ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

IDENTIFYING SPATIO-MOLECULAR DISTRIBUTION OF VERTEBRATE OLFACTORY BULB THROUGH SC-RNA SEQUENCING IN MOUSE AND ZEBRAFISH

ERSİN BERKAY PEKER

M.Sc. THESIS

DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

SUPERVISOR
Prof. Osman Uğur SEZERMAN

SECONDARY SUPERVISOR
Prof. Emre YAKSI

ISTANBUL-2024

ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

IDENTIFYING SPATIO-MOLECULAR DISTRIBUTION OF VERTEBRATE OLFACTORY BULB THROUGH SC-RNA SEQUENCING IN MOUSE AND ZEBRAFISH

ERSİN BERKAY PEKER

M.Sc. THESIS

DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

SUPERVISOR
Prof. Osman Uğur SEZERMAN

SECONDARY SUPERVISOR
Prof. Emre YAKSI

ISTANBUL-2024

Department: Department of Biostatistics and Bioinformatics
Program: Biostatistics and Bioinformatics Master of Science Program with Thesis
Thesis Title: Identifying Spatio-molecular Distribution of Vertebrate Olfactory Bulb through sc-RNA Sequencing in Mouse and Zebrafish
Student's name and Surname: Ersin Berkay Peker
Date of Defense: 25/07/2024

This is to certify that I have examined this copy of the master thesis. I confirm that she/he prepared after fulfilling the specified requirements in the associated legislations before the final examining committee whose signatures are below.

Jury Member (Supervisor of the thesis)	Prof. Osman Uğur Sezerman, Acıbadem Mehmet Ali Aydınlar University, School of Medicine, Basic Sciences, Biostatistics and Medical Informatics	Signature
Jury Member	Asst. Prof. Zeynep Tokcaer Keskin Acıbadem Mehmet Ali Aydınlar University, Faculty of Engineering and Natural Sciences, Molecular Biology and Genetics	Signature
Jury Member	Asst. Prof. Onur Emre Onat Bezmialem University, Faculty of Life Sciences, and Biotechnology, Molecular Biology and Genetics	Signature

DECLARATION

I declare that this thesis work is my own work, I had no unethical behavior at any stages from the planning to the writing of the thesis, I obtained all the information in this thesis in accordance with academic and ethicel rules, I cited all the information and comments that were not obtained with this thesis work, and I provided resources in the list of references. I also declare that there was no violation of any patents and copyrights during the study and writing of this thesis.

25.07.2024

"Ersin Berkay Peker"

PREFACE AND ACKNOWLEDGEMENT

During the process of this thesis, I have met with amazing people whose ideas helped me overcome the hurdles I encountered. Also, their presence gave me strength and made me realize that I am not alone in this.

I would like to thank TUBITAK for their support for my Master's degree with the BIDEB 2210-E programme.

I would first like to thank my secondary supervisor, Prof. Dr. Emre Yaksi who supported me constantly, always stayed positive and constructive whenever there was a problem. His great guidance has led me to find the light at the end of the tunnel, and brought me to today where I am. He had always been the greatest lantern that enlightened the path we walked on together, and he keeps doing so for his every other student.

I would like to thank my supervisor, Prof. Dr. Osman Uğur Sezerman whose lectures regarding bioinformatics helped me a lot to understand the concept of the methods that I used. His interesting ideas and hard questions made me realize the unlimited potential of how far math can reach.

I would like to thank Bjørn André Bredesen-Aa who supported me with his great knowledge at computational biology. His tools and methods revealed the full potential of this thesis. I enjoyed our every discussion during the process which made me realize that there are unlimited possibilities of how math can be applied to.

Lastly but most importantly, I would like to thank to my grandmother Emine Meral Olcayto who passed away in February 2023. I dedicate this thesis and all my efforts to her. May she rest in peace, may her tales be told in the most remote corners of science.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

ABCC9	ATP-Binding Cassette, Sub-family C Member 9
ACTA2	Actin Alpha 2, Smooth Muscle, Aorta
APOC1	Apolipoprotein C-I
APOEB	Apolipoprotein Eb
AP	Anterior Posterior
APOLD1	Apolipoprotein L Domain Containing 1
AZBA	Adult Zebrafish Brain Atlas
AZTEC	The Atlas of Zebrafish Transcriptomic Encephalic Cytoarchitecture
BSA	Bovine Serum Albumin
CBLN1	Cerebellin 1 Precursor Protein
CBLN4	Cerebellin 4 Precursor Protein
CALB1	Calbindin 1
CDH23	Cadherin Related 23 (Otocadherin)
CDHR1	Cadherin-Related Family Member 1
CD74A	CD74 Molecule, Major Histocompatibility Complex, Class II Invariant Chain A
CHRNA7	Cholinergic Receptor, Nicotinic, Alpha Polypeptide 7
CLDN5	Claudin 5
CLDN5A	Claudin 5a
CLDN11	Claudin 11
CCA	Canonical Correlation Analysis
CoCy	Comparative Cytoarchitectonics
CRH	Corticotropin Releasing Hormone
CX3CR1	C-X3-C Motif Chemokine Receptor 1
CX43	Gap Junction Protein Alpha 1b
DNA	Deoxyribonucleic Acid
DE	Differential Expression
EBF3	Early B Cell Factor 3
EggNOG	Non-supervised Orthologous Group

EOMESA	Eomesodermin Homolog A
ETC	External Tufted Cell
FEZF2	FEZ Family Zinc Finger 2
FABP7A	Fatty Acid Binding Protein 7, Brain, A
FLT1	FMS-like Tyrosine Kinase 1
GAD1B	Glutamate Decarboxylase 1b
GC	Granule Cell
GLM	Generalized Linear Model
GLRA1	Glycine Receptor, Alpha 1 Subunit
GLULA	Glutamate-ammonia ligase (glutamine synthase) A
GLUT	Glutamatergic Cell
GPC5	Glypican 5
GPR139	G Protein-Coupled Receptor 139
GRM6B	Glutamate Receptor, Metabotropic 6b
IGF2	Insulin-like Growth Factor 2
IGABA	Immature GABAergic Cell
IGFBPL1	Insulin-Like Growth Factor Binding Protein-Like 1
IVT	In Vitro Transcription
KCNH3	Potassium Voltage-Gated Channel, Subfamily H (Eag-Related), Member 3
KCNIP4	Kv Channel Interacting Protein 4
K-NN	K-nearest Neighbor
KMR	The K Mutual Ratio
MAL	Myelin and Lymphocyte Protein, T Cell Differentiation Protein
MATN3B	Matrilin 3b
MC	Mitral Cell
MERFISH	Multiplexed Error-robust Fluorescence in Situ Hybridization
ML	Medial Lateral
MOG	Myelin Oligodendrocyte Glycoprotein
NPY	Neuropeptide Y

NN	Non-neuronal
ORN	Olfactory Receptor Neuron
OB	Olfactory Bulb
PCA	Principal component analysis
PARM1	Prostate Androgen-Regulated Mucin-Like Protein 1
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PGC	Periglomerular Cell
PDE5A	Phosphodiesterase 5A, cGMP-Specific
PLP1	Proteolipid Protein (Myelin) 1
PLCXD3	Phosphatidylinositol-Specific Phospholipase C X Domain Containing 3
PNOC	Prepronociceptin
PRSS12	Serine Protease 12 Neurotrypsin (Motopsin)
PTGDS	Prostaglandin D2 Synthase (Brain)
PTK2B	PTK2 Protein Tyrosine Kinase 2 Beta
RGS5	Regulator of G-Protein Signaling 5
RNA	Ribonucleic Acid
RPRM	Reprimo, TP53 Dependent G2 Arrest Mediator Candidate
SCN5A	Sodium Channel, Voltage-Gated, Type V, Alpha
SLC1A2	Solute Carrier Family 1
SLC17A6	Solute Carrier Family 17
SLC17A6A	Solute Carrier Family 17 Member 6a
SLC17A6B	Solute Carrier Family 17 Member 6b
SLC17A7	Solute Carrier Family 17
SLC17A7A	Solute Carrier Family 17 Member 7a
SLC32A1	Solute Carrier Family 32
SNAP25A	Synaptosome Associated Protein 25a
SMOX	Spermine Oxidase
SOX11	SRY (Sex Determining Region Y)-Box 11
SV2B	Synaptic Vesicle Glycoprotein 2b
TAC1	Tachykinin Precursor 1

