UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Multi-Method Analysis of Cellular Composition and Cell—Cell Communication with MRI-Coupled Single-Cell Analysis

Permalink

https://escholarship.org/uc/item/7088n610

ISBN

9798173783844

Author

Lee, Yu Rim

Publication Date

2026-11-11

Peer reviewed|Thesis/dissertation

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

UNIVERSITY OF CALIFORNIA SAN DIEGO

Multi-Method Analysis of Cellular Composition and Cell–Cell Communication with MRI-Coupled Single-Cell Analysis

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Bioengineering

by

Yu Rim Lee

Committee in charge:

Professor Nathan Lewis, Chair
Professor Ester Kwon, Co-Chair
Professor Erika Cyphert
Professor Alyssa Taylor

2026

Copyright

Yu Rim Lee, 2026

All rights reserved.

The Thesis of Yu Rim Lee is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2026

DEDICATION

This thesis is dedicated to my father, Jun Seo Lee, and my aunt, Jinsuk Lee, for their unwavering support, and to my dogs, Hockney and Neleri, who offered no scientific input but ample emotional support.

TABLE OF CONTENTS

THESIS APPROVAL PAGE iii
DEDICATION iv
TABLE OF CONTENTS v
LIST OF FIGURES vii
LIST OF TABLES viii
LIST OF ABBREVIATIONS ix
ACKNOWLEDGEMENTS x
ABSTRACT OF THE THESIS xi
Chapter 1. Introduction 1
1.1 Single-Nucleus RNA Sequencing for Transcriptomic Profiling 1
1.2 Cell-to-Cell Communication Analysis and Current Tools 1
1.3 LIANA and Tensor-cell2cell for Multi-Context CCC Inference 2
Chapter 2. Exploratory Multi-Method Approaches to Cell Composition Scoring for Bias Mitigation 5
2.1 Introduction 5
2.1.1 Accounting for cell composition in CCC inference 5
2.1.2 Compositional Modeling as a Preprocessing Step for CCC Analysis 5
2.1.3 scCODA: Bayesian Compositional Analysis 6
2.1.4 MELD: Manifold-Based Perturbation Scoring 7
2.2 Methods 8
2.2.1 Participants and Data 8
2.2.2 Data Preprocessing 8
2.2.3 Raw Frequency Scoring 9
2.2.4 MELD scoring 9
2.2.5 scCODA scoring 10
2.2.6 cell_composition framework 10
2.3 Results 11
2.3.1 MELD captured condition-associated cell-state shifts 11
2.3.1.1 Spearman correlation and Scatter plot 11
2.3.1.2 Rank discrepancy comparison 14
2.3.2 scCODA-based compositional analysis highlighted Truncated Radial Glia 18
2.4 Discussions 19

Chapter 3. MRI-Informed Single-Cell Analysis of ASD BCOs....23
3.1 Introduction....23
3.2 Methods....24
3.2.1 Participants, snRNA-seq, and Neuroimaging Data....24
3.2.2 Ligand–Receptor Interaction Analysis and Tensor Decomposition....25
3.2.3 Yeo Network Parcellation....26
3.3 Results....27
3.3.1 Tensor-cell2cell with Brain MRI information....27
3.3.1.1 Month-1 Tensor-cell2cell Factors....27
3.3.1.2 Month-2 Tensor-cell2cell Factors....32
3.3.2 BCO Size and MRI Brain Volume Correlation....37
3.4 Discussions....39
Conclusion....42
Supplementary Materials....43
REFERENCES....56

LIST OF FIGURES

Figure 2.1. Frequency vs. MELD scores in TD and ASD 14

Figure 2.2. Frequency versus MELD rank discrepancies in ASD. Points are plotted by rank difference 16

Figure 2.3. Posterior inclusion probabilities for cell types in ASD 19

Figure 3.1. Tensor-cell2cell decomposition of month-1 BCO 28

Figure 3.2. Box plots showing the distribution of factor loadings for all ten factors across MRI brain volume groups 29

Figure 3.3. Dot plot showing GSEA results for KEGG pathways across Tensor-cell2cell factors derived from month-1 BCO snRNA-seq data 31

Figure 3.4. Tensor-cell2cell decomposition of month-2 BCO 33

Figure 3.5. Box plots showing the distribution of factor loadings for all sixteen factors across MRI brain volume groups 34

Figure 3.6. Dot plot showing GSEA results for KEGG pathways across Tensor-cell2cell factors derived from month-2 BCO snRNA-seq data 36

Figure 3.7. Scatter plots showing Spearman rank correlations between BCO size (μm²) and MRI-derived regional network volumes (mm³) 38

Figure 3.8. Bar plot summarizing Spearman ρ values between BCO size and MRI-derived regional network volumes across all seven Yeo networks (n = 4, month-1) 38

Supplementary Figure A. Elbow plots for Month-1 and Month-2 Tensor-cell2cell analysis ..44

LIST OF TABLES

Table 2.1. Frequency and MELD scores for tRG across ASD and TD conditions 12

Table 2.2. MELD and frequency differences (ASD − TD) for cell types with the greatest TD-associated signal reduction 13

Table 2.3. Discrepancy between frequency and MELD ranks in the ASD condition. Ranks are ordered descending (Rank 1 = highest) 17

Table 3.1. MRI-derived Mean Cortical Regional Volume, LH (mm^3) and size bin assignments for ASD donors included in Tensor-cell2cell analysis, stratified by organoid age 26

Table 3.2. MRI-derived regional network volumes and morphometric ranks for ASD donors included in the cross-modal correlation analysis (n = 4, month-1), based on the Yeo 7-network parcellation.....37

Supplementary Table A. Cell type composition scores across three methods (frequency, MELD, and scCODA) for all annotated cell types in ASD and TD brain cortical organoids 43

Supplementary Table B. Cell-type sender/receiver identities and top ligand-receptor pairs for notable Tensor-cell2cell factors derived from month-1 BCO snRNA-seq data.....45

Supplementary Table C. Cell-type sender/receiver identities and top ligand-receptor pairs for notable Tensor-cell2cell factors derived from month-2 BCO snRNA-seq data.....50

LIST OF ABBREVIATIONS

CCC	Cell-to-cell communication
BCOs	Brain Cortical Organoids
ASD	Autism Spectrum Disorder
tRG	Truncated Radial Glia
OPC	Oligodendrocyte Precursor Cells
NPC	Neural Progenitor Cell
EXT	Excitatory Neurons
INT	Interneurons
DMN	Default Mode Network

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Dr. Nathan Lewis, for his continued and insightful guidance and support throughout this thesis. I am also grateful to Dr. Ester Kwon, who not only served as co-chair but offered invaluable advice during a difficult period of this process. I would also like to thank my committee members, Dr. Erika Cyphert and Dr. Alyssa Taylor, for their time and feedback on this work. I am thankful to Dr. Hratch Baghdassarian, who developed the original code used in Chapter 2, and Dr. Caitlin Aamodt provided guidance for the analyses in Chapter 3. I would also like to thank the members of the Lewis Lab for their support throughout this process. Finally, I would like to thank the families and children who participated in this study. Their generous contribution made this study possible.

ABSTRACT OF THE THESIS

Multi-Method Analysis of Cellular Composition and Cell–Cell Communication with MRI-Coupled Single-Cell Analysis

by

Yu Rim Lee

Master of Science in Bioengineering

University of California San Diego, 2026

Professor Nathan Lewis, Chair
Professor Ester Kwon, Co-Chair

This thesis evaluates computational approaches for characterizing cell composition alterations and applies an integrated computational framework to investigate cell–cell communication (CCC) in brain cortical organoids (BCOs) derived from autism spectrum disorder (ASD) donors, linking CCC changes to donor-specific brain structure measured by magnetic resonance imaging (MRI). To assess cell type abundance-related bias that can confound CCC inference, a dedicated cell composition analysis framework was adapted, in which single-nucleus RNA sequencing (snRNA-seq) data of patient-derived BCOs were analyzed using three complementary composition-scoring methods: raw cell-type frequency, MELD (Manifold Enhancement of Latent Dimensions), and scCODA (Single-cell Compositional Data Analysis).

First, the three approaches were compared to characterize cell type composition differences between ASD and typically developing (TD) BCOs (N = 10 donors), testing whether compositional shifts were recurring across methods with different sensitivities to abundance-related bias. Truncated radial glia (tRG) emerged as a population of interest across the three approaches, showing lower relative abundance in ASD, differences in MELD condition-likelihood scores, and elevated posterior inclusion probabilities in scCODA. Next, MRI-informed CCC analysis was performed in ASD BCOs using LIANA (Ligand-receptor Interaction ANAlysis) and Tensor-cell2cell to identify signaling programs associated with donor brain volume across two organoid ages. Recurrent programs included NLGN-NRXN interactions, the MDK-PTPRZ1 trophic axis, axon guidance, extracellular matrix and integrin signaling, and WNT-associated pathways. In an exploratory cross-modal analysis, larger BCO size was inversely associated with default mode network (DMN) volume, suggesting a relationship between prenatal organoid growth dynamics and postnatal cortical network morphology. Although composition scores were retained as an exploratory analysis rather than integrated into the final CCC inference pipeline of Tensor-cell2cell, these findings provide a proof of concept for multi-method assessment of cell composition bias prior to CCC analysis and suggest that altered progenitor composition and intercellular signaling may contribute to organoid overgrowth and neuroanatomical variation in ASD.

Chapter 1. Introduction

1.1 Single-Nucleus RNA Sequencing for Transcriptomic Profiling

Recent advances in next-generation sequencing focusing at a single-cell level, in contrast to the bulk RNA-seq, have enabled high-resolution transcriptomic profiling. Particularly, the single-nucleus RNA sequencing (snRNA-seq) eliminated the restrictions and risks that came along with the scRNA-seq's cell dissociation methods to analyze the whole cells. Instead, snRNA-seq's focus on the nucleus expanded its applicability to the samples where fresh tissue is unavailable, such as brain organoids and frozen clinical specimens (Guo et al. 2023). The resultant expression profiling at high resolution, combined with in silico analysis, facilitate the profiling of cell-to-cell communication for condition-associated state shifts, opening new opportunities to systematically identify underlying pathophysiological cell signaling mechanisms (Armingol et al. 2024).

1.2 Cell-to-Cell Communication Analysis and Current Tools

Cell-to-cell communication (CCC) occurs through protein-protein interactions between secreted ligands and plasma membrane receptors. This subtype of cell-cell interactions via ligand–receptor interactions (LRIs) are essential for multicellular organisms to perform collective biological functions, from development to homeostasis. Despite their differences, every CCC tool shares two major components: a prior knowledge of ligand-receptor interactions, and a method to estimate CCC from those known interactions given the dataset at hand. To systematically unearth CCC patterns, the analysis starts from gene expression data by leveraging the concurrent expression of ligands and their cognate receptors. Once the gene expression

matrix is filtered to include only ligands and receptors, expression levels are aggregated across single cells within each cell type. Each LRI's interaction strength is then scored based on ligand expression in the sender cell type and receptor expression in the receiver cell type, and a communication score is assigned to each LRI in each sender-receiver cell-type pair. However, computational tools diverge in how they implement each of these steps, from the mathematical model used to compute communication scores, to the ligand-receptor database employed, and the statistical methods applied to determine significance (Armingol et al. 2024; Dimitrov et al. 2022).

Multiple scoring methods have been developed, including CellPhoneDB, Connectome, NATMI, SingleCellSignalR, CellChat, and Rank_Aggregate, each constructing a distinctive scoring system through permutation of cluster labels, regularizations, and scaling to prioritize interactions according to the input dataset. What these methods have in common is that they estimate the likelihood of communication based on the co-expression level of transcripts for sender and receiver cells (Cheng et al. 2023).

1.3 LIANA and Tensor-cell2cell for Multi-Context CCC Inference

Among the various computational methods, first-generation CCC tools are subject to two key limitations. First, most tools are limited to the analysis of single or paired samples at a time, preventing systematic comparison across multiple biological contexts such as disease severity, developmental time points, or treatment conditions. Second, these tools can only identify known LRI patterns present in predefined databases, unlike in vitro or in vivo experiments capable of

discovering novel interactions. This inherent dependence on the choice of ligand-receptor database introduces bias that is difficult to quantify or control.

To address the limitation of single or pairwise sample comparison, Tensor-cell2cell constructs a 4-dimensional tensor from CCC scores. On top of the 3D matrix composed of sender cells, receiver cells, ligand-receptor pairs, it adds the 4th dimension of biological contexts. This 4D framework enables the simultaneous comparison of CCC patterns across theoretically unlimited numbers of biological contexts, constrained only by computing power. Tensor-cell2cell has been applied to COVID-19 severity data, where it identified a CCC pattern most correlated with disease severity and revealed macrophages as the major pro-inflammatory sender cells, with CCL2–CCR1, CCL3–CCR5, and CCL8 as the top ligand-receptor pairs — findings that coincided with independent genomic analyses proposing CCR1 and CCR5 as therapeutic targets (Armingol et al. 2022).

To mitigate the inherent database dependency of CCC inference, LIANA (Ligand-receptor Interaction ANAlysis) aggregates a consensus across 17 curated ligand-receptor databases, including CellChatDB, CellPhoneDB, CellTalkDB, connectomeDB2020, ICELLNET, Kirouac2010, talklr, SIGNALink, and others. By integrating multiple resources and computing a rank aggregate score across scoring methods, LIANA acts as a platform to induce the "consensus" of databases. This substantially reduces the bias introduced by any single database or scoring method, improving the robustness and reproducibility of CCC inference (Dimitrov et al. 2022).

While Tensor-cell2cell provides a powerful tensor-based framework for extracting CCC patterns across multiple contexts, the biological interpretation of those patterns is dictated by the underlying ligand-receptor database. Coupling LIANA's consensus-based database aggregation with Tensor-cell2cell's multi-context pattern extraction therefore represents a principled strategy for robust and comprehensive CCC inference (Armingol et al. 2024; Dimitrov et al. 2022).

Chapter 2. Exploratory Multi-Method Approaches to Cell Composition Scoring for Bias Mitigation

2.1 Introduction

2.1.1 Accounting for cell composition in CCC inference

The scoring methodology adopted by the pipeline contributes significantly to CCC inference, with each approach carrying distinct advantages and limitations. Since the gene expression matrix of ligands and their cognate receptors is the primary input for CCC inference, interactions involving highly expressed ligand-receptor pairs are inherently overrepresented in the analysis of CCC patterns. However, highly expressed ligand-receptor pairs from the sequencing do not accurately reflect their biological significance in cell signaling in an organism. Expression levels are substantially influenced by the cell type composition of the sample, which can vary across sampling sites, donor age, disease state, and developmental or immune context, independent of any true biological signal (Armingol et al. 2021; Dimitrov et al. 2022). Without considering such factors, a CCC pattern may appear as the dominant signal not because of its biological significance, but simply due to the disproportionate abundance of its sender or receiver cell populations. Therefore, for biologically meaningful CCC inference in the context of a pathophysiology of interest, analysis pipelines must account for this compositional heterogeneity.

