution of Neuronal Cell Classes Based on Ddc and Th Gene Expression in Different Sexes and Bregma Locations.** This visualization presents the distribution of four cell classes across two animals: Animal ID 1 (Female, Bregma 0.16) and Animal ID 7 (Male, Bregma 0.11). Each cell is color-coded by class—Ddc- Th- (gray), Ddc- Th+ (blue), Ddc+ Th- (red), Ddc+ Th+ (green). The widespread distribution of these classes in various brain regions supports the hypothesis that TH+;Ddc- cells exocytose L-DOPA to nearby TH-;Ddc+ cells, facilitating dopamine synthesis across different neuronal clusters as proposed. This pattern underlines potential mechanistic interactions and supports the existence of a spatially mediated dopaminergic synthesis pathway in the brain.

Due to the large number of tissue slices, examining each one individually was impractical. To systematically assess the degree of mixing between Ddc+ Th- and Ddc- Th+ cells across all slices, we implemented a quantitative approach by measuring the distance from each Ddc- Th+ cell to its nearest Ddc+ Th- neighbor and averaging these distances. As shown in Figure 6, the relatively unimodal distribution of these distances supports the observation of a consistent mixture of these two neuronal types across all samples. This consistent pattern indicates potential intercellular interactions that facilitate dopamine synthesis, aligning with the proposed hypothesis.

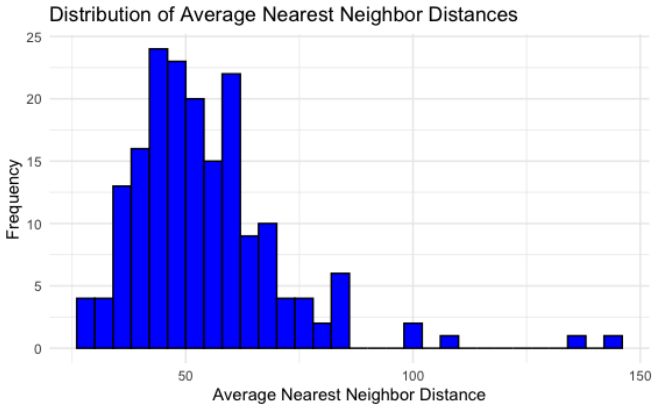

Figure 5: **Histogram of Average Nearest Neighbor Distances Between Ddc+ Th- and Ddc- Th+ Neuronal Classes Across Different Animals and Bregma Locations.** This histogram represents the distribution of average nearest neighbor distances measured in micrometers (µm). The distances were computed using point pattern analysis where each point represents a neuron labeled as either Ddc+ Th- or Ddc- Th+. The analysis was conducted for unique combinations of animal ID, bregma, and sex, focusing on the proximity between the two specified classes within each subset. The relatively unimodal distribution of distances supports the observation of a consistent mixture of these two neuronal types across all samples, indicating potential intercellular interactions that facilitate dopamine synthesis as proposed.

Observations of relatively large average nearest neighbor distances in Figure 5 prompted a closer examination of those specific tissue sections to analyze their cellular distribution patterns. In Figure 6, a snapshot of the slice with the highest average nearest neighbor distance is presented. Despite some degree of mixing between Ddc+ Th- and Ddc- Th+ cells, it is evident that these cells predominantly cluster in two distinct areas.

This clustering suggests a spatial segregation that might imply functional differentiation or a stage of developmental transition within the tissue. Such patterns could potentially influence the efficiency of dopamine synthesis if the physical separation affects the transfer of L-DOPA between TH+; Ddc- cells and TH-; Ddc+ cells. Future investigations might explore the biochemical and signaling pathways that underlie these spatial arrangements to better understand their impact on neuronal function and intercellular communication within the brain.

The observed clustering and spatial distribution patterns noted in the previous analysis raise questions about other factors that might influence the arrangement of neuronal types within the tissue. To explore potential variations in spatial dynamics further, we extended our analysis to examine whether there are sex-based differences in the distribution of Ddc+ Th- and Ddc- Th+ cells across different animal samples.