TACR3L	Tachykinin Receptor 3-like
TH	Tyrosine Hydroxylase
TMEM132D	Transmembrane Protein 132D
TPBG	Trophoblast Glycoprotein
TRH	Thyrotropin Releasing Hormone
TRPC7	Transient Receptor Potential Cation Channel Subfamily C, Member 7
TRP73	Transformation Related Protein 73
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique Molecular Identifier
VD	Ventral Dorsal
VWC2L	Von Willebrand Factor C Domain-Containing

LIST OF FIGURES

ABSTRACT

Identifying Spatio-Molecular Distribution of Vertebrate Olfactory Bulb Through Single Cell RNA Sequencing in Mouse and Zebrafish

The vertebrate olfactory bulb exhibits a topographical organization, where distinct odor classes are encoded in spatially distinct regions. This chemotopic (or odotopic) map plays a crucial role in olfactory computations. However, the molecular correlates of this organization in projection neurons and interneurons are not fully understood. By generating and utilizing single-cell and spatial transcriptomic atlases for both zebrafish and mouse, the molecular correlates of olfactory bulb topography was investigated. Basic cell types across species were identified and matched, revealing a continuum of gene expression diversity in classes of glutamatergic projection neurons and GABAergic interneurons. This molecular diversity correlates with topographical organization of the olfactory bulb, revealing higher transcriptomic similarity among nearby neurons. Both zebrafish and mouse projection neurons in spatially distinct olfactory bulb zones exhibit a high degree of molecular similarity, clearly different from spatially distant olfactory bulb zones. These molecularly identifiable groups of projection neurons are organized into non-overlapping regions across the dorsoventral and mediolateral axes of the olfactory bulbs. In zebrafish, these regions overlap with zones encoding alarm, food, and social odors. However, interneurons exhibit a different logic of molecular topography. In both species, the major axis of molecular diversity is the depth from the surface of the olfactory bulb, where interneurons at various depths exhibit prominent transcriptomic differences. While no obvious zonal organization of interneuron molecular topography along the dorsoventral and mediolateral axes of the mouse olfactory bulb was observed, the zebrafish olfactory bulb exhibits a prominent separation of molecularly different interneuron groups along these dimensions.

Keywords: Olfactory bulb, Chemotopy, Projection neuron, Interneuron, Single cell RNA sequencing

ÖZET

Omurgalıların Koku Soğanının Tek Hücre RNA Dizileme ile Bölgesel ve Moleküler Dağılımının Fare ve Zebra Balığında Tespit Edilmesi

Omurgalıların koku soğanı, belirgin koku türlerinin uzamsal olarak farklı bölgelerde kodlandığı bir topografik düzen gösterir. Bu kemotopik (veya odotopik) harita, koku işlemlerinde önemli bir rol oynar. Ancak, bu düzenin projeksiyon nöronları ve ara nöronlardaki moleküler karşılıkları tam olarak anlaşılamamıştır. Bu çalışmada zebra balığı ve fare için tek hücre RNA verisi ile uzamsal transkriptom atlasları kullanılarak, koku soğanının topografisinin moleküler karşılıkları araştırıldı. Projeksiyon nöronları ve ara nöronlar sınıflarındaki gen çeşitliliğini kullanarak farklı türler arasında temel hücre tiplerini tanımlandı ve eşleştirildi. Bu moleküler çeşitliliğin koku soğanının topografik düzeni ile bağlantılı olduğu ve bununla birlikte birbirlerine yakın olan nöronların daha yüksek benzerlik gösterdiği görüldü. Zebra balığı ve farede bölgesel olarak farklı olan projeksiyon nöronlarının, bölgesel olarak birbirlerine uzak olan nöronlara kıyasla daha yüksek moleküler benzerlik gösterdiği tespit edildi. Bu moleküler olarak tespit edilebilen projeksiyon nöronları dorsoventral ve mediolateral eksenlerinde birbirleriyle çakışmayan bölgeler içinde organize olmuştur. Zebra balığında bu bölgeler alarm, yiyecek ve sosyal kokuları kodlayan bölgeler ile örtüşmektedir. Ancak ara nöronlar farklı bir moleküler topografi mantığı gösterir. Her iki türde de moleküler çeşitliliğin ara ekseni koku soğanının derinliğidir, burada farklı derinliklerdeki ara nöronlar önemli transkriptomik farklılıklar gösterir. Fare koku soğanınında dorsoventral ve mediolateral eksenleri boyunca ara nöron moleküler topografyasının belirgin bir zonal organizasyonu gözlemlenmemesine rağmen, zebra balığı koku soğanında bu eksenler boyunca moleküler olarak farklı ara nöron gruplarının belirgin bir ayrımını görülmektedir.

Anahtar Sözcükler: Koku soğanı, Kemotopi, Projeksiyon nöron, Ara nöron, Tek hücre RNA dizileme

1 INTRODUCTION

1.1 Brain Topography and Sensory Systems

Brain topography plays a crucial role in the functioning of the brain, from sensory systems to higher-order cognitive processes (1). For instance, the visual and auditory systems exhibit topographical organization in the retina and cochlea, mapping different frequencies or visual receptive fields to the thalamus and subsequently to primary cortical areas (2,3). Similarly, the gustatory and somatosensory systems display topographical representations of sensory information (4,5). Even in higher-order brain regions important for cognitive function, such as the hippocampus and medial entorhinal cortex, dorsal-ventral topography corresponds to distinct functions, ranging from the encoding of space to mood and the encoding of different grid fields (6,7)

1.2 Developmental Origins of Functional Topography

The functional topography of brain circuits has developmental origins predominantly influenced and organized by molecular cues (8). Molecular signals guide the migration of neurons to their designated positions and direct axon pathfinding, ensuring precise innervation (9,10). This molecular orchestration of developmental phenomena is a crucial determinant of the functional topography observed in mature brain (11).

1.3 Topographical Organization in the Olfactory System

As with other sensory systems, the olfactory system exhibits a remarkable degree of topography (12) Olfactory receptor neurons expressing the same receptors project to spatially distinct locations in the olfactory bulb, known as olfactory glomeruli (13,14). Olfactory glomeruli are innervated by the apical dendrites of mitral cells, the projection neurons of the olfactory bulb (15). Individual mitral cells, both in zebrafish and rodents, receive information from one of many olfactory glomeruli and

communicate this information to distinct target regions in the anterior olfactory nucleus, piriform cortex, cortical amygdala, hypothalamus, and the habenula (16,17).