2.1.2 Compositional Modeling as a Preprocessing Step for CCC Analysis

To mitigate these compositional biases, this chapter evaluates the feasibility of incorporating compositional and density-based statistical methods, scCODA (Single-cell

Compositional Data Analysis) and MELD (Manifold Enhancement of Latent Dimensions), as a preprocessing step prior to LIANA and Tensor-cell2cell analysis. By characterizing cell type composition differences between 'ASD' (Autism Spectrum Disorder) and 'TD' (Typically Developing) conditions of toddlers, this chapter aims to determine whether compositional variation should be accounted for before CCC pattern extraction, and to evaluate the interpretability of two methodologically distinct approaches.

Both scCODA and MELD assess perturbation effects in condition versus control experiments, but at different resolutions. scCODA is based on a Bayesian hierarchical model for compositional data analysis, where the focus is on the relative changes in cell type abundances, whereas MELD is based on a more general approach of graph signal processing and local density estimation on a cell-cell similarity graph, enabling the smooth and continuous estimation of perturbation likelihoods (Büttner et al. 2021; Burkhardt et al. 2021). Having two distinct methodologies would be useful depending on whether the biological question involves discrete cell type shifts or continuous cell state transitions.

2.1.3 scCODA: Bayesian Compositional Analysis

scCODA is a Bayesian model for differential cell type composition across conditions, built on a hierarchical Dirichlet-Multinomial distribution for cell counts. It explicitly accounts for compositional constraints by jointly modeling all measured cell types, thereby mitigating the misleading negative correlations that arise when proportions are forced to sum to a constant. A

key strength of scCODA is its Bayesian formulation, which quantifies uncertainty in observed counts, applies sparsity-inducing priors for automatic variable selection, and avoids heavy reliance on asymptotic approximations. Through joint Bayesian inference across multiple cell types, scCODA identifies specific cell types whose abundances change between conditions while correcting for compositional bias (Büttner et al. 2021).

2.1.4 MELD: Manifold-Based Perturbation Scoring

MELD is a graph-based framework for quantifying cell-level perturbation responses. It assigns each cell a continuous perturbation score by estimating local density differences between conditions on a k-nearest-neighbor graph. By smoothing condition labels across the cell-cell similarity manifold, MELD produces a per-cell likelihood of belonging to each experimental condition. MELD couples these scores with vertex frequency clustering to identify regions of the manifold whose abundance and transcriptional state are most responsive to a perturbation. In disease contexts, MELD enables reframing of CCC analysis from general cell type summaries to specific regions of cell state space, facilitating the resolution of perturbation-associated signaling at a more detailed level. Compared to scCODA, which summarizes compositional changes at the cell type level, MELD captures continuous gradients of cell state, making the two methods complementary in their characterization of condition-associated variation (Burkhardt et al. 2021).

2.2 Methods

2.2.1 Participants and Data

Single-nucleus RNA sequencing data were obtained from brain cortical organoids (BCOs) derived from 6 ASD and 4 typically developing (TD) donors (total N = 10), as previously described in (Courchesne et al. 2024). BCOs were generated from iPSCs reprogrammed from peripheral blood mononuclear cells and grown for 30 and 60 days to capture early embryonic neural progenitor cells and neurons. All participants were ascertained, recruited, and phenotyped using uniform procedures at the UC San Diego Autism Center of Excellence.

2.2.2 Data Preprocessing

Raw count data were processed and stored using AnnData (Virshup et al. 2024). Quality control was applied to filter low-quality cells using the following thresholds: total UMI count < 10,000, mitochondrial gene percentage < 5%, and detected gene count between 200 and 5,500. Doublets were identified and removed using Scrublet with batch correction applied per sequencing pool (batch_key="orig.ident"). Counts were normalized to 10,000 reads per cell and log-transformed. Highly variable genes were selected using the Seurat flavor (n = 2,000, batch_key="orig.ident"), and principal component analysis (PCA) was performed using the top 20 components. Cell type identities were obtained from the anno_clusters annotation, which was renamed to cell_type prior to compositional analysis.

2.2.3 Raw Frequency Scoring

As a baseline estimate that does not account for the composition bias, raw frequency scores were calculated by dividing the number of cells assigned to each cell type within a condition by the total number of cells observed in that condition, formally defined as follows:

$$Frequency_{c,k} = \frac{n_{c,k}}{N_c}$$

$n_{c,k}$denotes the number of cells assigned to cell type k under condition c , and N_c denotes the total number of cells observed under condition c .Thus, $Frequency_{c,k}$ represents the relative abundance of cell type k under condition c .

2.2.4 MELD scoring

MELD was run with default parameters using the PCA embedding as the transcriptomic manifold. Per-cell likelihood scores were estimated via `meld.MELD().fit_transform()`. Likelihoods were normalized across conditions using `meld.utils.normalize_densities()`. For each cell, the likelihood score corresponding to its observed condition was retained, yielding a single composition score per cell. Cell type-level scores were obtained by aggregating per-cell scores within each annotated cell type. This approach captures subtle, continuous gradients in cell state shifts that are invisible to frequency-based methods.

2.2.5 scCODA scoring

For scCODA-based scoring, cell counts were summarized into a sample-by-cell-type count matrix using `sccoda.util.cell_composition_data.from_scanpy()`. A Dirichlet-Multinomial model was fitted with diagnosis (TD or ASD) as the covariate and reference cell type set to automatic selection. Posterior samples were drawn using Hamiltonian Monte Carlo (HMC). TD was used as the reference condition; intercept-derived expected counts represented TD composition, while effect estimates represented ASD-specific compositional changes. Cell types with a posterior inclusion probability ≥ 0.8 were considered potentially significant. For the reference condition, expected sample counts derived from the model intercept are used as composition scores, while for non-reference conditions, condition-specific expected sample counts from the effect estimates are used. This method provides uncertainty-aware compositional estimates with built-in correction for the interdependence of cell type proportions.

2.2.6 `cell_composition` framework

To integrate cell composition information into the CCC inference workflow, a dedicated CellComposition class was developed. This class accepts single-cell annotated data alongside a low-dimensional manifold representation and a user-defined condition label, and provides a unified interface for the three composition scoring strategies described above prior to LIANA and CCC analysis.

1. basic set up

```
import cell2cell as c2c
from cell_composition import CellComposition
seed = 888
```

2. loading the data

```
adata = sc.read_h5ad(data_folder + '/ASD_TD.h5ad')
```

```
# 3. preparation of manifold
manifold = pd.DataFrame(adata.obsm['X_pca'])
manifold.index = adata.obs_names

# 4. creation of CellComposition
cc = CellComposition(adata, context_label='diagnosis', context_order=['TD',
'ASD'], cellgroup_label='cell_type', manifold=manifold)

# 5. composition score calculation
cc.score_composition(score_type='sccoda')  # or 'meld', 'frequency'

# 6. check the result
cc.aggregate_composition_score['ASD']
```

Here, `Context` denotes the disease condition, `cell_type` denotes the annotated cell type identity, and the PCA embedding was used as the manifold representation. Composition scores were computed using the scCODA-based method to quantify cell type abundance changes across conditions.

2.3 Results

2.3.1 MELD captured condition-associated cell-state shifts

2.3.1.1 Spearman correlation and Scatter plot

Frequency scoring revealed that neural progenitor cells (NPCs) and maturing neurons were the most abundant cell types across both conditions, while truncated radial glia (tRG) showed markedly lower proportions in ASD (0.015–0.024) relative to TD (0.089–0.156).

Figure 2.1 presents scatter plot comparisons of frequency scores versus MELD scores across cell types under TD and ASD conditions. Spearman rank correlation revealed a reasonable level of agreement between the two methods in both conditions (TD: $\rho = 0.615$, $p = 0.002$; ASD: $\rho = 0.520$, $p = 0.013$), with a lower correlation in ASD indicating greater divergence between the two metrics. Cell types such as oligodendrocyte progenitor cells (OPCs) and stromal cells exhibited disproportionately high MELD scores relative to their frequency, indicating that frequency scoring alone may systematically underrepresent biologically relevant signals in ASD.

In Figure 2.1, the scatter plot appears more compressed in ASD than in TD, with frequency scores narrowing from approximately 0.02–0.16 to 0.02–0.10 and MELD scores converging from 0.35–0.95 to 0.45–0.72, suggesting a reduction in both cell type abundance variation and condition specificity in ASD. Reflecting MELD's resolution of continuous cell-state gradients, this compression may indicate a homogenization of cell composition in ASD. Specifically, the hierarchical dominance of certain progenitor populations, such as tRG, appears disrupted (Table 2.1).

Table 2.1. Frequency and MELD scores for tRG across ASD and TD conditions

Method	**Frequency**		**MELD**	
Context	**ASD**	**TD**	**ASD**	**TD**
00_Truncated_Radial_Glia	0.024	0.156	0.520	0.793
10_Truncated_Radial_Glia	0.015	0.089	0.465	0.932

Among the cell types showing the greatest reduction in TD-associated signals, the two tRG populations stood out as the only groups exhibiting concurrent, substantial decreases in both frequency and MELD scores in ASD relative to TD, with frequency reduced by 85% and 83% in

00_tRG and 10_tRG, respectively (Table 2.1, 2.2). In contrast, other cell types in this group, such as deep-layer excitatory neurons (18_Neurons_EXT_L56) and late radial glia (06_Late_Radial_Glia), showed much smaller decreases in both frequency (ΔFreq = −0.017 and −0.008, respectively) and MELD scores (ΔMELD = −0.033 and −0.028). Thus, the ASD-associated reduction was concentrated in tRG across both metrics, reflecting a concurrent loss of cell abundance and TD-associated transcriptomic condition specificity.

Table 2.2. MELD and frequency differences (ASD − TD) for cell types with the greatest TD-associated signal reduction.

	ΔMELD	**ΔFreq**
10_Truncated_Radial_Glia	−0.466	−0.074
00_Truncated_Radial_Glia	−0.273	−0.132
18_Neurons_EXT_L56	−0.033	−0.017
06_Late_Radial_Glia	−0.028	−0.008
16_Neurons_EXT_L6	−0.003	−0.003

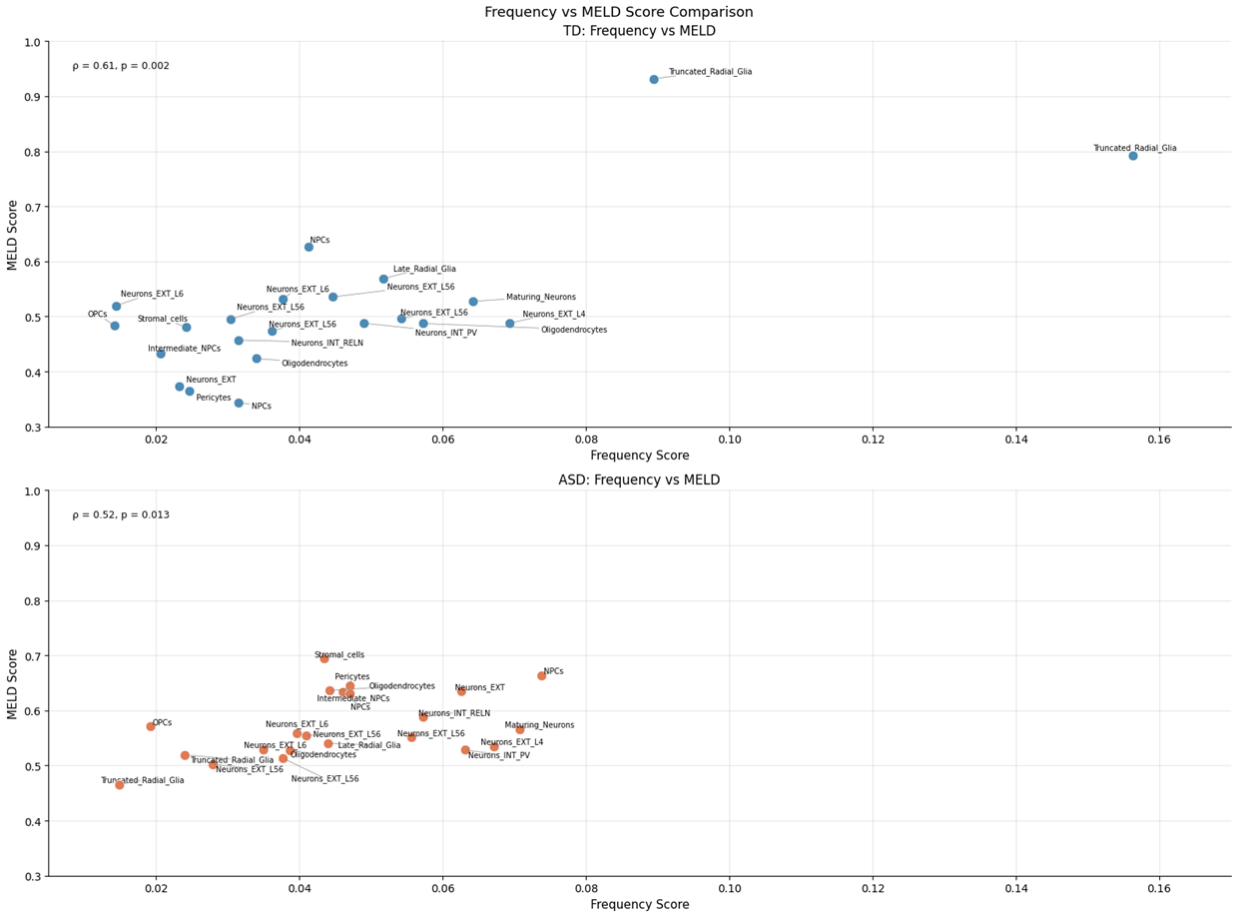


Figure 2.1. Frequency vs. MELD scores in TD (blue) and ASD (orange). Spearman correlation metrics of ρ = 0.61, p = 0.002 for TD and ρ = 0.52, p = 0.013.

2.3.1.2 Rank discrepancy comparison

To clarify the cell types underlying the rank discrepancy between the two scoring systems, rank differences (Frequency − MELD) between frequency and MELD scores were calculated for the ASD condition (Table 2.3, Figure 2.2). Here, positive rank differences denote cell types exhibiting disproportionately high MELD scores relative to their abundance. Cell types with the greatest positive rank discrepancy were stromal cells and OPCs (Δrank = +12 each), while the greatest negative discrepancies were observed in parvalbumin-expressing inhibitory

interneurons (03_Neurons_INT_PV, Δrank = −13) and layer 4 excitatory cortical neurons (02_Neurons_EXT_L4, Δrank = −12). This suggests that stromal cells and OPCs carried strong ASD-associated transcriptomic signals relative to their abundance, while the prevalence of parvalbumin-expressing inhibitory interneurons and layer 4 excitatory cortical neurons did not reflect ASD-specific transcriptomic activity.