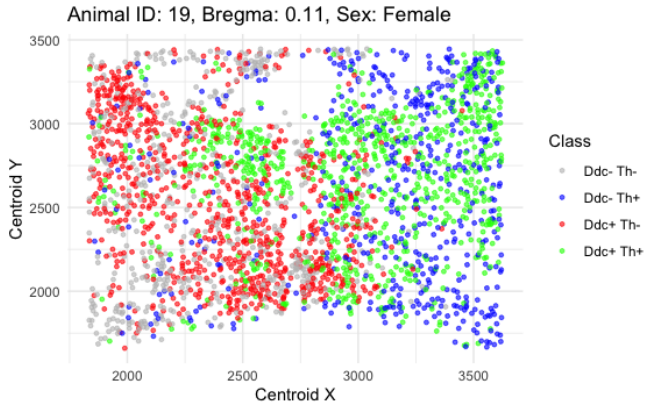

Figure 6: **Spatial Distribution of Neuronal Classes in Animal ID 19, Female, Bregma 0.11.** This plot illustrates the distribution of four neuronal classes—Ddc- Th- (gray), Ddc- Th+ (blue), Ddc+ Th- (red), and Ddc+ Th+ (green)—highlighting their spatial arrangement within the sample. This specific example shows the highest average nearest neighbor distances between the Ddc+ Th- and Ddc- Th+ classes among all analyzed samples. Notably, a certain level of separation between these two classes is observed, suggesting distinct spatial domains and potentially different functional interactions within this particular brain region.

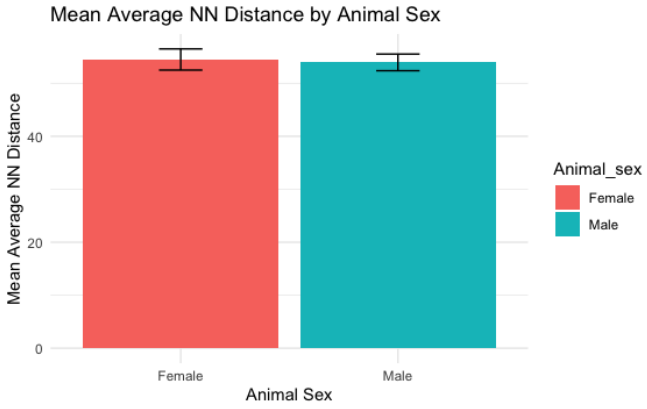

Figure 7: **Comparison of Mean Average Nearest Neighbor (NN) Distances Between Neuronal Classes by Animal Sex.** This bar chart displays the mean average nearest neighbor distances between the neuronal classes Ddc+ Th- and Ddc- Th+ across different samples, categorized by animal sex. Both female (red) and male (teal) animals show similar mean distances, indicating no significant sex-based differences in the spatial distribution of these neuronal classes. Error bars represent the standard error of the mean, underscoring the consistency of the measurements within each sex category.

Figure 7 presents a bar chart comparing the mean average nearest neighbor (NN) distances between these neuronal classes, categorized by animal sex. The data for both female (red) and male (teal) animals show similar mean distances, indicating no significant sex-based differences in the spatial distribution of these neuronal classes. The inclusion of error bars, representing the standard error of the mean, underscores the consistency of these measurements within each sex category. This consistency suggests that the spatial integration of these neuronal types is a robust feature across sexes, further supporting the universality of the observed intercellular interactions that facilitate dopamine synthesis.

Building on the analysis of spatial distributions by animal sex, we also examined the mean average nearest neighbor (NN) distances across different cell classes themselves. This additional layer of analysis was aimed at discerning whether intrinsic differences exist between the neuronal types in terms of their proximity to each other. The results revealed very similar mean NN distances between the different cell classes, suggesting a uniform spatial arrangement irrespective of the specific neuronal type. This uniformity indicates that the proximity required for intercellular communication and dopamine synthesis is maintained across different cell classes, reinforcing the hypothesis of a tightly regulated neurochemical environment within the tissue. This consistent spacing among all classes further supports the notion of a well-integrated dopaminergic network, critical for efficient neural functioning.