1.4 Chemotopic Maps and Functional Topography

This high degree of topographical organization is reflected in the functional topography of the olfactory bulb, where odorant receptors encoding distinct categories of odors innervate spatially distinct regions (18). These innervation patterns give rise to chemotopic or odotopic maps, where olfactory glomeruli and associated projection neurons (mitral and tufted cells) in spatially distinct regions of the olfactory bulb encode distinct categories of odorants (19,20) Such chemotopic maps are present in vertebrate olfactory bulbs and insect antennal lobes, underscoring the importance of this organization (21,22). The odotopic activation of particular parts of the bulb by particular odorants now has been confirmed by many techniques including the mapping of field potentials (23), unit recordings of mitral cells (23–25), up-regulation of immediate early gene products in the glomerular and granule cell layers (26–34), optical imaging of either endogenous responses (35–39) or voltage- and calcium-sensitive dye responses , and functional magnetic resonance imaging (40–44).

Odotopy has been extended to other classes of vertebrates such as zebrafish (45–47), as well as to insects such as honeybees (48–50) moths (51–55), ants (56) and *Drosophila* (57–59) Given this evidence, odorant-specific spatial patterning of glomerular activation is a basic characteristic of olfactory systems. Different categories of chemotopically encoded odors are crucial for regulating animal behavior (60,61) In zebrafish, distinct zones of the olfactory bulb have been shown to encode alarm odors, bile odors, amino acids, and food odors (62). Moreover, distinct olfactory bulb regions project to specific higher brain regions in vertebrates and insects (63,64). The topographical organization is not limited to olfactory bulb inputs and outputs but is also observed at the level of periglomerular interneurons and the granule cell layer (65). All these suggest that the function and development of the olfactory bulb are based on a well-defined and relatively stereotyped topography. Despite this high degree of topography in the olfactory bulb, the molecular correlates of this

topographical organization, both in the projection neurons and interneurons of the olfactory bulb, are not fully understood.

1.5 Research Objectives and Aim

The primary objective of this study is to identify and characterize the spatio-molecular distribution of cells within the olfactory bulb of mice and zebrafish using single-cell RNA sequencing. We hypothesize that distinct molecular signatures will correspond to specific spatial regions within the olfactory bulb, reflecting functional specialization. Our specific aims are:

To establish a framework to analyze single cell spatial transcriptome datasets from zebrafish and mouse olfactory bulbs.

To identify cell types of vertebrate olfactory bulbs and identify subclasses of these cell types along with relevant marker genes.

To identify whether these molecular organization of different cell types, and subclasses have topographical organization established by spatial transcriptomics analysis of datasets coming from zebrafish and mouse.

By achieving these aims, we seek to provide a comprehensive understanding of the organizational principles that govern olfactory processing in vertebrates and to identify potential targets for therapeutic intervention in sensory disorders.

2 BACKGROUND

2.1 Overview of Olfactory Systems in Vertebrates

Olfaction is a crucial sense for vertebrates, enabling them to detect and interpret a vast array of chemical signals from the environment (61,66). This sensory system is integral to survival, playing a key role in behaviors such as attraction, food searching, mate recognition, and danger avoidance (67,68). Different olfactory cues generate various types of behaviors, and this information is encoded in the mitral cell layers of the olfactory bulbs in both fish and mice (47,69).

The olfactory bulb acts as the first relay station for sensory input received from olfactory receptor neurons (ORNs) located in the nasal epithelium (18,70) In vertebrates, the olfactory bulb is a highly organized structure consisting of several distinct layers, each contributing to the transformation of chemical signals into neural signals that the brain can interpret (71,72). Mitral cells, located in the olfactory bulb, send their projections to higher brain regions responsible for different functions, such as cortical areas important for learning and hypothalamic and amygdala areas essential for innate behaviors (73,74).

This complex organization within the olfactory bulb allows for the precise discrimination and integration of olfactory information (18,21). The mitral cells play a pivotal role in this process, encoding and transmitting olfactory information to higher brain regions, which then interpret these signals to facilitate a range of behaviors critical for survival and reproduction (69,75). The olfactory system, therefore, enables animals to identify food sources, avoid predators, and engage in social and reproductive behaviors through the detection of pheromones, underscoring its fundamental importance in vertebrate biology (76,77).

2.2 Model Organisms Mouse and Zebrafish

Mice (Mus musculus) and zebrafish (Danio rerio) are among the most widely used model organisms in olfactory research, each offering unique advantages. Mice are highly relevant for studying mammalian olfaction due to their complex and well-characterized olfactory system, which shares many features with the human olfactory system (78,79). Their genetic similarity to humans and the availability of numerous genetic tools makes them an ideal model for investigating the molecular and cellular mechanisms underlying olfactory processing (80,81).

Zebrafish, in contrast, offer several complementary advantages. Their embryos are transparent and develop externally, making them highly amenable to in vivo imaging and developmental studies (82). Zebrafish also have a simpler olfactory system compared to mammals, which can facilitate the study of fundamental principles of olfactory processing (83). The rapid development and ease of genetic manipulation in zebrafish make them an excellent model for functional and genetic studies (84). By comparing the olfactory systems of mice and zebrafish, researchers can gain insights into the evolutionary conservation and divergence of olfactory mechanisms across vertebrates.

2.3 The Structure and Function of the Olfactory Bulb in Vertebrate Olfaction

The olfactory bulb (OB) is a critical structure in the vertebrate brain, responsible for the initial processing of olfactory information (74). Located at the anterior end of the brain, it receives direct input from olfactory receptor neurons (ORNs) located in the nasal epithelium (85). There are about 100 different types of odorant receptors in fish and 1000 in mammals (66). The axons of the ORNs project to the OB and form synapses with the mitral cells (MCs) in the olfactory glomeruli (69). Glomeruli are neuropil units in the outer layer of the OB that consist of ORN axon terminals, dendrites of MCs, and local inhibitory interneurons (74,86). ORNs expressing the same receptor type converge onto the same glomerulus in vertebrates (12).

The olfactory bulb is organized into several distinct layers, each with specific functions (69). The outermost layer, the glomerular layer, contains spherical structures called glomeruli, where ORN axons synapse with the dendrites of mitral and tufted cells (72) (Figure 1). Each glomerulus receives input from ORNs expressing the same odorant receptor gene, allowing for the spatial mapping of odorant signals (20). The external plexiform layer contains the dendrites of mitral and tufted cells, as well as various interneurons that modulate their activity (18). The mitral cell layer contains the cell bodies of mitral cells, which are the primary output neurons of the olfactory bulb, transmitting processed olfactory information to higher brain regions such as the olfactory cortex (74). The internal plexiform layer contains fibers and synapses that further modulate the activity of mitral and tufted cells (69). The deepest layer, the granule cell layer, contains granule cells, which are interneurons that provide inhibitory input to mitral and tufted cells, contributing to the fine-tuning of olfactory signals (19).