Overall, the choice of compositional scoring strategy may meaningfully influence the interpretation of downstream CCC patterns in ASD, as frequency-based approaches would prioritize abundant but potentially ASD-nonspecific neuronal populations over lower-abundance cell types such as stromal cells and OPCs that carry relatively stronger ASD-associated transcriptomic signals by MELD.

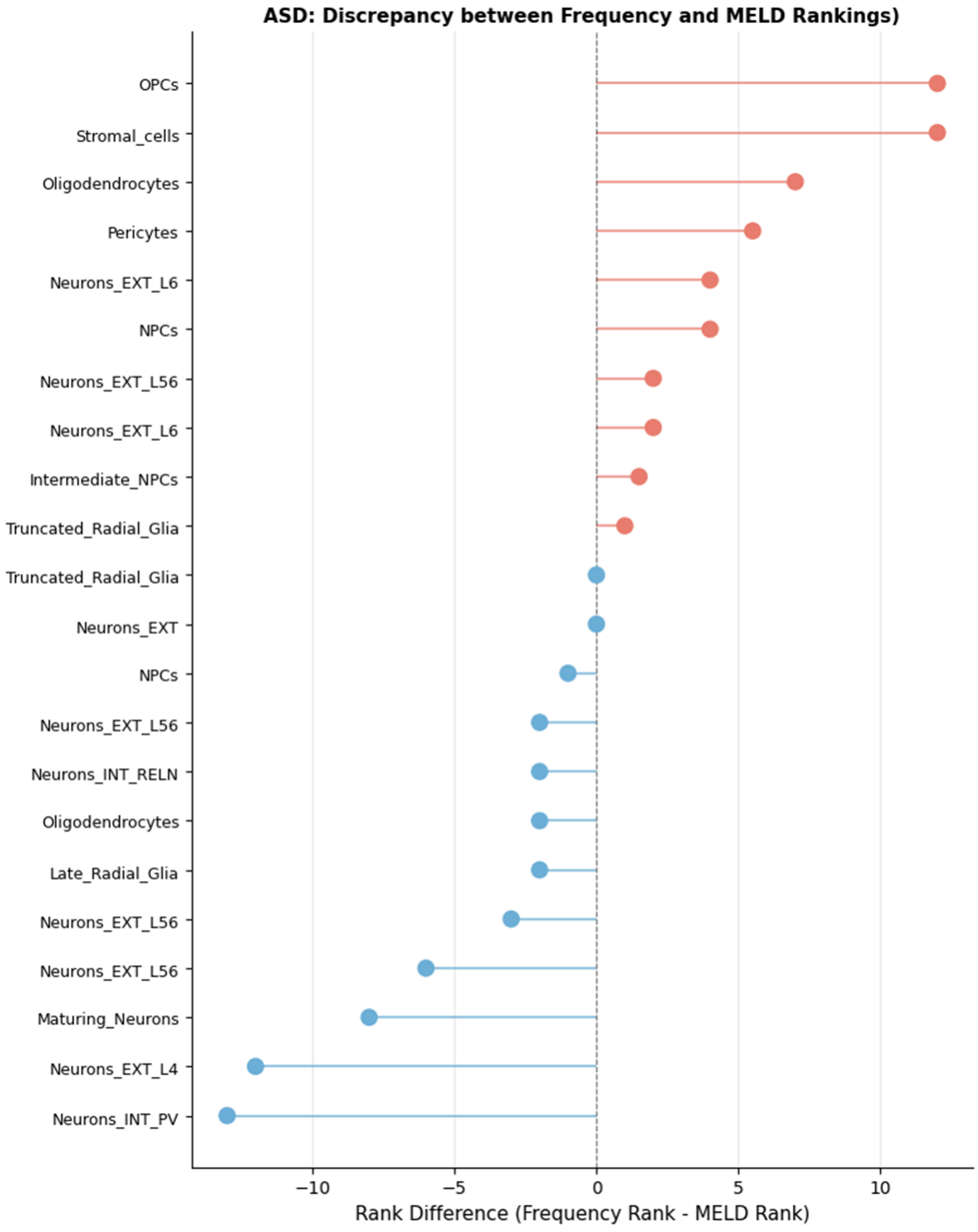


Figure 2.2. Frequency versus MELD rank discrepancies in ASD. Points are plotted by rank difference (Frequency rank - MELD rank), where positive values (red) indicate a higher MELD priority and negative values (blue) indicate a higher frequency priority.

Table 2.3. Discrepancy between frequency and MELD ranks in the ASD condition. Ranks are ordered descending (Rank 1 = highest). Positive rank differences (Frequency - MELD) denote cell types with high MELD scores relative to abundance; negative values indicate highly abundant cell types with weak ASD-association.

cell_type	freq_score	meld_score	freq_rank	meld_rank	Δrank
17_Stromal_cells	0.043	0.695	13	1	12
21_OPCs	0.019	0.571	21	9	12
12_Oligodendrocytes	0.044	0.636	11	4	7
14_Pericytes	0.047	0.646	8.5	3	5.5
11_NPCs	0.046	0.634	10	6	4
20_Neurons_EXT_L6	0.040	0.559	15	11	4
13_Neurons_EXT_L56	0.041	0.554	14	12	2
16_Neurons_EXT_L6	0.035	0.529	18	16	2
15_Intermediate_NPCs	0.047	0.631	8.5	7	1.5
00_Truncated_Radial_Glia	0.024	0.520	20	19	1
10_Truncated_Radial_Glia	0.015	0.465	22	22	0
07_Neurons_EXT	0.063	0.635	5	5	0
04_NPCs	0.074	0.664	1	2	-1
09_Neurons_INT_RELN	0.057	0.588	6	8	-2
08_Oligodendrocytes	0.039	0.527	16	18	-2
06_Late_Radial_Glia	0.044	0.541	12	14	-2
18_Neurons_EXT_L56	0.028	0.503	19	21	-2
19_Neurons_EXT_L56	0.038	0.513	17	20	-3
05_Neurons_EXT_L56	0.056	0.551	7	13	-6
01_Maturing_Neurons	0.071	0.566	2	10	-8
02_Neurons_EXT_L4	0.067	0.535	3	15	-12
03_Neurons_INT_PV	0.063	0.528	4	17	-13

2.3.2 scCODA-based compositional analysis highlighted Truncated Radial Glia

As shown in Figure 2.3, the majority of cell types exhibited inclusion probabilities below the 0.8 significance threshold, ranging from approximately 0.35 (maturing neurons) to 0.70 (excitatory cortical neurons). Notably, excitatory cortical neurons approached but did not exceed the threshold, suggesting a possible but statistically inconclusive compositional shift. tRG was the only cell type to exceed the threshold, with an inclusion probability of approximately 0.9, and was accordingly flagged as significantly altered.

This result is further contextualized by the scCODA-derived expected counts under the TD reference condition (Supplementary Table A). Compared to the most abundant cell types such as maturing neurons (expected count ≈ 1224) and neural progenitor cells (NPCs, ≈ 1069), tRG exhibited markedly lower expected counts (00_tRG: ≈ 530; 10_tRG: ≈ 356), indicating that tRG is relatively less abundant in the ASD condition. Notably, scCODA models covariate effects on cell-type compositions using a logit-normal spike-and-slab prior (Büttner et al. 2021). This prior structure tends to shrink many effects toward zero and only retain a subset as credible non-zero effects, leading to a sparse set of detected compositional changes (Biswas et al. 2022). The detection of tRG as significantly altered despite its low abundance and the model's conservative prior structure lends additional credibility to this finding.

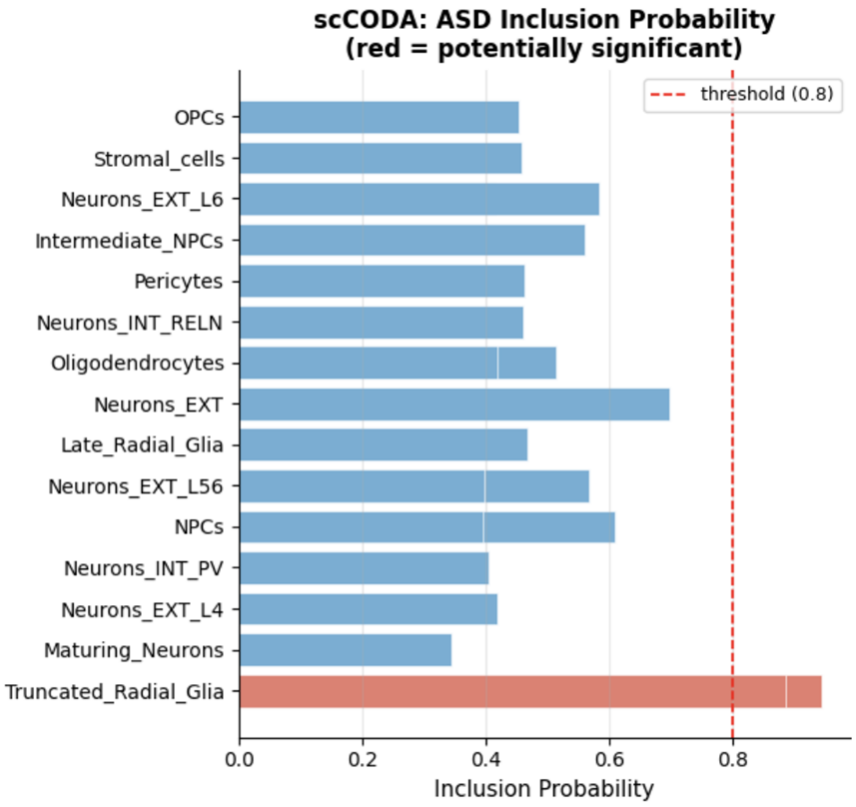


Figure 2.3. Posterior inclusion probabilities for cell types in ASD. Populations above the 0.8 significance threshold (dashed red line) are labeled in red; non-significant cell types are shown in blue.

2.4 Discussions

Integrating multiple compositional frameworks reveals complementary insights into cell-type-level perturbations in ASD cortical organoids. The divergence between frequency and MELD rankings underscores that frequency scoring may systematically miss condition-associated shifts in low-abundance but transcriptomically active populations, particularly OPCs, a glial progenitor population, and stromal cells.

The low frequency combined with high MELD scores of OPCs in the ASD condition

may reflect early oligodendrocyte lineage dysfunction. Oxidative stress in the ASD brain has been shown to impair oligodendrocyte differentiation and promote OPC proliferation, disrupting the myelination process (Galvez-Contreras et al. 2020). In line with impaired oligodendrocyte maturation, low levels of myelin basic protein have been implicated in ASD pathogenesis in mouse models, and impaired OPC maturation has been associated with reduced myelination in ASD-relevant brain regions (Graciarena et al. 2018). In this context, the high MELD score of OPCs despite their low abundance could reflect an accumulation of OPCs that fail to progress toward mature oligodendrocytes. Given that OPCs receive direct input via neuron-to-OPC synapses (Buchanan et al. 2023), transcriptomic changes in OPCs may ultimately disrupt neuron-glial communication in the developing cortex.

Stromal cells similarly exhibited disproportionately high MELD scores relative to their frequency in ASD, suggesting condition-associated transcriptomic activity in this population. This is consistent with emerging evidence that mesenchymal stromal cells affect immunoregulatory and immunosuppressive aspects that can modulate synapse formation and enhance synaptic function in ASD contexts. Furthermore, preliminary clinical trials have reported core symptom improvements following MSC infusions (Tamouza et al. 2022). However, whether the elevated ASD-associated transcriptomic signal in BCO stromal cells represents a compensatory response or a pathogenic state remains to be determined.

The tRG emerged as the most consistently altered population across all three composition scoring methods. Frequency scoring revealed a marked reduction in tRG abundance in ASD,

MELD captured a concurrent loss of TD-associated transcriptomic identity, and scCODA independently confirmed a statistically credible population-level compositional shift despite tRG's relatively low baseline abundance. Given the central role of radial glia in cortical neurogenesis and human cortical expansion, this convergence leads to the hypothesis that tRG depletion in ASD reflects a disruption of the progenitor hierarchy underlying cortical overgrowth, a proposition that warrants direct experimental validation in BCO perturbation models. Radial glia (RG) are the main progenitor cell type in the developing cortex, producing all neocortical neurons, and tRG is a progenitor subtype involved in brain expansion. For example, previous studies showed that Memo1 gene mutations disrupt the radial glial cell scaffold, resulting in neuronal misplacement and disorganized cortical layers in autism (Nakagawa et al. 2019; Meyerink et al. 2020).

Although all three approaches generated interpretable outputs, compositional modeling was retained as an exploratory analysis rather than a formal preprocessing step in the final CCC inference pipeline utilizing LIANA and Tensor-cell2cell. Several limitations of this chapter should be acknowledged. First, the sample size was small (n = 6 ASD donors in total and n = 4 per organoid age), with only one or two donors per size bin, which limits statistical power. Second, only two donors contributed organoids at both time points, and MRI was acquired at a single age per donor, precluding assessment of longitudinal CCC trajectories. Third, grouping donors into discrete size bins was a practical necessity given the sample size, but continuous modeling of MRI volume as a covariate would provide greater sensitivity to brain-size-associated CCC patterns. Fourth, TD donors were excluded from the tensor decomposition, so it remains unclear whether the identified CCC programs are specific to ASD. Finally, the BCO–

DMN association is preliminary and requires replication in a larger, balanced cohort, particularly given its divergence from previous group-level findings.

Chapter 3. MRI-Informed Single-Cell Analysis of ASD BCOs

3.1 Introduction

ASD is a highly heritable neurodevelopmental condition characterized by deficits in social communication and the presence of restricted, repetitive behaviors. As a spectrum disorder and a neurodivergent development that does not come with a definite set of conditions, ASD encompasses a wide range of symptom severities and developmental trajectories. Considering the developmental nature of the condition, it is critical to investigate the underlying biological mechanisms at early developmental stages (Girault and Piven 2020; Korkmaz 2011). Despite an estimated heritability of 80–90%, the specific genetic etiology remains unidentified in roughly 90% of ASD cases, and the molecular pathobiology underlying core social symptoms at the individual-patient level in early development is still largely unresolved (Sandin et al. 2017). This gap between high heritability and limited genetic resolution highlights the need for model systems that analyze early cortical development in a patient-specific manner.