Spatial Auto-correlation

There are several methods to compare and test whether there exists evidence of difference in spatial location between groups, and we implemented the spatial auto-correlation using Moran's I statistic. It is a metric which allows us to check whether the spatial data shows a pattern of clustered (I close to 1), dispersed (I close to -1), or randomly spread out (I around 0) [16] By computing the I statistic for Th+;Ddc and Th-;Ddc, we then set up a hypothesis for all distinct combination of 'Animal ID' and 'Bregma' (total 181) to test whether there is a difference of the statistics between these two group.

$$H_0 : I_{DDC+} = I_{DDC-}$$
$$H_1 : I_{DDC+} \neq I_{DDC-}$$

The null hypothesis states that there is no significant difference of spatial auto-correlation between Th+;Ddc- and Th-;Ddc+ whereas the alternative supports that there is a significant spatial difference between the groups. Permutation test was also adapted for the purpose of validating the hypothesis test. By randomly shuffling the data label (Th+;Ddc- and Th-;Ddc+), the I statistics and their difference were computed again for every iteration of the test to check whether the observed difference we got from the original data do show a significant pattern of spatial difference between the groups or simply detected by random chance. From the p-value yielded from the permutation test, we combined all the hypothesis tests we set up for each animal ID and Bregma region

to conduct multiple hypothesis testing, ultimately to see whether there is a statistically significant evidence of spatial distribution difference between the two groups overall. False Discovery Rate (FDR) was applied to do the multiple testing, and 78 out of 181 hypotheses yielded a p-value lower than the adjusted p-value. Consequently, the result suggested that there exists a spatial auto-correlation difference between the groups.

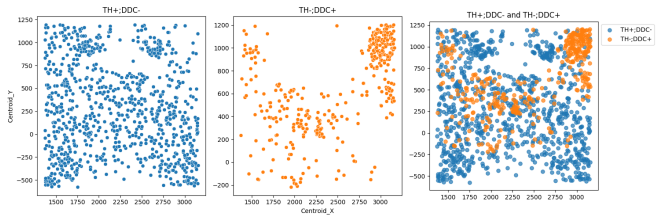

Figure 8: **Spatial Distribution of Th+;Ddc- and Th-;Ddc+ of Animal ID 36, Bregma 0.16** These scatterplots visualize the cells mapping of Th+;Ddc- (Blue) and Th-;Ddc+ (orange) of the sample with the largest Moran's I statistics difference. It is clear to detect the spatial difference between the groups, specifically the former group cells tend to be somewhat randomly spread (leftmost plot, I statistic: -0.109) while there is a cluster with cells concentrated at the top right corner for the latter group (middle plot, I statistic: 0.523).

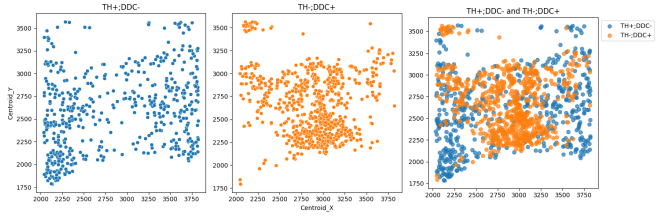

Figure 9: **Spatial Distribution of Th+;Ddc- and Th-;Ddc+ of Animal ID 20, Bregma 0.21** These scatterplots visualize the cells mapping of Th+;Ddc- (Blue) and Th-;Ddc+ (orange) of the sample with the second largest Moran's I statistics difference. As opposite to the previous figure, this sample has a positive I statistics (0.284, clustered) in Th+;Ddc- and negative I statistics (-0.147, dispersed) in Th-;Ddc+. Dispersed spatial indicates that the neighboring locations tend to have dissimilar values, which somewhat agrees that the cells for Th-;Ddc+ are mostly concentrated at the center while Th+;Ddc- cells seem to surround those cells.

To sum up, we were able to find the evidence of spatial difference between TH+;DDC- and TH-;DDC+ from both the test results and visualizations above. By connecting our findings to the research question, this strengthens that there is a notable process that affects the exocytosis activity of L-DOPA from Th+;Ddc- to Th-;Ddc+. However, the samples considered to be significant in terms of spatial difference had relatively small magnitude of Moran's I statistic around -0.3 to 0.3, which suggest somewhat weak to random patterns in the spatial visualizations in general. Furthermore, it is possible that the significant distribution disparities we saw from some samples might be due to unbalanced sample sizes between two groups. As a result, it is still necessary to be cautious even if the results suggest there exists notable evidence.