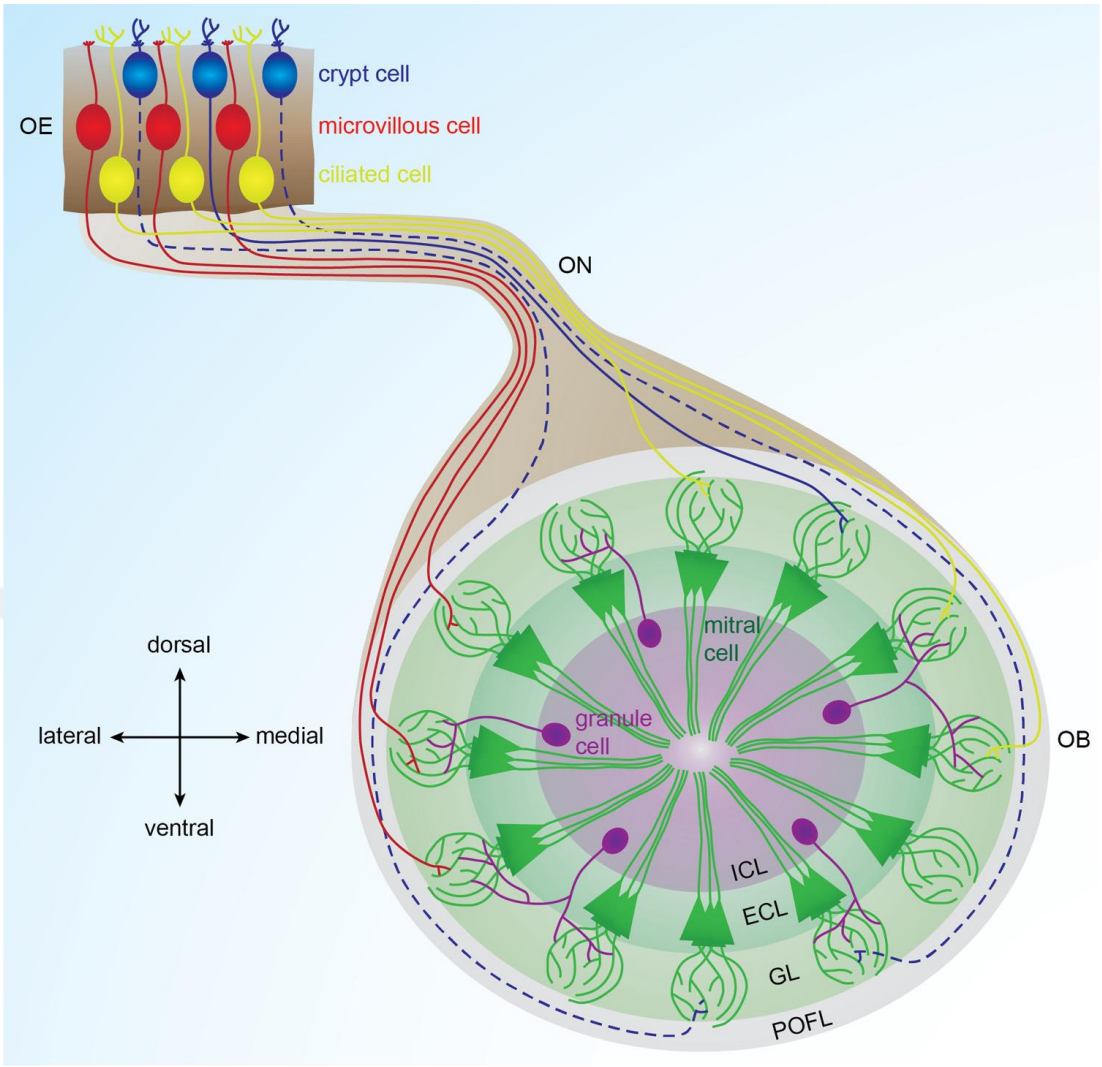

Figure 1. Organization of the olfactory bulb network (87).

In most vertebrates, individual MCs have large dendritic trees and receive input from multiple glomeruli (85,88) (Figure 2). The deeper layers of the OB contain many local inhibitory interneurons, such as granule cells (GCs) (74,85). GCs are axonless and form reciprocal synapses with MC dendrites (89). These interactions are excitatory from MCs to GCs and inhibitory from GCs to MCs (69).

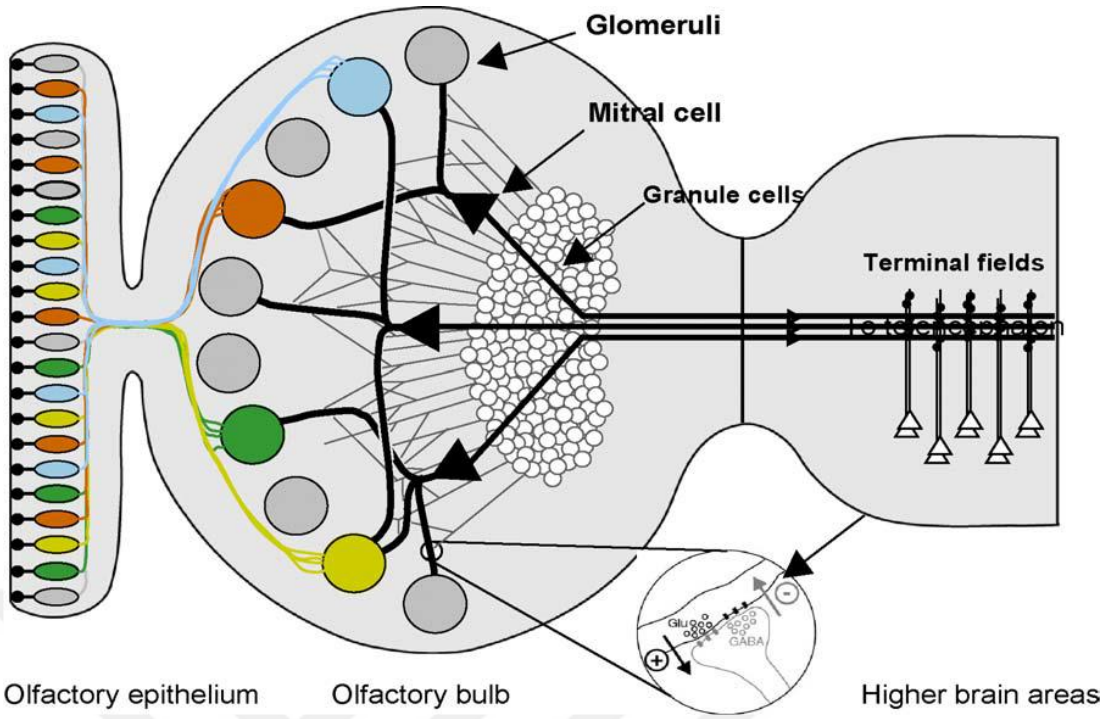

Figure 2. Overview of the olfactory system (90).

The precise organization and connectivity within these layers are critical for accurate odor perception and discrimination (18,75). The olfactory bulb functions as a sophisticated neural network, transforming raw sensory input into a format that higher brain regions can interpret and integrate with other sensory modalities (19,69). This sophisticated structure allows vertebrates to detect and interpret a wide array of chemical signals from the environment, facilitating behaviors such as food searching, mate recognition, and danger avoidance (66,77).

2.4 Combinatorial Odor Representation and Temporal Dynamics in the Olfactory Bulb

In the olfactory bulb (OB), odors are initially represented through combinatorial activity patterns across multiple glomeruli. Each glomerulus acts as a convergence point for ORNs that express the same odorant receptor, functioning as a key unit in the early stages of olfactory information processing (91). Glomeruli are responsive to various odorants, and conversely, single odorants can activate multiple glomeruli (47,92–100).

The glomeruli are arranged in a chemotopic manner, where glomeruli in specific regions of the OB are preferentially activated by odors from chemical classes (e.g., amino acids, bile acids, nucleotides in fish) (47,101). Chemically related odorants generate similar, overlapping patterns of neural activity (102).

Much like glomeruli, individual mitral cells (MCs) respond to a variety of odors (18,102–107). The responses of MCs to odors are characterized by slow, temporal firing patterns that alternate between different firing behaviors depending on the odor (102,105–108). Consequently, the activity patterns across the MC population change in an odor-specific manner during stimulus presentation (109). Early in the odor response, electrical recordings from multiple MCs show that related odorants evoke similar patterns of activity (102). Over time, however, these patterns become more distinct, allowing for a more specific representation of each odor (102).

2.5 Importance of Spatio-Molecular Distribution

The spatio-molecular distribution within the olfactory bulb is essential for understanding how olfactory information is encoded and processed (69,110). Each region of the olfactory bulb exhibits a unique molecular signature that influences its specific function (14,111). These molecular signatures include patterns of gene expression, protein distribution, and cellular composition (12,67). Mapping these patterns is crucial for elucidating the organizational principles that enable the olfactory bulb to process a wide variety of odorants accurately (112).