Previous studies have shown that patient-derived BCO size correlates with ASD social symptom severity and MRI-measured cortical overgrowth in vivo. Particularly, proliferation rates of iPSC-derived NPCs predict donor MRI brain volume, functionally linking in vitro neurogenesis dynamics to brain overgrowth in ASD. These findings validate BCOs as biologically meaningful proxies of individual neurodevelopmental trajectories (Courchesne et al. 2024). However, the CCC changes that contribute to these differences remain unknown.

To characterize CCC patterns within BCOs, snRNA-seq data were analyzed in conjunction with the donor brain volume data. Brain volume information gathered via the MRI

imaging were used to categorize donors into size bins that served as the context dimension in the tensor decomposition. Overall, this chapter aims to (1) identify CCC patterns associated with MRI brain volume variation across ASD BCO donors using LIANA and Tensor-cell2cell, and (2) examine whether BCO size-associated features are connected to the patients' neuroanatomic differences in ASD. This work aims to link prenatal molecular dysregulation with postnatal brain structural outcomes in ASD.

3.2 Methods

3.2.1 Participants, snRNA-seq, and Neuroimaging Data

Participant recruitment, BCO generation, snRNA-seq, and structural MRI procedures have been described in detail previously (Courchesne et al. 2024). Briefly, snRNA-seq data from 6 ASD donors were utilized in this chapter; TD donors were excluded as the analysis focused on MRI brain volume stratification within the ASD group. Donors were classified into two clinical outcome groups following Courchesne et al. 2024: ASDgood, corresponding to the mild autism subtype characterized by improving social and language abilities and milder BCO enlargement; and ASDpoor, corresponding to the profound autism subtype characterized by persistently severe social symptoms, reduced language ability, and the greatest BCO overgrowth. MRI-derived regional brain volumes were obtained from the same cohort (N = 9 ASD toddlers) and calculated using FreeSurfer 5.3 as described by Courchesne et al. 2024. Regional gray matter volumes of the 34 left-hemisphere Desikan-Killiany cortical regions were corrected for age at scan and sex, and their mean was used as a summary measure of donor cortical volume. Samples were analyzed separately by organoid age: month-1 (n = 4: B8T5A, K6L4M, X3N7Y, E6Y3N) and

month-2 (n = 4: K6L4M, H5A4A, X3N7Y, C6G4V). Two donors (K6L4M, X3N7Y) contributed data at both time points. For two donors with ambiguous month-0 labels (C6G4V, E6Y3N), time point assignments were corrected to month-2 and month-1, respectively, based on dimensionality reduction analysis.

3.2.2 Ligand–Receptor Interaction Analysis and Tensor Decomposition

The ligand-receptor interaction analysis was performed using LIANA (Dimitrov et al. 2022) by applying the rank aggregate method with the consensus resource. Ligands and receptors were required to be expressed in at least 10% of cells within a given cell type, with a minimum of 5 cells per cell type per sample, and statistical significance was assessed using 100 permutations. Prior to tensor construction, magnitude rank scores were inverse-transformed (1 − score) so that higher values reflect stronger interactions. The resulting interaction scores were then used as input for Tensor-cell2cell, applied separately to month-1 and month-2 samples. The donor_month label (e.g., K6L4M_1, X3N7Y_2) was used as the primary sample identifier, and donors were stratified into three size bins based on the mean cortical regional volume (Table 3.1).

The optimal rank was determined via elbow analysis on reconstruction error (upper rank = 25), and tensor factorization was performed using a PyTorch backend with GPU acceleration (random_state = 888). Gene Set Enrichment Analysis (GSEA) was subsequently performed on ligand-receptor pair loadings from each factor using KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways (permutations = 999, min_size = 15). Factor context loadings were then

compared across MRI brain volume groups to identify CCC programs associated with individual differences in donor brain volume.

Table 3.1. MRI-derived Mean Cortical Regional Volume, LH (mm^3) and size bin assignments for ASD donors included in Tensor-cell2cell analysis, stratified by organoid age.

Donor ID	Time Point (Month)	Mean Cortical Regional Volume, LH (mm^3)	Size Bin
K6L4M	1	5,645	Small
B8T5A	1	6,304	Medium
X3N7Y	1	6,536	Large
E6Y3N	1	6,751	Large
K6L4M	2	5,645	Small
H5A4A	2	5,802	Medium
X3N7Y	2	6,536	Large
C6G4V	2	6,921	Large

3.2.3 Yeo Network Parcellation

MRI regional volumes derived from FreeSurfer 5.3 were mapped to the Yeo 7-network parcellation using the Desikan-Killiany atlas as an approximation (Buckner et al. 2011; Yeo et al. 2011). Each of the seven canonical networks (visual, somatomotor, dorsal attention, ventral attention, limbic, frontoparietal control, and default mode) was represented by the mean volume of corresponding cortical regions in the left hemisphere. Spearman rank correlations were computed between BCO size and each network's mean volume across donors with both BCO and MRI data available (n = 4 for month-1).

3.3 Results

3.3.1 Tensor-cell2cell with Brain MRI information

3.3.1.1 Month-1 Tensor-cell2cell Factors

Among the ten factors, only Factor 2 achieved statistical significance, showing enrichment in the large brain volume group (Figure 3.2). In this factor, excitatory neurons and oligodendrocytes drove signaling toward OPCs, maturing neurons, and pericytes, with FAM3C–LAMP1, NCAM1–PTPRZ1, and ADAM10–EPHA3 as the top ligand-receptor pairs (Supplementary Table B). In contrast, Factor 1 was enriched in the small brain volume group and was characterized by NPC-to-interneuron signaling via NTN1–DCC, SLIT2–DCC, and SLIT2–ROBO, which are key components of axon guidance pathways. Although they did not achieve statistical significance, Factors 9 and 10 exhibited higher loadings in the large brain volume group. Factor 9 showed reciprocal communication between interneurons and excitatory neurons, featuring NLGN4X–NRXN1 and NXPH1–NRXN1. Factor 10 was driven by oligodendrocytes and stromal cells signaling to deep-layer neurons and NPCs via MDK–LRP2, MDK–PTPRZ1, and PTN–PTPRZ1, reflecting a trophic signaling axis centered on the MDK–PTPRZ1 interaction.

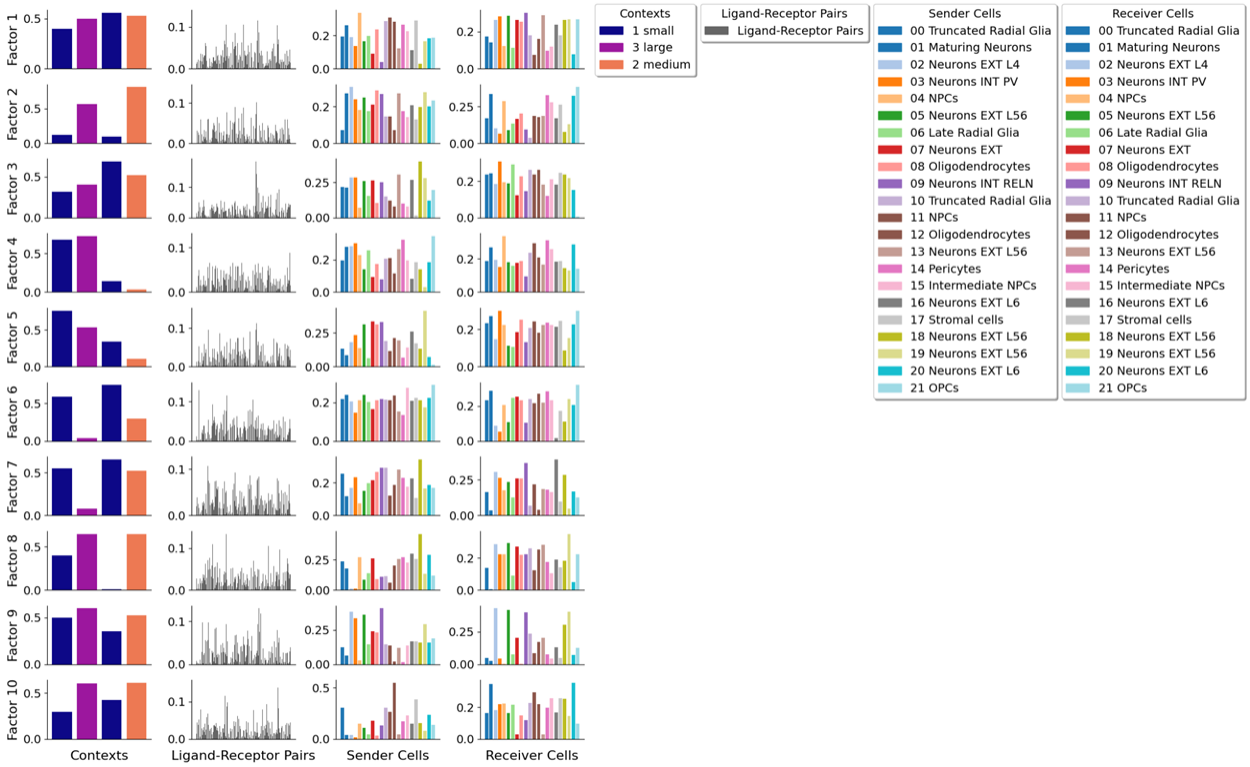


Figure 3.1. Tensor-cell2cell decomposition of month-1 BCO. Factors obtained after decomposing the 4D-communication tensor from snRNA-seq data of ASD BCOs stratified by donor MRI cortical volume. Ten factors were selected based on elbow analysis (Supplementary Figure A). n = 4 donors; MRI brain volume bins: small, medium, large.

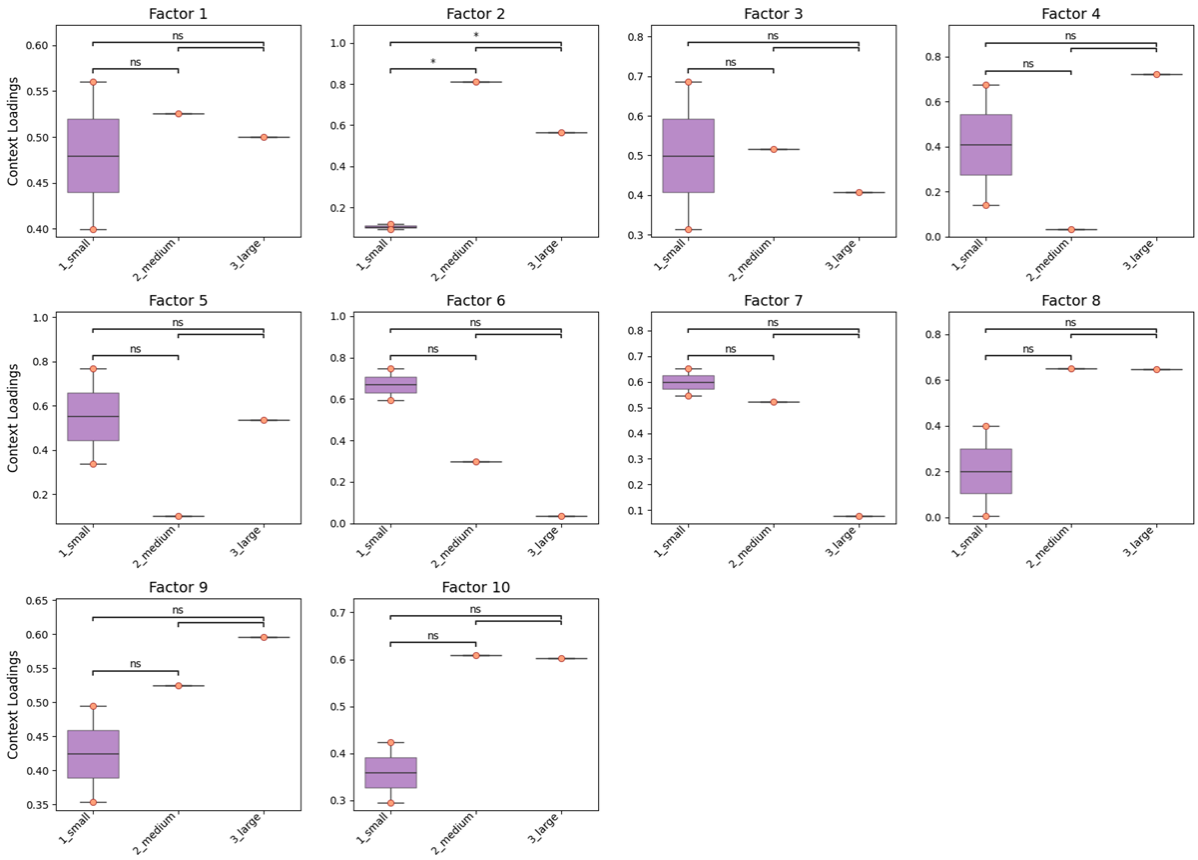


Figure 3.2. Box plots showing the distribution of factor loadings for all ten factors across MRI brain volume groups (small, medium, large). Statistical comparisons are shown between groups (ns, not significant; * $p < 0.05$).

In GSEA analysis (Figure 3.3), Factor 1 showed significant positive enrichment for axon guidance and cell adhesion molecules (CAMs), consistent with its NPC-to-interneuron signaling profile via NTN1–DCC and SLIT2–ROBO1. The positive normalized enrichment score (NES) indicates that guidance- and adhesion-related LR pairs are concentrated among the top-ranked LR pairs of Factor 1, suggesting that this factor captures an axon guidance and neuronal wiring program. On the other hand, Factor 2 showed negative enrichment for the Notch signaling in the large brain volume group. Given the role of Notch signaling in maintaining NPC self-renewal and regulating the timing of neuronal differentiation (Roese-Koerner et al. 2017), reduced Notch activity may reflect a shift toward increased or premature differentiation in neurodevelopmental

trajectories. Factor 9 showed significant positive enrichment for CAMs, consistent with NLGN4X–NRXN1 and NXPH1–NRXN1 being its top contributors. Additionally, Factor 10 showed significant positive enrichment for axon guidance as the only enriched pathway.