Bregma	**Count**	$+/-$	$-/+$	**Avg I Diff**
-0.29	4	1	3	0.288
-0.24	1	0	1	0.225
-0.19	4	2	2	0.290
-0.14	4	1	0	0.172
-0.09	6	2	1	0.236
-0.04	4	4	0	0.340
0.01	3	2	1	0.409
0.06	3	1	1	0.321
0.11	12	6	3	0.263
0.16	14	6	7	0.309
0.21	13	8	1	0.308
0.26	10	4	6	0.267

Table 1: **Summary Table of Significant Hypotheses** The table above shows the total count of significant hypotheses (first column), number of hypotheses with positive Moran's I statistic for Th+;Ddc- and negative for Th-;Ddc+ (+/-, third column), vice versa (-/+, fourth column), and the average magnitude of I statistics difference between the groups (fifth column) by Bregma. It is clear that Bregma 0.16 had the most significant hypotheses, followed by 0.21 and 0.11 with relatively large average differences. Specifically for Bregma 0.21, it shows notably consistent trend of positive I statistics for Th+;Ddc- and negative for Th-;Ddc+ (8 out of 13), which could be a signal that the spatial distribution of exocytosis activity tends to change from cluster-concentrated to disperse or random spread out.

3 Discussion

The findings from our scRNA-seq and MERFISH analyses offer insightful revelations about the spatial dynamics and molecular characteristics of partially dopaminergic neurons within the POA. The scRNA-seq data reveals that Ddc and Th genes are expressed across multiple cellular clusters instead of associating exclusively within a single cluster. This suggests a subtle interaction of these genes in neuron function and dopaminergic signaling pathways, possibly reflecting a spectrum of neuron maturation stages or functional states within the brain tissue [17, 18].

Using UMAP for visualization, we can further understand the cellular heterogeneity and the spatial organization of neuronal types. Cells are divided into principal clusters based on Ddc and Th expression profiles, which emphasizes the potential for diverse functional roles among these cells dictated by their molecular signatures. This clustering pattern highlights the plasticity within dopaminergic circuits, which could be crucial for the adaptive responses of neural networks in the brain [19,20].

Integrating scRNA-seq findings with spatial data from the MERFISH dataset allowed us to bridge the gap

between molecular expression and physical positioning within the brain. Although there are some missing values in gene expression in the MERFISH, the machine learning approaches we tried help us interprete this biological dataset [21, 22].. Applying Logistic Regression, the one shows a high accuracy, we found that the cells are closely positioned. This spatial distribution is significant for further analysis of interaction among partial dopaminergic neurons.

The spatial distribution analysis of the target dataset points out there is no significant sex-based differences in the arrangement of neuronal classes, which further emphasizes the fundamental nature of dopaminergic signaling mechanisms as consistent across sexes. This finding is particularly relevant in the context of neurological diseases that may exhibit sex-specific prevalence or symptoms, suggesting that the basic dopaminergic architecture remains stable across genders, possibly affected more by external factors or later- life experiences than by inherent sex-linked genetic differences [23, 24].

Moreover, evaluating the distance between every Ddc-Th+ and its nearest neighbor Ddc+ Th- and calculating the mean, we figured out potential intereaction between cells that promote dopamine synthesis exists. This could imply a more integrative and less compartmentalized arrangement within the dopaminergic systems, potentially facilitating versatile interactions necessary for complex brain functions such as mood regulation, reward processing, and motor control [25, 26].

Utilizing Moran's I statistics and multiple hypothesis testing to detect spatial auto-correlation between Th+;Ddc- and Th-;dc+ cells across various brain sections and obtain quantitative measures of their proximity, we were able to provide notable evidence of spatial difference of dopamine synthesis. Although there was no general trend detected across the cells, sufficient number of cells still mapped a consistent pattern of spatial distribution from concentrated to dispersed/random or vice versa. The changes of both the cells' location and overall shape may be an indication that L-DOPA exocytosis activity from Th+;Ddc- cells to nearby Th-;Ddc+ cells could be a crucial mechanism affecting and enabling efficient dopamine production within spatially structured neuronal networks. This could represent a foundational mechanism underlying the robustness of dopaminergic neurotransmission, especially in regions critical for motor and cognitive functions [27, 28]. Alternatively, this could be an indication that the L-DOPA and dopamine biosynthetic landscape is dynamic and has evolved in a manner allowing it to facilitate both completion of dopamine biosynthesis when necessary and supporting the alternative physiological functions of L-DOPA when needed. [6,7] Nonetheless, one should be aware of adapting the analyses as there are many factors which can affect the results such as different methods of data pre-processing, hyperparameters selection, etc.

