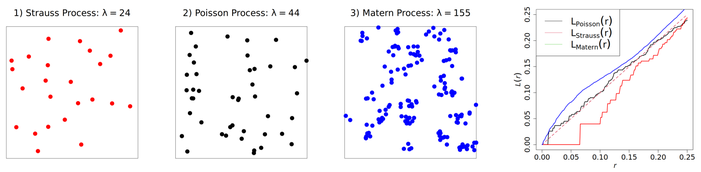
Spatial transcriptomic data reveals pure cell types via the mosaic hypothesis

Yiliu Wang[1,2]; Christof Koch[1,3]; Uygar Sümbül[1]
[1]Allen Institute, [2]University of Washington, [3]Tiny Blue Dot Foundation

Abstract

Clustering in high-dimensional feature spaces, such as those defined by single-cell RNA-seq data, is often inconclusive. In this work, we try to define cell types from the spatial prospect via the mosaics hypothesis.

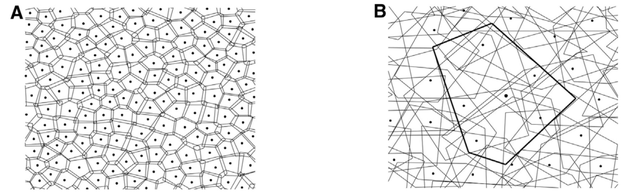
Mosaics hypothesis modeled using Strauss point process and examined by L-function. Cells of one type are neither widely spaced, nor crowded together.



Introduction

In the retina, a layered structure, neurons of the same discrete type avoid spatial proximity with each other. While this principle, which is independent of clustering in feature space, has been a gold standard for retinal cell types, its applicability to the cortex has been only sparsely explored.

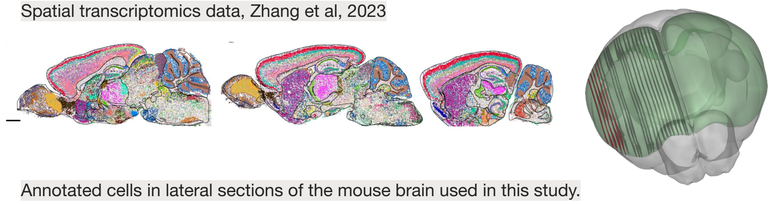
Tiling v.s. Mosaicking in Retina, Seung and Sümbül, 2014



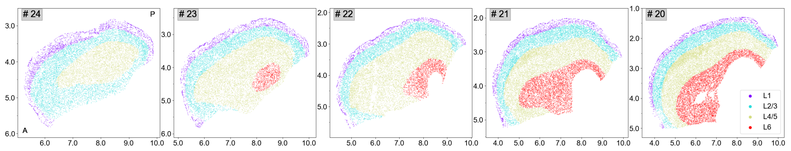
Photoreceptor, bipolar, horizontal, amacrine cells

Ganglion cells

Spatial transcriptomics data, Zhang et al 2023



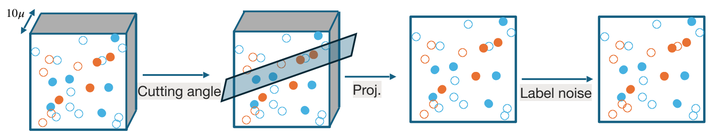
Annotated cells in lateral sections of the mouse brain used in this study.



Data Limitations

There are various types of noises in the dataset,

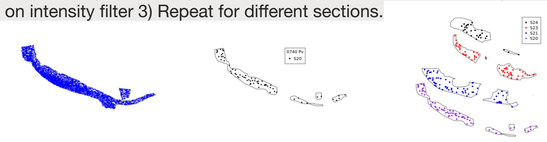
- 3D to 2D projection error
- Segmentation and profiling
- Mixing-up, inaccurate reference and algorithmic noise
- Region bias



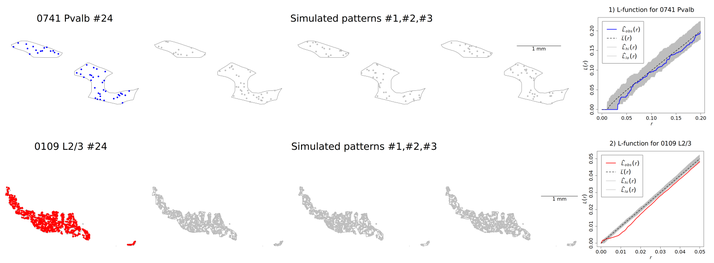
We study cell types that are less affected by these noises.

Self-avoidance at cluster level

Window selection procedure 1) Use all cells to define the boundaries 2) Refine the window based on intensity filter 3) Repeat for different sections.



Hypothesis testing procedure for two exemplar GABAergic and Glutamatergic clusters.