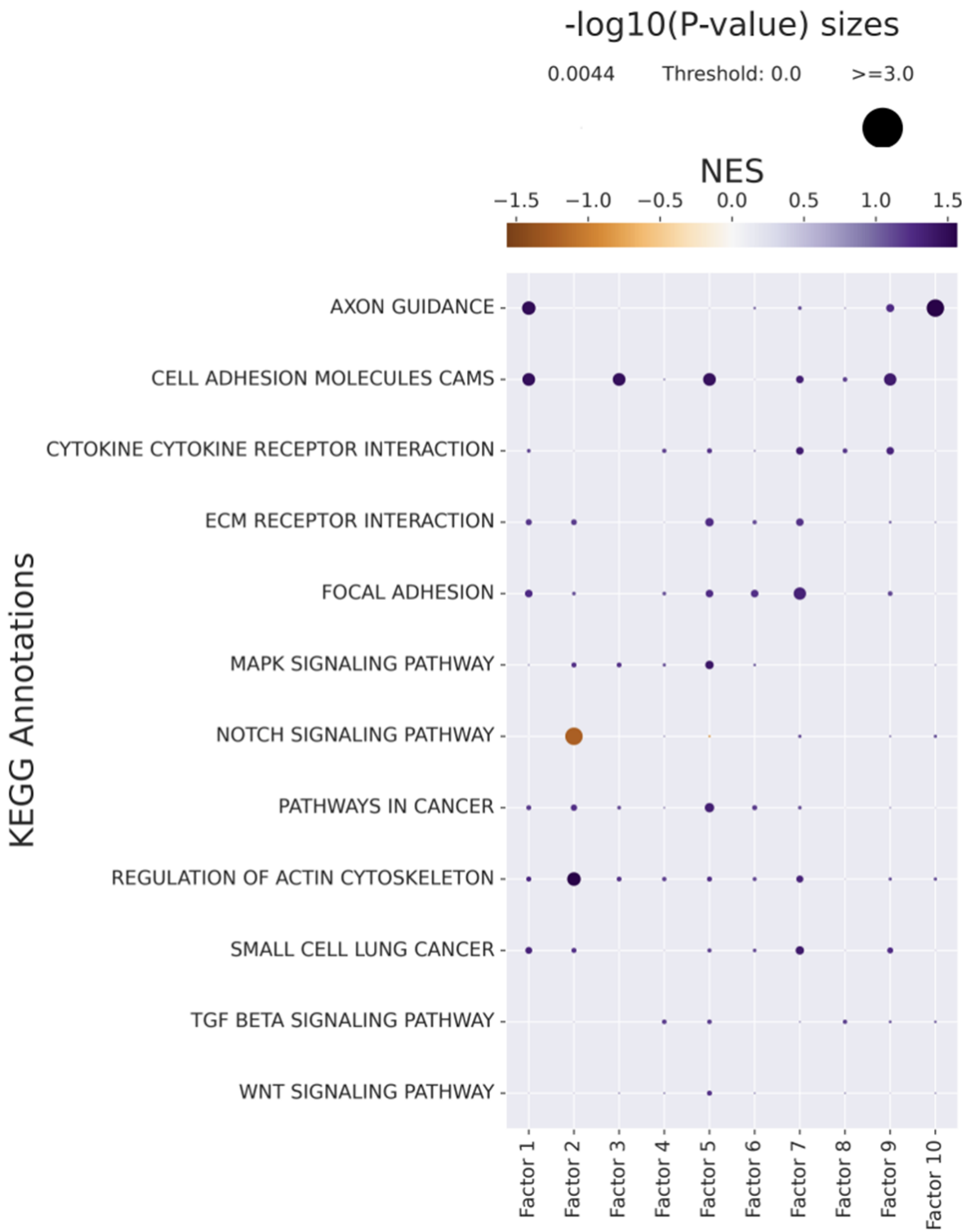


Figure 3.3. Dot plot showing GSEA results for KEGG pathways across Tensor-cell2cell factors derived from month-1 BCO snRNA-seq data. Dot size indicates statistical significance (−log10 p-value; threshold = 1.301, corresponding to $p < 0.05$). Dot color indicates NES, with purple representing positive enrichment and orange representing negative enrichment.

3.3.1.2 Month-2 Tensor-cell2cell Factors

In the month-2 data, Factor 11 was enriched in the large brain volume group, where tRG and deep-layer excitatory neurons acted as the primary senders and interneurons as the receivers, with NLGN2–NRXN1 and NLGN4X–NRXN1 as top pairs. While month-1 Factor 9 was characterized by reciprocal interneuron-excitatory neuron communication via NLGN4X–NRXN1 and NXPH1–NRXN1, Factor 11 incorporated tRG as an additional sender. Similarly, Factor 1 featured NLGN4X–NRXN1 and NXPH1–NRXN1 as top pairs, but with interneurons and tRG acting as senders and excitatory neurons as receivers. Compared to month-1 Factor 9, Factors 1 and 11 collectively reveal a more extensive expansion of the neurexin-centered (NLGN/NXPH–NRXN) network at month 2, characterized by the addition of tRG as an extra sender and the coordinated contribution of multiple pairs including NLGN4X–NRXN1, NXPH1–NRXN1, and NLGN2–NRXN1.

Additionally, Factors 14 and 16 both exhibited prominent MDK-related signaling. Factor 14 involved late radial glia and stromal cells sending signals to progenitor cells, specifically NPCs and OPCs, through LAMA2–ITGAV_ITGB8 and MDK–LRP2. Factor 16 featured NPCs sending signals to maturing neurons through MDK–PTPRZ1 and DLK1–NOTCH1. The MDK–PTPRZ1 pair appeared as a top ligand-receptor interaction in both month-1 Factor 10 and month-2 Factor 16, but with different sender populations: oligodendrocytes, stromal cells, and tRG at month 1, and NPCs at month 2. This suggests that multiple cell types engage in MDK–PTPRZ1 signaling across organoid development, with higher factor loadings in the larger brain volume group.

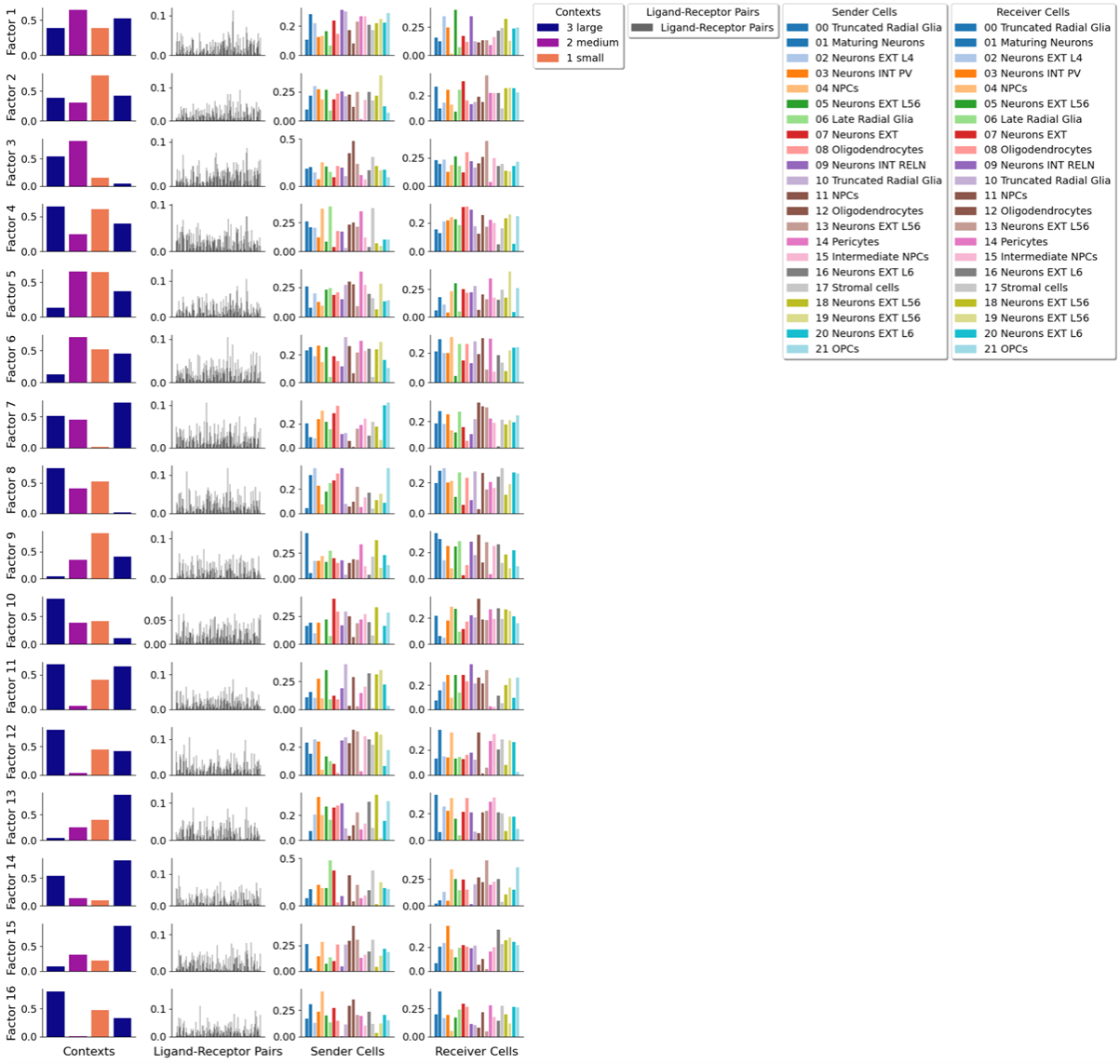


Figure 3.4. Tensor-cell2cell decomposition of month-2 BCO. Factors obtained after decomposing the 4D-communication tensor from snRNA-seq data of ASD BCOs stratified by donor MRI cortical volume. Sixteen factors were selected based on elbow analysis (Supplementary Figure A). n = 4 donors; MRI brain volume bins: small, medium, large.

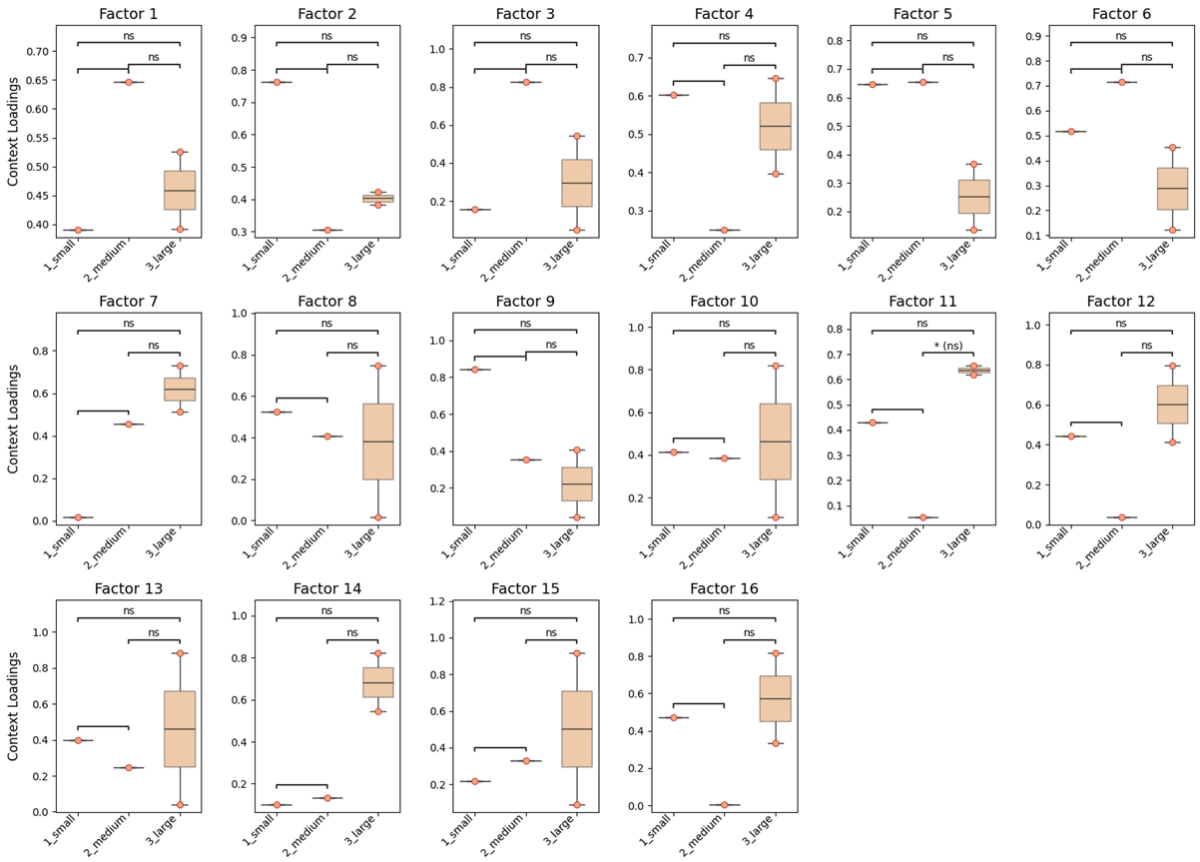


Figure 3.5. Box plots showing the distribution of factor loadings for all sixteen factors across MRI brain volume groups (small, medium, large). Each dot represents an individual donor. Statistical comparisons are shown between groups (ns, not significant; * $p < 0.05$).

The WNT Signaling Pathway was the most recurrently enriched pathway, appearing in Factors 3, 8, 10, and 15. Its broad recurrence across multiple factors and its absence from the month-1 GSEA results suggest that WNT-mediated regulation of progenitor proliferation and differentiation may be more prominent at month 2 than at month 1 in ASD BCOs.

For the GSEA of month-2 factors (Figure 3.6), CAMs were consistently enriched across Factors 1 and 11, matching the dominant neurexin-centered pairs identified, including the neuroligin–neurexin pairs NLGN4X–NRXN1 and NLGN2–NRXN1 and the neurexophilin–neurexin pair NXPH1–NRXN1. The WNT signaling pathway was the most recurrently enriched

program, appearing in Factors 3, 8, 10, and 15. Its recurrence across multiple factors, combined with its absence from the month-1 GSEA results, suggests that WNT-mediated regulation of progenitor proliferation and differentiation may become more prominent at month 2 compared to month 1 in ASD BCOs.

Factor 7 showed enrichment for the melanoma pathway, with FGF13–FGFR2 as the top LR pair driven by OPCs and oligodendrocytes as the primary senders. Since FGFR2 is critical for the proliferation, self-renewal, and differentiation of cortical neural stem cells during corticogenesis (Stevens et al. 2010), this enrichment may reflect FGF-mediated regulation of progenitor activity rather than melanocyte-specific biology.

Factor 14 displayed enrichment for dilated cardiomyopathy, with late radial glia and stromal cells as the primary senders. The dominant integrin-ECM pairs, LAMA2–ITGAV_ITGB8 and COL4A2–ITGAV_ITGB8, share cytoskeletal anchoring and mechanosignaling components with the dilated cardiomyopathy pathway and neurodevelopment. Neuroglial expression of ITGAV and ITGB8 has been shown to be required for brain angiogenesis and vascular stability (Schwartz 2010; Tata et al. 2015). As with Factor 7, this enrichment could reflect convergent CCC during cortical development rather than cardiac pathology.