Spatial distribution refers to the precise localization of different cell types and molecular markers within the olfactory bulb (19). This spatial organization is believed to support functional specialization, with specific regions dedicated to processing different types of olfactory information (16). Molecular distribution, on the other hand, involves the specific expression profiles of genes and proteins within these regions (113). By understanding both spatial and molecular distributions, researchers can gain insights into the fundamental mechanisms of olfactory processing and how these mechanisms might be conserved or diversified across different species (114).

2.6 Chemotopic Organization in Fish and Mice

Studies, particularly in fish, show that different subregions of the olfactory bulb encode information from different classes of odors (47). In fish, this chemotopic organization is highly pronounced: glomeruli in the dorsal part of the olfactory bulb primarily encode danger-related odors, glomeruli in the lateral part encode food-related odors such as amino acids, and glomeruli in the medial and ventral parts encode bile acids, which are important for kin recognition and social interactions (115). Chemotopic information is somewhat limited in mice, though there is evidence that dorsal olfactory bulb glomeruli, especially anterior ones, are associated with innate behaviors (16). Both zebrafish and mouse systems demonstrate that distinct classes of chemical groups are encoded in distinct spatially organized parts of the olfactory bulb, referred to as chemotopy (19).

2.7 Functional Topography and Molecular Correlates

While there is clear functional topography in the olfactory system of both species, the extent to which this functional topography correlates with the molecular features of mitral cells and underlying interneurons remains unclear. One possibility is that all mitral cells are molecularly identical, raising questions about how they organize so precisely and form functional topography, as well as how they find their targets. Alternatively, there may be general classes of mitral cells, with distinct subclasses either expressing specific markers or exhibiting combinatorial expressions of multiple genes, distinguishing them from other subclasses located in different topographical locations within the olfactory bulb.

2.8 Molecular Diversity in Mitral and Interneuron Cells

The existence of molecular differences among mitral cells is uncertain, as is the existence of such differences among interneurons, such as periglomerular cells or granule cells (19). The question arises whether there are different classes of periglomerular and granule cells and if these classes follow any molecular topography. These are the questions addressed in this master's thesis by studying the spatial transcriptomes and single-cell transcriptomes of both zebrafish and mouse olfactory bulbs. The goal is to identify whether glutamatergic and GABAergic cells have clear molecular differences correlated with their spatial locations within the olfactory bulb.

2.9 Single-Cell RNA Sequencing (scRNA-seq) Technology

Established in 2009, single-cell RNA sequencing (scRNA-seq) has become a widely used tool in biological research across diverse fields (116). This technology offers a detailed, genome-wide profile of RNA molecules within individual cells, facilitating comparisons of transcriptomes among single cells (117). This capability is essential for understanding transcriptional similarities and differences among cell groups, thus aiding in the analysis of cellular heterogeneity (117). scRNA-seq is particularly valuable for studying unique cells, such as neurons in the brain, and for gaining insights into gene expression by identifying co-regulated gene modules based on gene co-expression patterns in single cells (117).

The scRNA-seq process involves isolating cells from tissue samples and capturing individual cells within nanoscale droplets, each containing a bead with unique molecular identifiers (118). This step ensures that high-quality individual cells are extracted to obtain precise genetic data (119). Methods for cell isolation and capture vary depending on the cell and tissue type and may involve whole-cell isolation or cell-specific nuclei isolation (117). Each cell is isolated in a separate reaction mixture, allowing RNA from each cell to be processed individually (120). In this mixture, RNA transcripts are converted into complementary DNA (cDNA), with each cDNA molecule assigned a unique barcode (118).

Following barcoding, cDNA is amplified using techniques such as in vitro transcription (IVT) or polymerase chain reaction (PCR) (121). A common method for cDNA amplification is PCR amplification with SMART technology (122). To minimize amplification biases, each mRNA molecule is barcoded with unique molecular identifiers (UMIs) (123). Post-sequencing, the data is analyzed to map and categorize the gene expression landscape within a heterogeneous cell population (118).

2.10 Comparison of scRNA-seq and Bulk-cell Sequencing

Single-cell experiments provide significantly more precise data compared to bulk-cell sequencing, which averages expression levels across large groups of cells (124). Unlike bulk sequencing, which prepares libraries from thousands of cells, scRNA-seq generates cell-specific libraries, enabling detailed investigation of DNA and RNA functionalities within distinct cellular subsets (124). Traditional bulk RNA/DNA transcriptome analyses capture only the overall signal levels from tissues or organs, failing to identify variations between individual cells (118). scRNA-seq captures single cells, facilitating the detection of gene expression variability and uncovering cellular heterogeneity within a sample (125). This method also increases the likelihood of detecting rare cell populations within tissues (126).

In neuroscience, scRNA-seq is employed to map the diverse cell types in the brain and understand their functions (118). As previously mentioned, scRNA-seq is a powerful tool that identifies the molecular characteristics of various cell types by examining the gene expression of single cells extracted from brain tissue (118). Despite its advantages, scRNA-seq has limitations due to its lack of spatial context, making it challenging to fully understand the arrangement and interactions of these cells in producing specific brain functions (127).

2.11 Combining scRNA-seq with Spatial Transcriptomics

To address the limitations of scRNA-seq, it is crucial to combine its data with spatial information about the location of each cell type within the brain (127). Spatial transcriptomics techniques map the spatial organization of cell types within intact tissues (128). Over the past decade, advancements in imaging and sequencing technologies have led to the development of several spatial transcriptomics tools, enabling the creation of detailed functional atlases of the brain and providing deeper insights into neural circuits (127).

Spatial transcriptomics involves several techniques, including in situ hybridization, in situ sequencing, and multiplexed error-robust fluorescence in situ hybridization (MERFISH) (128). These methods allow researchers to visualize and quantify RNA molecules directly within their tissue context, preserving spatial information (128). The integration of spatial transcriptomics with scRNA-seq data provides a comprehensive view of gene expression, revealing how different cell types are spatially arranged and how they interact within the tissue environment (128).

Combining spatial transcriptomics with scRNA-seq has profound implications for understanding brain function (127). It allows researchers to correlate specific gene expression patterns with precise cellular locations and interactions, enhancing our understanding of neural circuits and their roles in behavior and disease (127). This integrated approach can identify spatially distinct gene expression signatures, uncovering new insights into the molecular mechanisms underlying various brain functions and pathologies (127).

3 MATERIALS AND METHODS

3.1 Zebrafish Datasets

Two datasets for zebrafish were used in this thesis. Single-cell RNA sequencing dataset, and spatial transcriptomics dataset. R programming language was used for the analysis of these datasets (R version 4.2.2) (129). The analysis was done by utilizing Seurat, Comparative Cytoarchitectonics (CoCy) (130), and clustree R toolkits (131).