L-functions of the observed two patterns and simulated baselines

Cluster	Section	p-value
0109 L2/3 IT CTX Glut_2	20–24	7.5e-08
0074 L4/5 IT CTX Glut_1	20–24	1.1e-06
0079 L4/5 IT CTX Glut_2	20,21,24	4.1e-05
0455 L6 CT CTX Glut_5	20,21	3.9e-04
0116 L2/3 IT CTX Glut_3	21,23,24	1.1e-03
0092 L4/5 IT CTX Glut_4	23,24	3.0e-03
0061 L5 IT CTX Glut_3	20	5.0e-03
0114 L2/3 IT CTX Glut_3	24	5.0e-03
0443 L6 CT CTX Glut_2	21	5.0e-03
0086 L4/5 IT CTX Glut_3	23,24	6.6e-03
0375 L5 ET CTX Glut_5	23	0.011
0026 L5/6 IT TPE-ENT Glut_2	23	0.015
0439 L6 CT CTX Glut_1	21	0.024
0007 IT EP-CLA Glut_1	20	0.027
0359 L5 ET CTX Glut_2	23	0.039
0003 CLA-EPd-CTX Car3 Glut_1	20,21	0.044
0453 L6 CT CTX Glut_5	20	0.061
0004 CLA-EPd-CTX Car3 Glut_1	23, 21	0.14

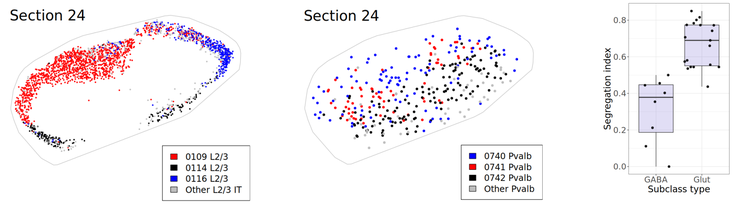
A majority of Glutamatergic clusters and all GABAergic clusters display self-avoidance.

0754 Pvalb Gaba_8	20, 21	0.002
0740 Pvalb Gaba_3	20,21,23,24	0.004
0741 Pvalb Gaba_3	24	0.007
0624 Vip Gaba_1	24	0.011
0641 Vip Gaba_5	22	0.013
0645 Vip Gaba_6	20,24	0.015
0747 Pvalb Gaba_4	20	0.017
0742 Pvalb Gaba_3	23,24	0.018
0811 Sst Gaba_13	21	0.018
0709 Lamp5 Gaba_1	22	0.029

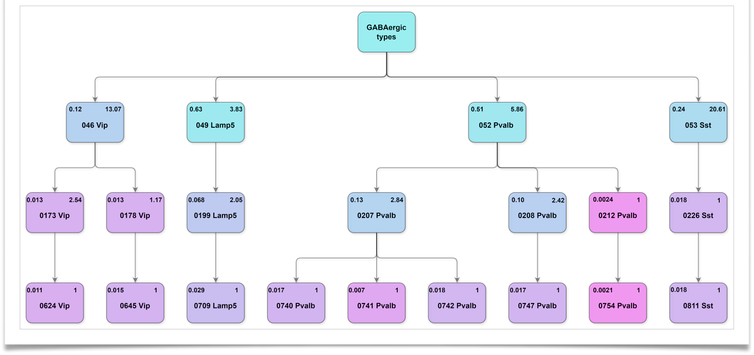
Disappearance of self-avoidance at mixture level

Given that many cell types defined at the finest transcriptional level display self-avoidance, we next study whether this is true at coarser levels of the transcriptional hierarchy as well.

Segregation effect at the subclass level.



Increase in p-values for GABAergic mixture types.When moving up the hierarchy, pure cell types merge into supertypes and subclasse, and spatial avoidance disappears.



Discussions

Summary of our contributions:

- A statistical point process analysis framework for spatial transcriptomics data.
- Novel filters: introduce theoretically grounded cell type filters based on relative abundance and intensity of high-confidence cells.
- Statistical test for spatial organization of cell bodies: provide evidence for mosaic hypothesis at the t-type level.
- Subclass and supertype level analysis: self-avoidance behavior disappears for mixture of cell types.
- A potential gold standard metric for evaluating purity of cell types.

We wish to thank the founder of the Allen Institute for Brain Science, P. G. Allen, for his vision, encouragement and support. The original motivation to pursue this question arose from discussions between Francis Crick and C.K. many decades earlier. We wish to thank Zizhen Yao, Michael Kunst, and Kelly Jin for useful discussions, Forrest Collman and Leila Elabbady for guidance on the MICrONS dataset, Forrest Collman and Stephen Smith for valuable feedback on the manuscript. Y.W. is a Shanahan Foundation Fellow.

Contact

<Yiliu Wang>
<Allen Institute for Brain Science>
BioRxiv paper: https://www.biorxiv.org/content/10.1101/2024.08.09.607193v1
Email: yiliu.wang@alleninstitute.org

References

1. Y. Wang, C. Koch and U. Sümbül, Spatial transcriptomic data reveals pure cell types via the mosaic hypothesis, 2024
2. F. Crick, Cell types in the neocortex: The tiling hypothesis, 1983
3. R. L. Rockhill, T. Euler, and R. H. Masland, Spatial order within but not between types of retinal neurons, 2000
4. H. S. Seung and U. Sümbül, Neuronal cell types and connectivity: lessons from the retina, 2014
5. H. Maruoka et al, Lattice system of functionally distinct cell types in the neocortex, 2017
6. Z. Yao et al, A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain, 2023
7. M. Zhang et al, Molecularly defined and spatially resolved cell atlas of the whole mouse brain, 2023
8. Q. Wang et al, The allen mouse brain common coordinate framework: A 3d reference atlas, 2020
9. A. Baddeley et al, Spatial point patterns: Spatial Point Patterns: Methodology and Applications with R, 2015
10. J. Illian et al, Statistical analysis and modelling of spatial point patterns, 2008
11. Y. Benjamini and Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, 1995
12. H. Zeng, What is a cell type and how to define it?, 2022

alleninstitute.org