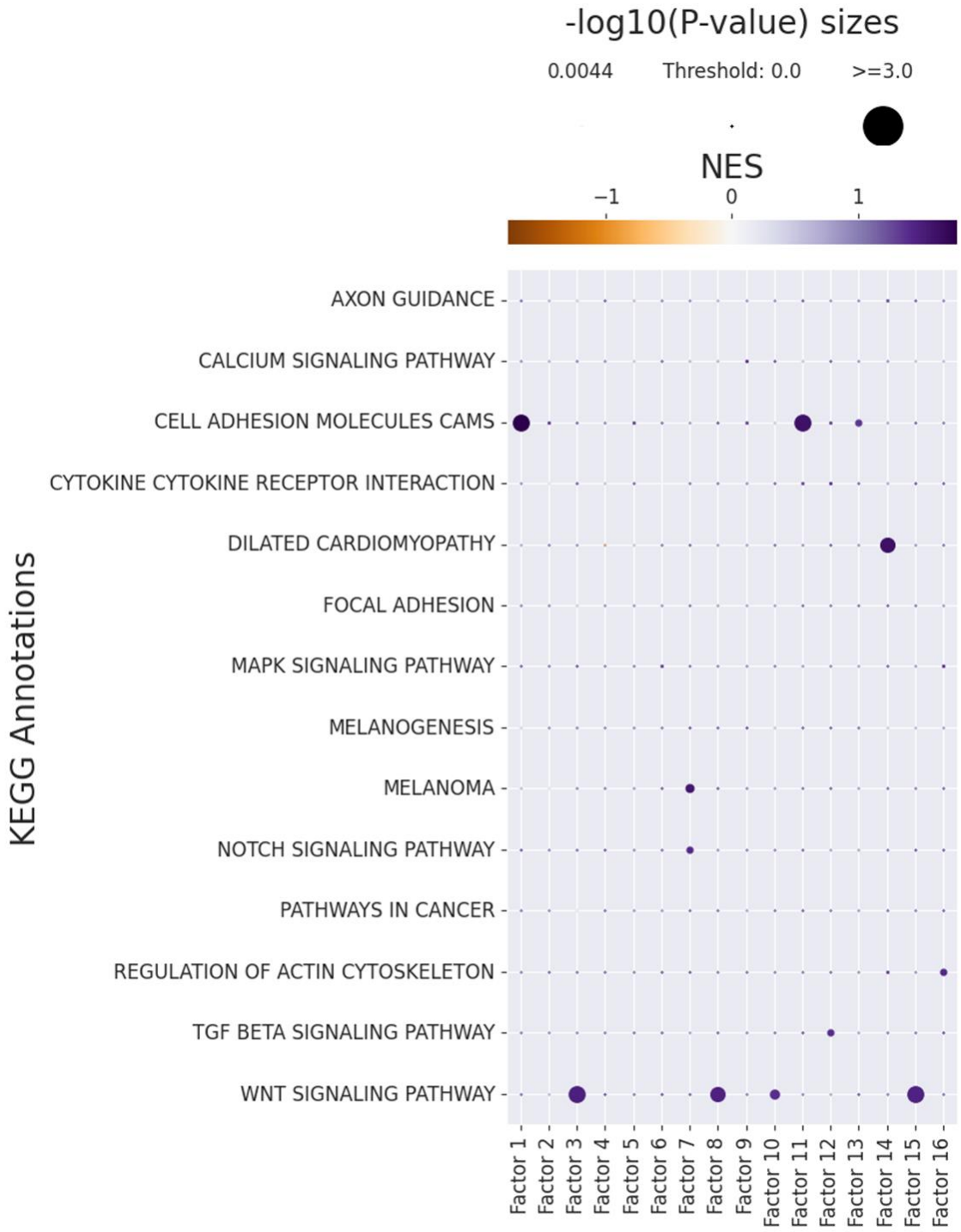


Figure 3.6. Dot plot showing GSEA results for KEGG pathways across Tensor-cell2cell factors derived from month-2 BCO snRNA-seq data. Dot size indicates statistical significance (−log10 p-value; threshold = 1.301, corresponding to $p < 0.05$). Dot color indicates normalized enrichment score (NES), with purple representing positive enrichment and orange representing negative enrichment.

3.3.2 BCO Size and MRI Brain Volume Correlation

Among ASD donors with both BCO measurements and MRI data available for month-1 (n = 4), BCO size was correlated with MRI regional volumes mapped to the Yeo 7-network parcellation (Table 3.2, Figure 3.7). Six of the seven networks showed negative rank correlations with BCO size (ρ range: −1.00 to −0.20), while the ventral attention network showed no correlation (ρ = 0.00). The default mode network (DMN) showed the strongest inverse rank association with BCO size (ρ = −1.00), with larger BCOs consistently associated with smaller DMN volumes across all four donors; however, this association did not reach statistical significance in an exact two-sided permutation test (p = 0.083, n = 4). The frontoparietal, visual, dorsal attention, and limbic networks each showed moderate negative trends (ρ = −0.40), while the somatomotor network showed a weaker negative trend (ρ = −0.20). These preliminary findings suggest a directional link between in vitro organoid growth dynamics and cortical network volumes in ASD, providing an initial cross-modal bridge between prenatal neurodevelopmental signals and postnatal brain structural outcomes.

Table 3.2. MRI-derived regional network volumes and morphometric ranks for ASD donors included in the cross-modal correlation analysis (n = 4, month-1), based on the Yeo 7-network parcellation.

Donor	BCO Size (μm²)	BCO Rank	DMN Volume (mm³)	DMN Rank	Prognosis
K6L4M	1108.9	1	3788.4	4	ASDgood
E6Y3N	1064.3	2	4332.9	3	ASDgood
X3N7Y	889.9	3	4413.7	2	ASDpoor
B8T5A	760.4	4	4504.8	1	ASDgood

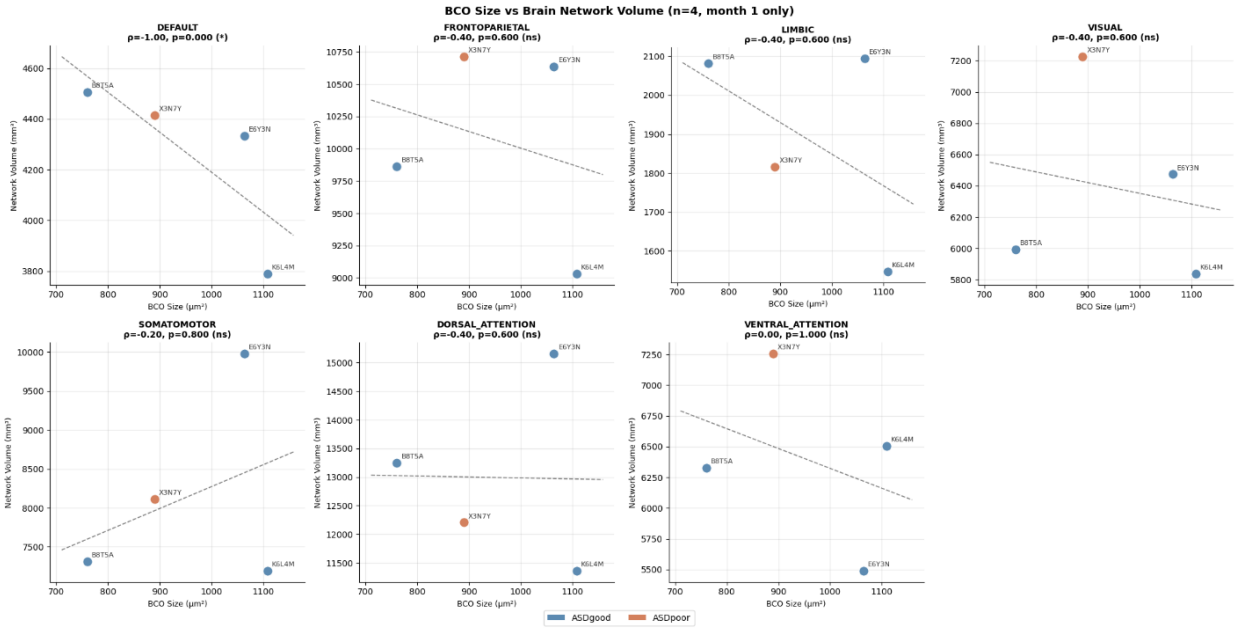


Figure 3.7. Scatter plots showing Spearman rank correlations between BCO size (μm²) and MRI-derived regional network volumes (mm³). Donors are colored by clinical prognosis (ASDgood, blue; ASDpoor, orange). Dashed lines indicate linear regression fits. Spearman ρ and asymptotic p-values are annotated for each network; for the DMN, the exact two-sided p-value is 0.083 (n = 4)

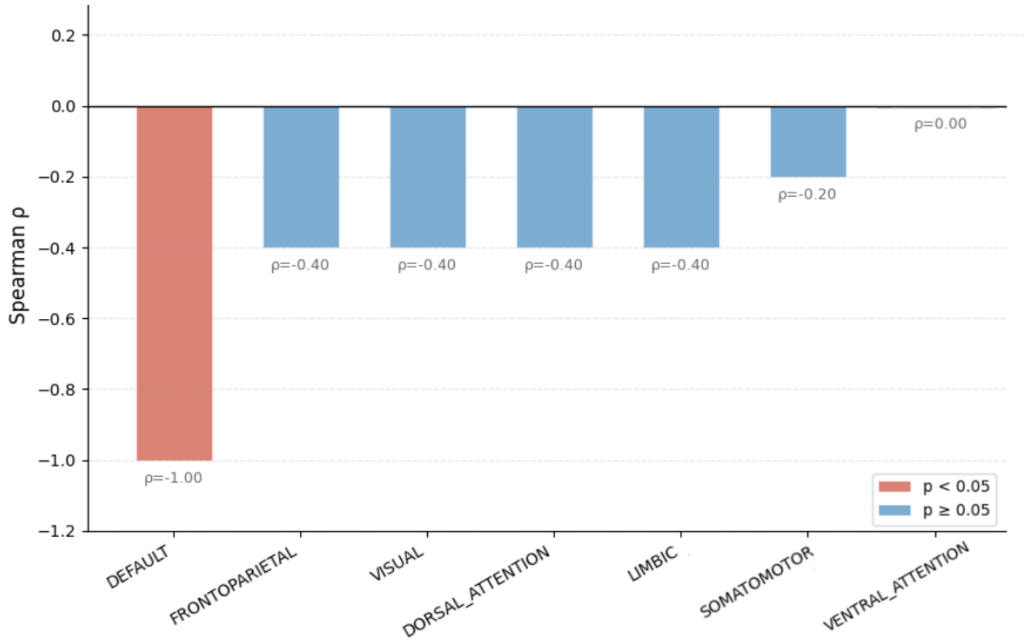


Figure 3.8. Bar plot summarizing Spearman ρ values between BCO size and MRI-derived regional network volumes across all seven Yeo networks (n = 4, month-1).

3.4 Discussions

Overall, three CCC patterns stood out in this chapter: neurexin–neuroligin signaling represented by NLGN–NRXN pairs, recurrent MDK signaling through MDK–PTPRZ1 and MDK–LRP2, and axon guidance signaling (month-1 Factors 1 and 10). The recurrence of NLGN–NRXN pairs across multiple factors at both organoid ages is consistent with the interaction between presynaptic neurexins and postsynaptic neuroligins being one of the most affected pathways in ASD. In vivo, homozygous knock-in mice carrying the NLGN2 R215H missense mutation exhibited growth retardation, anxiety-like behavior, and impaired spatial learning and memory (Chen et al. 2017). The prominent recurrence of NLGN4X among the top LR pairs coincides with the finding that NLGN4 carries the largest number of ASD-associated mutations among the neuroligin genes (Trobiani et al. 2020). This convergence of genetic evidence and the present CCC analysis supports the importance of neuroligins in the neuronal network differences underlying ASD pathophysiology.

The MDK axis is a second recurrent feature. MDK–PTPRZ1 appeared among the top LR pairs of month-1 Factor 10 and month-2 Factor 16, and MDK–LRP2 in month-1 Factor 10 and month-2 Factor 14, spanning glial/stromal-to-neuron and NPC-to-neuron signaling contexts. Midkine (MDK) is a heparin-binding growth factor involved in neurogenesis, neurite outgrowth, and cell migration, and serum MDK concentrations have been reported to be significantly higher in individuals with ASD than in controls (Esnafoglu and Cirrik 2018). MDK–PTPRZ1 signaling has also been shown to regulate neural plate convergence and midline formation during early neural development (Le et al. 2025). Together, these observations suggest that MDK-mediated trophic signaling may be an active component of CCC in ASD BCOs.

Axon guidance was the third recurrent program, showing positive enrichment in month-1 Factors 1 and 10, driven by NTN1–DCC, SLIT2–DCC, and SLIT2–ROBO1 in Factor 1 and by SLIT2–PLXNA1 and SLIT2–APP in Factor 10. Dysregulation of axonal growth and guidance has been proposed as a contributor to ASD etiology (McFadden and Minshew 2013), and histopathological analyses have reported alterations in axon density, size heterogeneity, and morphology in the cortex of individuals with ASD (Yazdanbakhsh et al. 2025). The presence of guidance programs in early BCOs suggests that wiring-related signaling alterations may already be detectable at embryonic-like stages, although their association with specific brain volume groups requires confirmation in larger cohorts.

Courchesne et al. (2024) demonstrated that BCO size correlated with donor social symptom severity, establishing BCO size as an index of cortical overgrowth tendency in ASD. In this context, the inverse rank association between BCO size and DMN volume observed here ($\rho = -1.00$, $n = 4$) suggests that donors whose organoids showed a greater overgrowth tendency had smaller DMN volumes. However, this association did not reach statistical significance (exact two-sided $p = 0.083$), which is the minimum attainable p-value with four donors, and should therefore be regarded as preliminary. Notably, previous studies reported the opposite tendency, with autistic preschoolers exhibiting larger DMN volumes than controls (Zielinski et al. 2012). This discrepancy may reflect differences between individual-level and group-level analyses, differences in participant age, and the approximate Desikan-Killiany-based mapping of the DMN used here. Regardless, DMN dysfunction has been consistently linked to impairments in self-referential processing and theory of mind in ASD (Padmanabhan et al. 2017). Taken together

with the tRG reduction observed in Chapter 2, these observations raise the possibility that altered progenitor dynamics contribute to both organoid growth differences and postnatal DMN structural variation, a hypothesis that requires validation in larger cohorts.