Future research should aim to expand upon these findings by incorporating more dynamic assessments of neuron activity and interactions, potentially through live imaging or more sophisticated spatial transcriptomics techniques. Additionally, exploring the biochemical pathways that enable L-DOPA transfer and dopamine synthesis within these identified neuron clusters could unveil new therapeutic targets for enhancing dopaminergic function in pathological states such as Parkinson's disease and depression [29, 30].

In conclusion, our comprehensive analysis not only sheds light on the cellular architecture and molecular diversity of dopaminergic neurons but also enhances our understanding of the spatial and functional integration within the brain. This integration is vital for the coordinated activity necessary for effective neurotransmission and overall brain health, offering valuable insights into the complexities of neural systems biology [31, 32].

Contributions

EH performed the scRNA-seq analysis, developed the classifier, implemented the average nearest neighbor algorithm, and wrote the corresponding results and discussion sections. JY performed the Spatial Auto-correlation analysis, authored that section, and contributed to the discussion. SB wrote the introduction and edited the document. SS contributed to the discussion part.

References

[1] David S. Goldstein. Catecholamines 101. *Clinical Autonomic Research*, 20(6):331–352, 2010.

[2] Jeffrey R. Moffitt, Dhananjay Bambah-Mukku, Stephen W. Eichhorn, Eric Vaughn, Karthik Shekhar, Julio D. Perez, Nimrod D. Rubinstein, Junjie Hao, Aviv Regev, Catherine Dulac, and Xiaowei Zhuang. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science*, 362(6416):165–171, 2004.

[3] Stephen X. Zhang, Andrew Lutas, Shang Yang, Adriana Diaz, Hugo Fluhr, Georg Nagel, Shiqiang Gao, and Mark L. Andermann. Hypothalamic dopamine neurons motivate mating through persistent camp signalling. *Nature*, 597(7875):245–249, 2021.

[4] Irma Gvilia, Feng Xu, and Ronald Szymisiak. Homeostatic regulation of sleep: A role for preoptic area neurons. *The Journal of Neuoscience*, 26(37):9426–9433, 2006.

[5] Renata C.H. Barros, Luiz G.S. Branco, and Evelin C. Cárnio. Evidence for thermoregulation by dopamine d1 and d2 receptors in the anteroventral preoptic region during normoxia and hypoxia. *Brain Research*, 1030(2), 2018.

[6] Jeremy P.E. Spencer, Andrew Jenner, John Butler, Okezie I. Aruoma, David T. Dexter, Peter Jenner, and Barry Halliwell. Evaluation of the pro-oxidant and antioxidant actions of l-dopa and dopamine in

vitro: Implications for parkinson's disease. *Free Radical Research*, 24(2):95–105, 1996.

[7] Lili Arabuli, Petra Lovecka, Rudolf Jezek, Viktorovam Jitka, Tomas Macek, Petra Junkova, Ramaz Gakhokidze, Fariborz Sharifianjazi, Amirhossein Esmaeilkhanian, Peyman Salahsour, Parnian Poursafa, and Parisa Sabouri. Ache inhibitory effect, antioxidant and anti-inflammatory properties of cyclen and l-dopa related compounds: Targeting in neurodegenerative disease. *Journal of Molecular Structure*, 1287, 2023.

[8] Satija Lab. Pbmc 3k guided tutorial. `https://satijalab.org/seurat/articles/pbmc3k_tutorial`, 2023. Accessed: 2023-05-12.

[9] Leland McInnes and John Healy. Umap: Uniform manifold approximation and projection for dimension reduction. `https://umap-learn.readthedocs.io`, 2023. Accessed: 2023-05-12.