3.1.1 Single-cell RNA sequencing dataset

This dataset was received from Cağhan Kızıl Lab. Five 7-month-old adult wild-type (AB) zebrafish of mixed sex were used in this study. The dissected olfactory bulbs were immediately placed in ice-cold PBS. For tissue dissociation, we utilized the Neural Tissue Dissociation Kit (Miltenyi Biotec, Cat. No. 130-092-628). Post-dissociation, the cells were filtered through a 40 μM cell strainer into 10 mL of 2% bovine serum albumin (BSA) in phosphate buffered saline (PBS) solution and then centrifuged at 300 g for 10 minutes. The resulting cell pellet was resuspended in 4% BSA in PBS. During the FACS stage, cell viability was assessed using Sytox Blue (Invitrogen, Cat. No. S34857) and Dycle Ruby (Invitrogen, Cat. No. V10309) dyes, and flow cytometry was performed. Cells were sorted on ice with minimal incubation time using a Sony MA900-FP sorter (histogram data attached). The sorted cells were immediately subjected to single-cell analysis using the 10X Chromium system, with library preparation conducted according to the manufacturer's protocol. The libraries were sequenced on an Illumina NovaSeq 6000 platform. A total of 10,488 cells were obtained. The raw sequencing data were processed using the Cell Ranger Single Cell Software Suite (10X Genomics, v6.1.2) and aligned to the zebrafish GRCz11 (release 105) genome. The Seurat object was created by excluding cells with fewer than 200 genes and genes expressed by fewer than 3 cells. In this study, the data from the five 7-month-old adult wild-type zebrafish of mixed sex do not need to be integrated because they were treated as biological replicates rather than distinct experimental groups. Each zebrafish underwent the same experimental procedures, including tissue

dissection, cell dissociation, and single-cell RNA sequencing (scRNA-seq) using identical conditions and protocols. This approach ensures that the data generated from each individual zebrafish are comparable and can be collectively analyzed to provide a comprehensive view of the olfactory bulb cell population.

3.1.2 Spatial transcriptomics dataset

The dataset used for the spatial transcriptomic analysis is named The Atlas of Zebrafish Transcriptomic Encephalic Cytoarchitecture (AZTEC). This project does not cover the generation of AZTEC data since the data is still unpublished and is currently being drafted (Bredesen-Aa, 2024). However, its generation will briefly be mentioned. Only olfactory bulb region of the AZTEC data is used in this thesis.

Adult zebrafish spatial transcriptomics data was generated using Resolve Biosciences' "The Molecular Cartography Platform" (Bredesen-Aa, 2024). This cutting-edge, highly multiplex spatial analysis technology provides detailed contextual datasets that expose molecular interactions at a subcellular level while preserving the integrity of the tissue samples. The platform offers a high-resolution view of transcriptomic activity, allowing for the analysis of hundreds of genes in a single experiment (Benedict Nilges, Resolve Biosciences). For this study, six-month-old zebrafish were euthanized in cold water, and their brains were dissected and fixed according to Resolve's guidelines. The brains were incubated in PAXgene® fixation solution, stabilized with PAXgene® tissue stabilizer, and cryoprotected in a sucrose solution before being cryo-embedded and sectioned (132). The sections were collected on precooled coverslips and sent to Resolve GmbH in Germany for Molecular Cartography analysis of Pax-gene-fixed samples. The analysis followed the manufacturer's protocol, involving probe hybridization, imaging, and color development to decode gene-specific signals (132). The Molecular Cartography sections were then aligned with the Adult Zebrafish Brain Atlas (AZBA) (133) using custom software (130). Using Cellpose (134) on DAPI stains, nuclei were identified. Nuclear transcriptomes were determined by assessing the transcript enrichment for each gene within individual nuclei, and the spatial positions of nuclei were calculated by determining the centroids of the identified nuclei (130).

As mentioned, only olfactory bulb section of AZTEC data was used in this thesis. AZTEC data includes coronal sections from 2 animals labeled as “animal.1”, and “animal.2”. These animals were 6-month-old adult zebrafishes. The integration of AZTEC data regarding two animals were not applied because no batch effects were detected (130). Although no integration method was applied, the compatibility of two animals in AZTEC was checked before after downstream analysis. AZTEC data has 99 genes, and 31904 cells. This data was divided into two in order to analyze GABAergic and glutamatergic cell types separately.

3.2 Mouse Datasets

Two datasets for mouse were used in this thesis. Single-cell RNA sequencing dataset, and spatial transcriptomics dataset. R programming language was used for the analysis of these datasets (R version 4.2.2) (129). The analysis was done by utilizing Seurat, Comparative Cytoarchitectonics (CoCy) (130), and clustree R toolkits (131).

3.2.1 Single-cell RNA sequencing dataset

The mouse single-cell RNA sequencing dataset was divided into several brain regions by Allen Brain Institute (135). The “OLF” labeled data is used in this thesis suggesting that it only contains cells belong to olfactory bulb. The file was download in h5ad format and turned into Seurat object using Seurat R toolkit (136) . Then “OLF” labeled metadata was attached to existing Seurat object. The dataset is available at the Allen Brain Cell Atlas (135).

Additionally, some cells that do not belong to olfactory bulb were excluded. The reason for this is that since this data is used to analyze whole mouse brain, some labeling of olfactory bulb cells are incorrectly done. Therefore, in class section of metadata, only "03 MOB-DG-IMN", "12 MOB-CR Glut", "28 Astro-Epen", "29 Oligo", "30 OEG", "31 Vascular", "32 Immune" sections were used. Although a lot of cells were excluded, the data still contained some cells that do not belong to olfactory bulb. Therefore, another filtering step is applied. In subclass section "279 DG-PIR Ex IMN", "291 Ependymal NN", "306 Lymphoid NN", "304 Monocytes NN", "305 DC NN" sections were excluded from the data. This dataset includes 26479 genes, and 73086 cells in total. This dataset consists of two animals, one male, and one female. The integration of these two animals were applied therefore no integration method was applied however, the pre-applied integration was checked.

3.2.2 Spatial transcriptomics dataset

This spatial transcriptomics dataset was generated using multiplexed error-robust fluorescence in situ hybridization (MERFISH) (137). The dataset is available at the Allen Brain Cell Atlas (137).

Unlike mouse single cell RNA sequencing data, MERFISH data was not divided into sections in Allen Brain Atlas therefore whole brain was downloaded in h5ad format then later transformed into Seurat object. The corresponding metadata was attached to existing Seurat object. To target olfactory bulb, in “brain_section_label” section of metadata, only “Zhuang-ABCA-1.015”, “Zhuang-ABCA-1.016”, “Zhuang-ABCA-1.017”, “Zhuang-ABCA-1.018”, “Zhuang-ABCA-1.019”, “Zhuang-ABCA-1.020” sections were used, so in general there are 6 coronal sections in the data. This data only includes one animal therefore no integration method was required. MERFISH data has 1122 genes, and 56097 cells in total. This data was divided into two in order to analyze GABAergic and glutamatergic cell types separately.

3.3 Normalization and Variance Stabilization

The first step of our analysis started with normalization and variance stabilization. Since scRNA-seq data is subject to technical variability such as cell capture efficiency, amplification biases, and sequencing depth, it is important to address to these differences in cells properly otherwise the results of the analysis might be misleading, and thus might not reflect the true biological differences between cells (138).

After normalization it is important to apply variance stabilization. Since the variability in gene expression data is not consistent across different expression levels, variance stabilization ensures that highly expressed genes do not dominate the variability structure of the data thus taking lowly expressed genes into account as well (138). Overall, both normalization and variance stabilization enhance the interpretability and reliability of analysis revealing accurate biological information thus improving the quality of upcoming clusters in the analysis (138).

Traditional normalization methods usually utilize scaling gene expression counts in order to explain differences in sequencing depth and other technical factors. Such approaches might include log-normalization, where counts are divided by the total counts per cell and then log-transformed and global-scaling normalization, where counts are scaled to a fixed median or mean value. These approaches do not explain the mean-variance relation, and usually assume that technical noise is constant across genes. Therefore, such methods might lead inaccuracies in analyses.