Several limitations of this chapter should be acknowledged. First, the sample size was small (n = 6 ASD donors in total and n = 4 per organoid age), with only one or two donors per size bin, which limits statistical power. Second, only two donors contributed organoids at both time points, and MRI was acquired at a single age per donor, precluding assessment of longitudinal CCC trajectories. Third, grouping donors into discrete size bins was a practical necessity given the sample size, but continuous modeling of MRI volume as a covariate would provide greater sensitivity to brain-size-associated CCC patterns. Fourth, TD donors were excluded from the CCC analysis, so it remains unclear whether the identified CCC programs are specific to ASD. Finally, the BCO–DMN association is preliminary and requires replication in a larger, balanced cohort, particularly given its divergence from previous group-level findings.

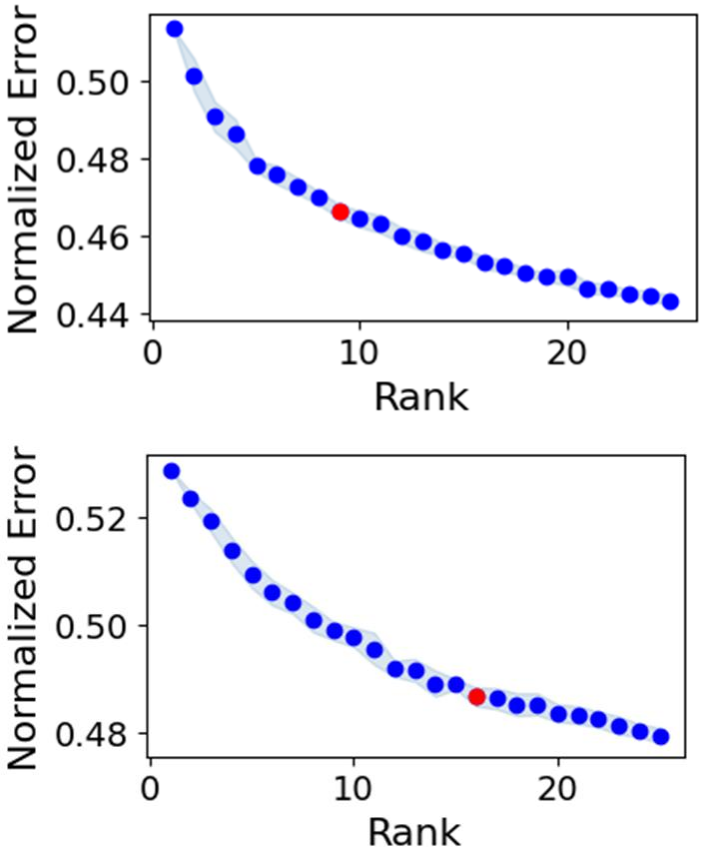
Conclusion

Several directions emerge from this work. Most immediately, replication in larger cohorts with continuous MRI covariate modeling and inclusion of TD donors in the tensor decomposition framework would address the core statistical and comparative limitations of the current analysis. Beyond replication, functional validation of the MDK–PTPRZ1 trophic axis and NLGN–NRXN synaptic adhesion programs through perturbation experiments in BCOs would establish causal rather than associative links between these CCC patterns and ASD-associated brain overgrowth. Finally, extending this pipeline to spatial transcriptomics data would enable the spatial localization of the CCC patterns identified here, providing laminar and regional resolution on where ASD-associated signaling dysregulation is most pronounced in the developing cortex.

Supplementary Materials

Supplementary Table A. Cell type composition scores across three methods (frequency, MELD, and scCODA) for all annotated cell types in ASD and TD brain cortical organoids.

method	frequency		meld		sccoda
context	ASD	TD	ASD	TD	ASD (expected count)
00_Truncated_Radial_Glia	0.024	0.156	0.520	0.793	529.860
01_Maturing_Neurons	0.071	0.064	0.566	0.528	1224.898
02_Neurons_EXT_L4	0.067	0.069	0.535	0.488	1217.570
03_Neurons_INT_PV	0.063	0.049	0.529	0.488	1027.221
04_NPCs	0.074	0.032	0.664	0.344	1069.143
05_Neurons_EXT_L56	0.056	0.054	0.551	0.496	1023.121
06_Late_Radial_Glia	0.044	0.052	0.541	0.569	827.668
07_Neurons_EXT	0.063	0.023	0.635	0.374	861.446
08_Oligodendrocytes	0.039	0.057	0.527	0.487	756.432
09_Neurons_INT_RELN	0.057	0.032	0.588	0.457	845.233
10_Truncated_Radial_Glia	0.015	0.089	0.465	0.932	355.887
11_NPCs	0.046	0.041	0.634	0.627	792.041
12_Oligodendrocytes	0.044	0.034	0.636	0.424	724.595
13_Neurons_EXT_L56	0.041	0.036	0.555	0.474	701.073
14_Pericytes	0.047	0.025	0.646	0.364	679.674
15_Intermediate_NPCs	0.047	0.021	0.631	0.432	655.641
16_Neurons_EXT_L6	0.035	0.038	0.529	0.532	643.301
17_Stromal_cells	0.044	0.024	0.695	0.481	642.016
18_Neurons_EXT_L56	0.028	0.045	0.503	0.536	568.848
19_Neurons_EXT_L56	0.038	0.030	0.513	0.495	624.913
20_Neurons_EXT_L6	0.040	0.014	0.559	0.519	520.408
21_OPCs	0.019	0.014	0.571	0.483	310.013



Supplementary Figure A. Elbow plots for Month-1 and Month-2 Tensor-cell2cell analysis

Supplementary Table B. Cell-type sender/receiver identities and top ligand-receptor pairs for notable Tensor-cell2cell factors derived from month-1 BCO snRNA-seq data

Factor 1						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	04_NPCs	0.339	09_Neurons_INT_RELN	0.301	NTN1–DCC	0.134
2	11_NPCs	0.309	13_Neurons_EXT_L56	0.286	SLIT2–DCC	0.128
3	17_Stromal_cells	0.293	05_Neurons_EXT_L56	0.284	IL1RAPL1–PTPRD	0.128
4					SLIT2–ROBO1	0.124
5					SLIT2–APP	0.123
Factor 2						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	02_Neurons_EXT_L4	0.307	21_OPCs	0.383	FAM3C–LAMP1	0.140
2	08_Oligodendrocytes	0.288	01_Maturing_Neurons	0.334	SPTAN1–PTPRA	0.108
3	19_Neurons_EXT_L56	0.278	14_Pericytes	0.328	NCAM1–GFRA1	0.105
4					ADAM10–EPHA3	0.103
5					NCAM1–PTPRZ1	0.102

Supplementary Table B (cont.)

Factor 3						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	18_Neurons_EXT_L56	0.396	03_Neurons_INT_PV	0.308	NCAM1–FGFR2	0.183
2	13_Neurons_EXT_L56	0.303	06_Late_Radial_Glia	0.294	NCAM1–PTPRA	0.169
3	03_Neurons_INT_PV	0.285	10_Truncated_Radial_Glia	0.264	NCAM1–CACNA1C	0.158
4					NCAM1–NCAM2	0.156
5					NCAM1–PTPRZ1	0.142
Factor 4						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	21_OPCs	0.346	04_NPCs	0.337	GDF7–BMPR1A_BMPR2	0.126
2	14_Pericytes	0.324	14_Pericytes	0.312	ADAM17–ERBB4	0.124
3	03_Neurons_INT_PV	0.302	11_NPCs	0.294	VEGFA–ITGAV	0.117
4					NTRK3–PTPRF	0.117
5					BMP7–ACVR2B_BMPR1A	0.107

Supplementary Table B (cont.)

Factor 5						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	19_Neurons_EXT_L56	0.416	21_OPCs	0.302	RIMS2–ABCA1	0.145
2	07_Neurons_EXT	0.333	03_Neurons_INT_PV	0.299	NCAM1–PTPRZ1	0.113
3	09_Neurons_INT_RELN	0.331	01_Maturing_Neurons	0.271	NRXN1–NLGN1	0.110
4					FAM3C–CLEC2D	0.108
5					COL1A2–ITGA2B	0.106
Factor 6						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	21_OPCs	0.291	21_OPCs	0.321	LMAN1–MCFD2	0.142
2	15_Intermediate_NPCs	0.276	01_Maturing_Neurons	0.287	ADAM17–ERBB4	0.140
3	01_Maturing_Neurons	0.239	14_Pericytes	0.284	ADAM10–EPHA3	0.129
4					ADAM9–ITGAV	0.128
5					EFNA5–EPHA3	0.115

Supplementary Table B (cont.)

Factor 7						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	18_Neurons_EXT_L56	0.343	16_Neurons_EXT_L6	0.393	NXPH1–NRXN1	0.121
2	10_Truncated_Radial_Glia	0.292	09_Neurons_INT_RELN	0.368	EGF–LDLR	0.120
3	09_Neurons_INT_RELN	0.291	02_Neurons_EXT_L4	0.304	BMP7–ACVR1_ACVR2B	0.107
4					DLL1–NOTCH2	0.106
5					FN1–ITGA5_ITGB1	0.097
Factor 8						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	18_Neurons_EXT_L56	0.462	19_Neurons_EXT_L56	0.351	DRAXIN–DCC	0.134
2	16_Neurons_EXT_L6	0.298	05_Neurons_EXT_L56	0.292	SEMA4G–PLXNB2	0.130
3	20_Neurons_EXT_L6	0.287	02_Neurons_EXT_L4	0.286	RSPO2–ZNRF3	0.114
4					CNTN2–NRCAM	0.110
5					CALM3–KCNQ1	0.110

Supplementary Table B (cont.)

Factor 9						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	09_Neurons_INT_RELN	0.406	02_Neurons_EXT_L4	0.427	NLGN4X–NRXN1	0.130
2	02_Neurons_EXT_L4	0.381	05_Neurons_EXT_L56	0.416	NCAM1–CACNA1C	0.122
3	05_Neurons_EXT_L56	0.358	19_Neurons_EXT_L56	0.403	NRG3–ERBB2_ERBB4	0.119
4					NXPH1–NRXN1	0.117
5					NCAM1–NCAM2	0.099
Factor 10						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	12_Oligodendrocytes	0.548	20_Neurons_EXT_L6	0.355	MDK–LRP2	0.150
2	17_Stromal_cells	0.384	01_Maturing_Neurons	0.346	MDK–PTPRZ1	0.148
3	00_Truncated_Radial_Glia	0.306	11_NPCs	0.297	SLIT2–PLXNA1	0.138
4					PTN–PTPRZ1	0.135
5					SLIT2–APP	0.134

Supplementary Table C. Cell-type sender/receiver identities and top ligand-receptor pairs for notable Tensor-cell2cell factors derived from month-2 BCO snRNA-seq data

Factor 1						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	09_Neurons_INT_RELN	0.308	05_Neurons_EXT_L56	0.402	NLGN4X–NRXN1	0.114
2	10_Truncated_Radial_Glia	0.297	02_Neurons_EXT_L4	0.348	NXPH1–NRXN1	0.113
3	21_OPCs	0.287	09_Neurons_INT_RELN	0.347	NLGN1–NRXN1	0.113
4					NRXN3–NLGN1	0.099
5					NCAM1–NCAM2	0.089
Factor 2						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	19_Neurons_EXT_L56	0.390	13_Neurons_EXT_L56	0.359	RIMS2–ABCA1	0.137
2	02_Neurons_EXT_L4	0.301	07_Neurons_EXT	0.314	NCAM1–ROBO3	0.125
3	03_Neurons_INT_PV	0.271	00_Truncated_Radial_Glia	0.271	IL1RAPL1–PTPRS	0.118
4					NRG1–NETO2	0.111
5					NRG1–ADGRL1	0.103
Factor 3						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	12_Oligodendrocytes	0.481	13_Neurons_EXT_L56	0.402	ADAM17–ERBB4	0.102
2	11_NPCs	0.351	08_Oligodendrocytes	0.300	SHANK1–SSTR2	0.093
3	17_Stromal_cells	0.312	05_Neurons_EXT_L56	0.265	NLGN1–NRXN3	0.090
4					SLIT3–ROBO1	0.087
5					TGFB2–ENG	0.082

Supplementary Table C (cont.)

Factor 4						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	06_Late_Radial_Glia	0.392	08_Oligodendrocytes	0.316	BMP5–PTPRD	0.097
2	17_Stromal_cells	0.375	07_Neurons_EXT	0.309	WNT3–FZD6_LRP5	0.093
3	04_NPCs	0.371	09_Neurons_INT_RELN	0.293	DRAXIN–DCC	0.092
4					NTN1–DCC	0.089
5					LAMC1–ITGA7_ITGB1	0.088
Factor 5						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	14_Pericytes	0.383	19_Neurons_EXT_L56	0.407	DSCAM–DCC	0.126
2	11_NPCs	0.298	14_Pericytes	0.346	SLIT2–DCC	0.106
3	19_Neurons_EXT_L56	0.281	05_Neurons_EXT_L56	0.302	APP–DCC	0.100
4					TNXB–ITGA9_ITGB1	0.100
5					SLIT2–ROBO2	0.099
Factor 6						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	10_Truncated_Radial_Glia	0.328	04_NPCs	0.307	NCAM1–PTPRZ1	0.104
2	14_Pericytes	0.301	12_Oligodendrocytes	0.300	NTN1–NEO1	0.104
3	19_Neurons_EXT_L56	0.295	14_Pericytes	0.295	NRXN1–NLGN1	0.095
4					ADAM17–ERBB4	0.094
5					NTNG1–LRRC4C	0.087

Supplementary Table C (cont.)

Factor 7						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	21_OPCs	0.378	11_NPCs	0.341	FGF13–FGFR2	0.106
2	20_Neurons_EXT_L6	0.356	12_Oligodendrocytes	0.314	LAMA2–DAG1	0.091
3	08_Oligodendrocytes	0.351	13_Neurons_EXT_L56	0.304	GNAI2–ADCY1	0.086
4					RSPO2–ZNRF3	0.084
5					CALM1–KCNQ1	0.082
Factor 8						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	09_Neurons_INT_RELN	0.373	17_Stromal_cells	0.296	NCAM1–PTPRZ1	0.118
2	02_Neurons_EXT_L4	0.372	02_Neurons_EXT_L4	0.294	NCAM1–GFRA1	0.109
3	21_OPCs	0.369	01_Maturing_Neurons	0.276	BSG–SLC16A7	0.109
4					NCAM1–CACNA1C	0.106
5					RGMA–NEO1	0.101
Factor 9						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	00_Truncated_Radial_Glia	0.444	00_Truncated_Radial_Glia	0.344	WNT9A–FZD6_LRP5	0.113
2	18_Neurons_EXT_L56	0.378	11_NPCs	0.333	APOE–ABCA1	0.102
3	14_Pericytes	0.335	01_Maturing_Neurons	0.301	NTRK3–PTPRD	0.097
4					PTN–PTPRZ1	0.095
5					SLITRK2–PTPRD	0.093

Supplementary Table C (cont.)