[10] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[11] Christopher M Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[12] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. *Introduction to Data Mining*. Pearson Education, 2005.

[13] Max Kuhn and Kjell Johnson. *Applied Predictive Modeling*. Springer, 2013.

[14] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics, 2009.

[15] Ankit Dixit. *Ensemble Machine Learning: Methods and Applications*. Springer, 2018.

[16] Arthur Getis and J. K. Ord. The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24(3), 1992.

[17] Stephan Lammel, Byung Kook Lim, Caroline Ran, Kevin W Huang, M. James Betley, Kay M Tye, Karl Deisseroth, and Robert C Malenka. Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. *Neuron*, 82(5):1039–1053, 2014.

[18] Anthony A Grace. Regulation of dopamine neuron activity and its implications for the functional roles of dopamine. *Biochemical Society Transactions*, 2016.

[19] Jean-Martin Beaulieu, Tatyana D Sotnikova, Raul R Gainetdinov, and Marc G Caron. An akt/beta-arrestin 2/pp2a signaling complex mediates dopaminergic neurotransmission and behavior. *Cell*, 122(2):261–273, 2005.

[20] David Sulzer. Multiple hit hypotheses for dopamine neuron loss in parkinson's disease. *Trends in Neurosciences*, 30(5):244–250, 2007.

[21] Xuyu Zhang, Tian Li, Fei Liu, Yu Chen, Jian Yao, Zixiao Li, Yanyi Huang, and Jian Wang. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*, 37:547–555, 2019.

[22] Taesup Kim, Youngjun Song, and Kwang-Hyun Cho. Deep learning approaches to biological image reconstruction. *Bioinformatics*, 2020.

[23] Glenda E Gillies and Simon McArthur. Sex differences in parkinson's disease and other movement disorders. *Experimental Neurology*, 259:44–56, 2014.

[24] Jill B Becker and Elena Chartoff. Sex differences in drug abuse. *Frontiers in Neuroendocrinology*, 35(1):36–47, 2016.

[25] Roshan Cools, Angela C Roberts, and Trevor W Robbins. Serotonin and dopamine: Unifying affective, activational, and decision functions. *Neuropsychopharmacology*, 36(1):98–113, 2011.

[26] Roy A Wise. Dopamine, learning and motivation. *Nature Reviews Neuroscience*, 5(6):483–494, 2014.

[27] Charles R Gerfen and D James Surmeier. Basal ganglia. *The Rat Nervous System*, pages 441–484, 2011.

[28] André Nieoullon. Dopamine and the regulation of cognition and attention. *Progress in Neurobiology*, 67(1):53–83, 2002.

[29] Jose A Obeso, Maria Stamelou, Christopher G Goetz, Werner Poewe, Anthony E Lang, Daniel Weintraub, David Burn, Glenda M Halliday, Erwan Bezard, Serge Przedborski, Stephane Lehericy, David J Brooks, John C Rothwell, Mark Hallett, Mahlon R DeLong, Connie Marras, Caroline M Tanner, G Webster Ross, J William Langston, Christine Klein, Vincenzo Bonifati, Joseph Jankovic, Andres M Lozano, Günter Deuschl, Hagai Bergman, Eduardo Tolosa, Mayela Rodriguez-Violante, Stanley Fahn, Paul Krack, Connie Marras, Daniel Mendelsohn, Peter Riederer, Anthony HV Schapira, and Werner Poewe. Past, present, and future of parkinson's disease: A special essay on the 200th anniversary of the shaking palsy. *Movement Disorders*, 32(9):1264–1310, 2017.

[30] Barbara Picconi, Diego Centonze, Kerstin Håkansson, Giorgio Bernardi, Paolo Calabresi, Paul Greengard, and Gilberto Fisone. Loss of bidirectional striatal synaptic plasticity in l-dopa-induced dyskinesia. *Nature Neuroscience*, 9(6):501, 2008.

[31] Karl Deisseroth. Optogenetics: 10 years of microbial opsins in neuroscience. *Nature Neuroscience*, 18(9):1213–1225, 2015.

[32] D James Surmeier, Jose A Obeso, and Glenda M Halliday. Dopaminergic modulation of striatal networks in health and parkinson's disease. *Current Opinion in Neurobiology*, 48:88–96, 2017.