Therefore, SCTransform method was used in both species in this project. SCTransform is a method developed by Seurat R toolkit developers. On the contrary of traditional methods, SCTransform is a more sophisticated technique which models gen e expression counts with a generalized linear model (GLM). SCTransform explains mean-variance relationship in the data by fitting a negative binomial regression model to each gene, with observed counts modeled as a function of sequencing depth and other technical covariates. Furthermore, the analysis of varying sequencing depths and technical noise is more reliable using SCTransform due to

including a regularization step to stabilize variance across different expression levels (138).

SCTransform by default chooses 3000 variable genes. By using this method, 3000 genes out of 21477 genes were chosen as variable genes in the zebrafish scRNA-seq dataset. All 99 genes were chosen as variable genes in the AZTEC dataset (Figure). 3000 genes out of 26479 genes were chosen as variable genes in the mouse scRNA-seq dataset. All 1122 genes were chosen as variable genes in the MERFISH dataset.

3.4 Principal Component Analysis

Principal component analysis (PCA) is a statistical technique that identifies significant patterns and trends in multidimensional data. By focusing on the variables that exhibit the highest variance, PCA transforms a large set of informative variables into a smaller set of principal components (PCs). The first principal component captures the maximum variance in the data, while the second principal component captures the next highest variance independently of the first (139).

The process of PCA starts with standardizing the data to ensure all features are on a common scale. Then, it computes the covariance matrix to uncover relationships between the variables. From this covariance matrix, eigenvectors and eigenvalues are extracted, where the eigenvectors represent the principal components, and the eigenvalues indicate the magnitude of variance captured by each component. Finally, the original data is projected onto the new space defined by these principal components (139).

PCA is effective in reducing the dimensionality of a dataset while preserving as much essential information as possible. By emphasizing high variance, PCA helps filter out insignificant fluctuations or noise, leading to cleaner and more interpretable data. Additionally, PCA facilitates the reduction of data to two or three principal components, enabling visual examination through plotting. This simplification enhances dataset exploration, visualization, and subsequent analyses (139).

Moreover, PCA is widely used for feature reduction, data compression, and as a preliminary step in machine learning pipelines. It helps in identifying hidden patterns and simplifying complex datasets, making it easier to perform clustering, classification, and regression tasks. Overall, PCA is a powerful tool for improving data interpretability and efficiency in various analytical workflows.

In this project, in order to determine the optimal PC numbers, ElbowPlot() function in Seurat package was used. PCs are ranked in elbow plot based on their percentage of variance, and then function generates a plot with PCs' standard deviations. The elbow point in this plot indicates the optimal number of PC because after the elbow threshold, the analysis will not benefit much from increasing the number of PCs since the standard deviation remains almost the same (136). Therefore, first 20 PCs were chosen for zebrafish scRNA-seq data. First 30 PCs were chosen for both mouse scRNA-seq data, and MERFISH data. However, PCA was not used for AZTEC data because AZTEC data has 99 genes only. Because applying PCA to AZTEC data may potentially lead to loss of information (140).

3.5 Finding Neighbors and Clusters

In order to detect different types of clusters and cell types, Seurat v5 R toolkit was used (136). For this purpose, the "FindNeightbors()" function was used. This function identifies clusters of similar cells by constructing a shared nearest neighbor (SNN) graph (136). This process works in two steps. Firstly, k-nearest neighbor (K-NN) graph is calculated for cells based on their feature space (136). This feature space is derived from PCA (136). Secondly, it refines this graph into an SNN graph, where edges between cells are weighted by the number of shared nearest neighbors (136). This graph structure is needed for downstream clustering and visualization since the processed information here will be used in Uniform Manifold Approximation and Projection (UMAP) for visualization purposes subsequently (136).

After this step, in order to identify clusters of similar cells, the “FindClusters()” function was used (136). This clustering processed uses a community detection method, in this case it is Louvain algorithm. Louvain algorithm iteratively optimizes the modularity of the network partitioning (136). With this way, the cells are grouped into clusters according to their connectivity in the SNN graph (136). There is also an important parameter called resolution. This parameter is used to fine-tune the cluster granularity meaning that lower values yielding less cluster and vice versa (136).

3.6 Choosing Optimal Resolution via Clustree

Selection of optimal resolution parameter is critical. Resolution parameter changes the granularity of clusters making them divide depending on resolution number used. If lower numbers are used, some biological information might be missed. On the other hand, if high numbers are used, unstable tiny clusters might be observed.

To select the optimal number of resolution, the “clustree()” function in clustree R toolkit was used (131). This function creates a tree-like diagram where each node represents a cluster at a particular resolution (131). The edges between nodes in tree-like diagram show how one cluster changes as the resolution are increased or decreased (131). Some cluster may split into different clusters, and some may merge and form bigger clusters (131). With that tree-like diagram formation, we can visualize, and have an idea about the stability of the clusters (131). Therefore, this function aids user to detect the resolution number in which clusters remain stable (131).

In this project, the resolution parameters of the data were selected as 0.8,0.9 0.6,0.6 for zebrafish scRNA-seq, AZTEC, mouse scRNA-seq, and MERFISH data respectively. However, MERFISH data was subdivided into glutamatergic and GABAergic data to analyze glutamatergic cells separately. In this only glutamatergic data resolution number was chosen as 0.4. Additionally, all over cluster analyses were done using 3.0 resolution.

3.7 Taxonomy and Visualization

Normally, in Seurat R toolkit, clusters are ordered based on their size (136). However, in this project clusters were ordered based on their genetic similarity by using CoCy R tool kit (130). This process is done by computing the centroid, which is the average gene expression profile of all the cells within that cluster. These centroids are compared to one another after the centroids for all clusters are calculated. Then hierarchical clustering is applied to centroids (130).

Since single cell RNA sequencing analysis uses high-dimensional data, its visualization is not ultimately correct. Because visualization methods reduce this high-dimensional data information into 2 dimensions thus making it interpretable for the user. In this project Uniform Manifold Approximation and Projection (UMAP) technique was used for visualization by utilizing "RunUMAP()" function in Seurat v5 R toolkit (136). UMAP optimizes the layout of points (cells) in the lower-dimensional space to preserve the local and global structure of the data. The result is a scatter plot where cells that are similar in their gene expression profiles are placed closer together, allowing for intuitive visualization of cell clusters and relationships. This aids in identifying distinct cell populations and understanding their heterogeneity (141).

To relate spatial information to molecular diversity, a graph called "Correlation vs Distance" was used by utilizing CoCy R toolkit. (130). The working principle of this graph is that each cell in the dataset has spatial location coordinates, and for each pair of cells, the function calculates the correlation between their feature vectors (e.g., gene expression) and the Euclidean distance between their spatial positions. Afterwards, the distances are plotted in x-axis while the pair-wise correlations are plotted in y-axis. Correlation vs Distance graph helps us understand how feature similarity changes with physical distance, revealing spatial dependencies and patterns.

Spatial images for AZTEC and MERFISH datasets were created by using CoCy R toolkit (130) “CoCyPlotAnatomicalClusterLocation2DMulti” function was used to generate 2D plots by utilizing cell coordinates (130). Additionally, “CoCyPlotEmbedding3D” function was used to create 3D plots in which cells are assigned to a different coloring based on their spatial positions and are represented in UMAP (130).

3.8 Differential Expression Testing

Next step is differential expression (DE) analysis. This analysis identifies differentially expressed genes between clusters. For this purpose, the “FindAllMarkers()”function was used in Seurat v5 R toolkit (136). This function uses statistical test, in this case Wilcoxon rank-sum test. Wilcoxon rank-sum test evaluates whether the expression levels of a gene in one cluster are significantly different from those in other clusters (136). Wilcoxon rank-sum test was preferred because it was known that it performs better than other DE tests for population-level RNA-seq studies with large sample sizes (142). Once DE genes are identified, the “DoHeatmap()” function is used to visualize the top 10 marker genes for each cluster (Hao et al., 2024). This function generates a heatmap where rows represent genes, columns represent cells, and color intensity indicates expression levels. The heatmap provides an intuitive visualization of how these top genes are differentially expressed across clusters, highlighting the unique gene expression profiles that define each cluster (136). Using the “FeaturePlot()” in Seurat v5 R toolkits, selected marker genes ,and their expressions are shown in UMAP (136). Moreover, the gene expression values were also shown in spatial plots created by using the “CoCyPlotAnatomicalFeature2DMulti()” function in CoCy R toolkit (130).