Factor 10						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	07_Neurons_EXT	0.405	11_NPCs	0.348	WNT3–FZD6_LRP5	0.094
2	18_Neurons_EXT_L56	0.328	04_NPCs	0.288	SPON1–APP	0.090
3	10_Truncated_Radial_Glia	0.289	16_Neurons_EXT_L6	0.275	APP–LRP1	0.086
4					NCAM1–NCAM2	0.082
5					TNFSF13B–TFRC	0.080
Factor 11						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	10_Truncated_Radial_Glia	0.403	09_Neurons_INT_RELN	0.371	NLGN2–NRXN1	0.130
2	19_Neurons_EXT_L56	0.351	13_Neurons_EXT_L56	0.323	CALM3–KCNQ1	0.127
3	05_Neurons_EXT_L56	0.349	05_Neurons_EXT_L56	0.285	SHH–LRP2	0.103
4					HSPG2–DAG1	0.100
5					NLGN4X–NRXN1	0.096
Factor 12						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	12_Oligodendrocytes	0.321	01_Maturing_Neurons	0.362	GDF11–ACVR2B_TGFBR1	0.127
2	13_Neurons_EXT_L56	0.308	04_NPCs	0.340	NRXN3–NLGN2	0.112
3	18_Neurons_EXT_L56	0.304	11_NPCs	0.338	CALM1–PDE1A	0.110
4					CALM3–PDE1A	0.106
5					SHH–BOC	0.098

Supplementary Table C (cont.)

Factor 13						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	18_Neurons_EXT_L56	0.362	00_Truncated_Radial_Glia	0.341	NRXN1–NLGN3	0.121
2	03_Neurons_INT_PV	0.347	15_Intermediate_NPCs	0.322	COL4A1–ITGAV_ITGB8	0.115
3	21_OPCs	0.314	04_NPCs	0.318	DSCAM–DCC	0.112
4					NCAM1–CACNA1C	0.105
5					FAT4–DCHS1	0.102
Factor 14						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	06_Late_Radial_Glia	0.479	13_Neurons_EXT_L56	0.427	LAMA2–ITGAV_ITGB8	0.144
2	07_Neurons_EXT	0.373	21_OPCs	0.363	MDK–LRP2	0.142
3	17_Stromal_cells	0.373	04_NPCs	0.342	COL4A2–ITGAV_ITGB8	0.102
4					LRPAP1–LDLR	0.102
5					EFNA5–EPHA3	0.101
Factor 15						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	12_Oligodendrocytes	0.443	03_Neurons_INT_PV	0.377	SLIT2–APP	0.124
2	13_Neurons_EXT_L56	0.309	16_Neurons_EXT_L6	0.346	SLIT2–ROBO2	0.120
3	17_Stromal_cells	0.308	19_Neurons_EXT_L56	0.280	CNTN2–CNTN1	0.113
4					PDGFC–PDGFRB	0.108
5					SLIT2–ROBO1	0.106

Supplementary Table C (cont.)

Factor 16						
Rank	**Sender Cell Type**	**Score**	**Receiver Cell Type**	**Score**	**Ligand–Receptor Pair**	**Score**
1	04_NPCs	0.430	01_Maturing_Neurons	0.408	MDK–PTPRZ1	0.164
2	12_Oligodendrocytes	0.352	07_Neurons_EXT	0.298	AFDN–EPHB2	0.132
3	01_Maturing_Neurons	0.307	14_Pericytes	0.280	COL4A3–CD47	0.130
4					DLK1–NOTCH1	0.129
5					BMP5–BMPR1A_BMPR2	0.125

REFERENCES

Armingol, Erick, Hratch M. Baghdassarian, and Nathan E. Lewis. 2024. “The Diversification of Methods for Studying Cell-Cell Interactions and Communication.” *Nature Reviews. Genetics* 25 (6): 381–400.

Armingol, Erick, Hratch M. Baghdassarian, Cameron Martino, Araceli Perez-Lopez, Caitlin Aamodt, Rob Knight, and Nathan E. Lewis. 2022. “Context-Aware Deconvolution of Cell-Cell Communication with Tensor-cell2cell.” Nature Communications 13 (1): 3665.

Armingol, Erick, Adam Officer, Olivier Harismendy, and Nathan E. Lewis. 2021. “Deciphering Cell-Cell Interactions and Communication from Gene Expression.” *Nature Reviews. Genetics* 22 (2): 71–88.

Biswas, Niloy, Lester Mackey, and Xiao-Li Meng. 2022. “Scalable Spike-and-Slab.” In *arXiv [stat.CO]*. April 4. arXiv. https://doi.org/10.48550/arXiv.2204.01668.

Buchanan, Joann, Nuno Maçarico da Costa, and Lucas Cheadle. 2023. “Emerging Roles of Oligodendrocyte Precursor Cells in Neural Circuit Development and Remodeling.” *Trends in Neurosciences* 46 (8): 628–639.

Buckner, Randy L., Fenna M. Krienen, Angela Castellanos, Julio C. Diaz, and B. T. Thomas Yeo. 2011. “The Organization of the Human Cerebellum Estimated by Intrinsic Functional Connectivity.” *Journal of Neurophysiology* 106 (5): 2322–2345.

Burkhardt, Daniel B., Jay S. Stanley 3rd, Alexander Tong, Ana Luisa Perdigoto, Scott A. Gigante, Kevan C. Herold, Guy Wolf, Antonio J. Giraldez, David van Dijk, and Smita Krishnaswamy. 2021. “Quantifying the Effect of Experimental Perturbations at Single-Cell Resolution.” Nature Biotechnology 39 (5): 619–629.

Büttner, M., J. Ostner, C. L. Müller, F. J. Theis, and B. Schubert. 2021. “scCODA Is a Bayesian Model for Compositional Single-Cell Data Analysis.” *Nature Communications* 12 (1): 6876.

Chen, Chia-Hsiang, Pin-Wei Lee, Hsiao-Mei Liao, and Pi-Kai Chang. 2017. “Neuroligin 2 R215H Mutant Mice Manifest Anxiety, Increased Prepulse Inhibition, and Impaired Spatial Learning and Memory.” Frontiers in Psychiatry 8: 257

Cheng, Changde, Wenan Chen, Hongjian Jin, and Xiang Chen. 2023. “A Review of Single-Cell RNA-Seq Annotation, Integration, and Cell-Cell Communication.” *Cells (Basel, Switzerland)* 12 (15): 1970.

Courchesne, Eric, Vani Taluja, Sanaz Nazari, Caitlin M. Aamodt, Karen Pierce, Kuaikuai Duan, Sunny Stophaeros, Linda Lopez, Cynthia Carter Barnes, Jaden Troxel, Kathleen Campbell, Tianyun Wang, Kendra Hoekzema, Evan E. Eichler, Joao V. Nani, Wirla Pontes, Sandra Sanchez Sanchez, Michael V. Lombardo, Janaina S. de Souza, Mirian A. F. Hayashi, and Alysson R. Muotri. 2024. “Embryonic Origin of Two ASD Subtypes of Social

Symptom Severity: The Larger the Brain Cortical Organoid Size, the More Severe the Social Symptoms." Molecular Autism 15 (1): 22.

Dimitrov, Daniel, Dénes Türei, Martin Garrido-Rodriguez, Paul L. Burmedi, James S. Nagai, Charlotte Boys, Ricardo O. Ramirez Flores, Hyojin Kim, Bence Szalai, Ivan G. Costa, Alberto Valdeolivas, Aurélien Dugourd, and Julio Saez-Rodriguez. 2022. "Comparison of Methods and Resources for Cell-Cell Communication Inference from Single-Cell RNA-Seq Data." Nature Communications 13 (1): 3224.

Esnafoglu, Erman, and Selma Cirrik. 2018. "Increased Serum Midkine Levels in Autism Spectrum Disorder Patients." *The International Journal of Neuroscience* 128 (7): 677–681.

Galvez-Contreras, Alma Y., David Zarate-Lopez, Ana L. Torres-Chavez, and Oscar Gonzalez-Perez. 2020. "Role of Oligodendrocytes and Myelin in the Pathophysiology of Autism Spectrum Disorder." *Brain Sciences* 10 (12): E951.

Girault, Jessica B., and Joseph Piven. 2020. "The Neurodevelopment of Autism from Infancy through Toddlerhood." *Neuroimaging Clinics of North America* 30 (1): 97–114.

Graciarena, Mariana, Araceli Seiffe, Brahim Nait-Oumesmar, and Amaicha M. Depino. 2018. "Hypomyelination and Oligodendroglial Alterations in a Mouse Model of Autism Spectrum Disorder." *Frontiers in Cellular Neuroscience* 12: 517.

Guo, Yunxia, Wenjia Wang, Kaiqiang Ye, Liyong He, Qinyu Ge, Yan Huang, and Xiangwei Zhao. 2023. "Single-Nucleus RNA-Seq: Open the Era of Great Navigation for FFPE Tissue." International Journal of Molecular Sciences 24 (18): 13744.

Korkmaz, Baris. 2011. "Theory of Mind and Neurodevelopmental Disorders of Childhood." *Pediatric Research* 69 (5 Pt 2): 101R–8R.

Le, Yao, Kavitha Rajasekhar, Tricia Y. J. Loo, Timothy E. Saunders, Thorsten Wohland, and Christoph Winkler. 2025. "Midkine-a Interacts with Ptprz1b to Regulate Neural Plate Convergence and Midline Formation in the Developing Zebrafish Hindbrain." *Developmental Biology* 521 (May): 52–74.

McFadden, Kathryn, and Nancy J. Minshew. 2013. "Evidence for Dysregulation of Axonal Growth and Guidance in the Etiology of ASD." *Frontiers in Human Neuroscience* 7 (October): 671.

Meyerink, Brandon L., Neeraj K. Tiwari, and Louis-Jan Pilaz. 2020. "Ariadne's Thread in the Developing Cerebral Cortex: Mechanisms Enabling the Guiding Role of the Radial Glia Basal Process during Neuron Migration." *Cells (Basel, Switzerland)* 10 (1): E3.

Nakagawa, Naoki, Charlotte Plestant, Keiko Yabuno-Nakagawa, Jingjun Li, Janice Lee, Chu-Wei Huang, Amelia Lee, Oleh Krupa, Aditi Adhikari, Suriya Thompson, Tamille Rhynes, Victoria Arevalo, Jason L. Stein, Zoltán Molnár, Ali Badache, and E. S. Anton. 2019. "Memo1-Mediated Tiling of Radial Glial Cells Facilitates Cerebral Cortical Development." Neuron 103 (5): 836–852.e5.

Padmanabhan, Aarthi, Charles J. Lynch, Marie Schaer, and Vinod Menon. 2017. “The Default Mode Network in Autism.” *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 2 (6): 476–486.

Roese-Koerner, Beate, Laura Stappert, and Oliver Brüstle. 2017. “Notch/Hes Signaling and miR-9 Engage in Complex Feedback Interactions Controlling Neural Progenitor Cell Proliferation and Differentiation.” *Neurogenesis (Austin, Tex.)* 4 (1): e1313647.

Sandin, Sven, Paul Lichtenstein, Ralf Kuja-Halkola, Christina Hultman, Henrik Larsson, and Abraham Reichenberg. 2017. “The Heritability of Autism Spectrum Disorder.” *The Journal of the American Medical Association* 318 (12): 1182–1184.

Schwartz, Martin Alexander. 2010. “Integrins and Extracellular Matrix in Mechanotransduction.” *Cold Spring Harbor Perspectives in Biology* 2 (12): a005066.

Stevens, Hanna E., Karen M. Smith, Brian G. Rash, and Flora M. Vaccarino. 2010. “Neural Stem Cell Regulation, Fibroblast Growth Factors, and the Developmental Origins of Neuropsychiatric Disorders.” *Frontiers in Neuroscience* 4 (September). https://doi.org/10.3389/fnins.2010.00059.

Tamouza, Ryad, Fernanda Volt, Jean-Romain Richard, Ching-Lien Wu, Jihène Bouassida, Wahid Boukouaci, Pauline Lansiaux, Barbara Cappelli, Graziana Maria Scigliuolo, Hanadi Rafii, Chantal Kenzey, Esma Mezouad, Soumia Naamoune, Leila Chami, Florian Lejuste, Dominique Farge, and Eliane Gluckman. 2022. “Possible Effect of the Use of Mesenchymal Stromal Cells in the Treatment of Autism Spectrum Disorders: A Review.” Frontiers in Cell and Developmental Biology 10 (July): 809686.

Tata, Mathew, Christiana Ruhrberg, and Alessandro Fantin. 2015. “Vascularisation of the Central Nervous System.” *Mechanisms of Development* 138 Pt 1 (November): 26–36.

Trobiani, Laura, Maria Meringolo, Tamara Diamanti, Yves Bourne, Pascale Marchot, Giuseppina Martella, Luciana Dini, Antonio Pisani, Antonella De Jaco, and Paola Bonsi. 2020. “The Neuroligins and the Synaptic Pathway in Autism Spectrum Disorder.” Neuroscience and Biobehavioral Reviews 119 (December): 37–51.

Virshup, Isaac, Sergei Rybakov, Fabian J. Theis, Philipp Angerer, and F. Alexander Wolf. 2024. “anndata: Access and Store Annotated Data Matrices.” Journal of Open Source Software 9 (101): 4371.

Yazdanbakhsh, Arash, Kim T. M. Dang, Kelvin Kuang, Tingru Lian, Xuefeng Liu, Songlin Xie, and Basilis Zikopoulos. 2025. “Artificial Intelligence Networks Combining Histopathology and Machine Learning Can Extract Axon Pathology in Autism Spectrum Disorder.” Autism Research: Official Journal of the International Society for Autism Research 18 (11): 2210–2230.

Yeo, B. T. Thomas, Fenna M. Krienen, Jorge Sepulcre, Mert R. Sabuncu, Danial Lashkari, Marisa Hollinshead, Joshua L. Roffman, Jordan W. Smoller, Lilla Zöllei, Jonathan R. Polimeni, Bruce Fischl, Hesheng Liu, and Randy L. Buckner. 2011. “The Organization of

the Human Cerebral Cortex Estimated by Intrinsic Functional Connectivity." Journal of Neurophysiology 106 (3): 1125–1165.

Zielinski, Brandon A., Jeffrey S. Anderson, Alyson L. Froehlich, Molly B. D. Prigge, Jared A. Nielsen, Jason R. Cooperrider, Annahir N. Cariello, P. Thomas Fletcher, Andrew L. Alexander, Nicholas Lange, Erin D. Bigler, and Janet E. Lainhart. 2012. "scMRI Reveals Large-Scale Brain Network Abnormalities in Autism." PloS One 7 (11): e49172.