3.9 Dataset Integration and Alluvial Plots

In order to compare and relate the clusters in different datasets, their integration is critical. For this purpose, CoCy R toolkit was used to verify spatial transcriptomics datasets clusters by comparing them with scRNA-seq transcriptomics datasets clusters. (130). Moreover, it is also possible to compare different species in order to explain potential homologies in between (130).

Since it is not possible to integrate different species without identifying orthologous groups of genes, Evolutionary genealogy of genes: Non-supervised Orthologous Group (EggNOG) was used (143). The process starts with the identification of orthologous groups, which are sets of genes that have evolved from a common ancestral gene and typically retain similar functions across different species. By clustering genes into these groups, EggNOG captures the evolutionary history and functional conservation of genes across a wide range of organisms (143). Afterwards, EggNOG-mapped versions of the Seurat objects were created with the EggNOG mapping code that is included in CoCy (130).

In this project Seurat's anchor-based Canonical Correlation Analysis (CCA) integration was a method used to integrate multiple single-cell RNA-seq datasets (136). It is known that CCA method works efficiently compared to other integration methods. This approach first identifies anchors, which are pairs of cells from different datasets that are similar to each other based on their gene expression profiles (136). These anchors are identified through canonical correlation analysis, which finds shared sources of variation between the datasets (136). Once anchors are established, they are used to transform the datasets into a common integrated space, effectively correcting batch effects and ensuring that similar cell types from different datasets are aligned (136).

In order to correlate clusters in between integrated datasets The K Mutual Ratio (KMR) method was used (130). The K Mutual Ratio (KMR) method quantifies the overlap between two datasets by analyzing their nearest neighbors. Initially, the k nearest neighbors (KNNs) between the datasets are identified. For each sample in dataset A, its KNNs in dataset B are found. These KNNs are then filtered to retain only those samples in B that also have at least one KNN in A with the same cluster label. These filtered samples are counted as mutual neighbors. For a cluster label lA in dataset A and lB in dataset B, mA represents the number of samples in A with label lA that have at least one mutual neighbor with label lB. Similarly, mB counts the samples in B whose KNNs in A have a KNN in B with the same label. The KMR is calculated by summing mA and mB, providing a measure of the mutual overlap between clusters in the two datasets. This method involves calculating the KMR for each combination of cluster labels and normalizing the resulting matrix, producing a matrix of ratios of matches between clusters. This matrix helps to understand how well the samples in each dataset correspond to each other based on their nearest neighbors, offering a robust metric for evaluating dataset integration and cluster similarity (130).

Other than matrices, CoCy R toolkit is able to create alluvial plots to visualize the correlations in between datasets (130). Alluvial plots are generated based on a cluster comparison matrix M for two datasets. For an alluvial plot of two datasets A and B, two columns are drawn on the left and right, respectively. Each column is divided into rows that represent cluster labels. For each combination of clusters i and j from dataset A and B, respectively, a connection is drawn between clusters i and j. The thickness of each connection is calculated on each side as the value in comparison matrix M, normalized by the sum of comparisons for the cluster compared on each side. In order to reduce crossing of connections, CoCy sorts clusters in one dataset by means of hierarchical clustering, and then sorts clusters of the other dataset by assigning ownership for clusters based on highest matching. CoCy implements a modular framework for generating alluvial plots that simultaneously enables the visualization of various dataset features, including gene expression dot plots and cell spatial locations (130).

4 RESULTS

4.1 Analysis of scRNA-seq and Spatial Transcriptomics Datasets Revealed Multiple Distinct Cell Types of Olfactory Bulb in Zebrafish and Mouse

4.1.1 Analysis of zebrafish datasets

4.1.1.1 Analysis of zebrafish single-cell RNA sequencing dataset

Firstly, SCTranform normalization was applied, and variable genes across the dataset were detected, and first 3000 genes were selected as variable genes. Top 10 most variable genes were highlighted which are *apoc1, fabp7a, fabp11a, ccl34b.1, cd74a, zgc:92066, apoeb, matn3b, si:ch211-214p16.1, L0C100149863* (Figure 3).

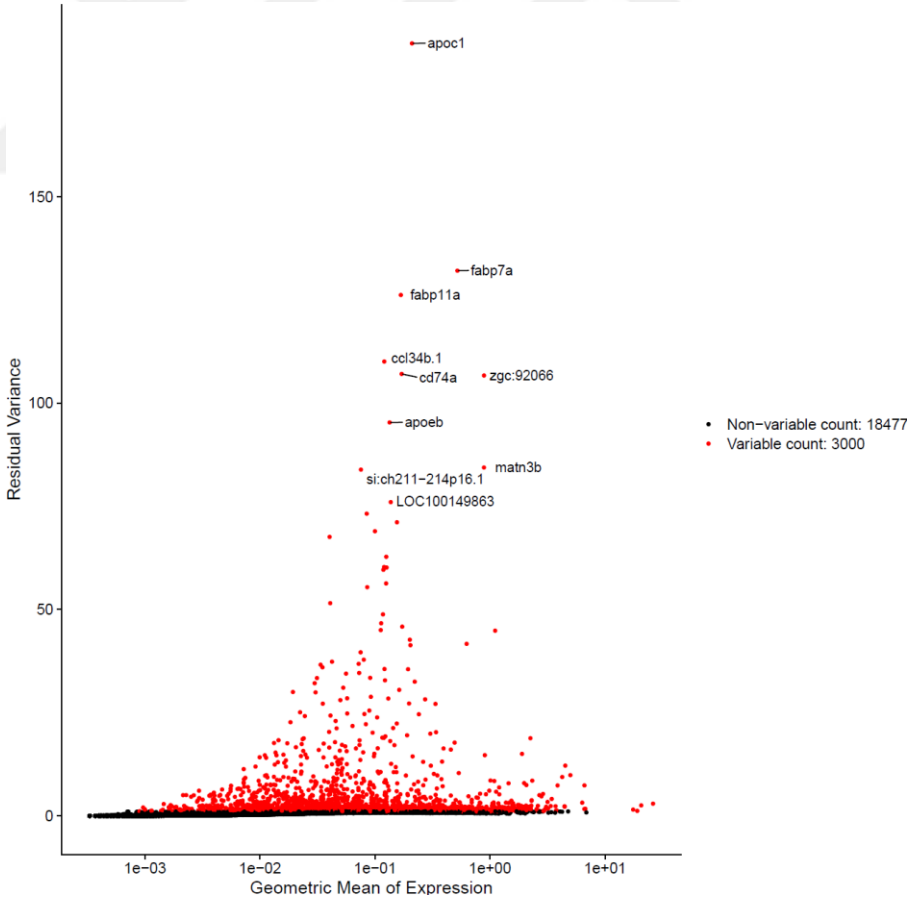

Figure 3. Analysis of the zebrafish scRNA-seq variable genes. Red dots representing variable features, and black dots representing non-variable genes.

After normalization step, Principal Component Analysis (PCA) was applied to the dataset. The number of Principle Components (PCs) used was determined considering the elbow plot, and first 20 PCs were selected (Figure 4).

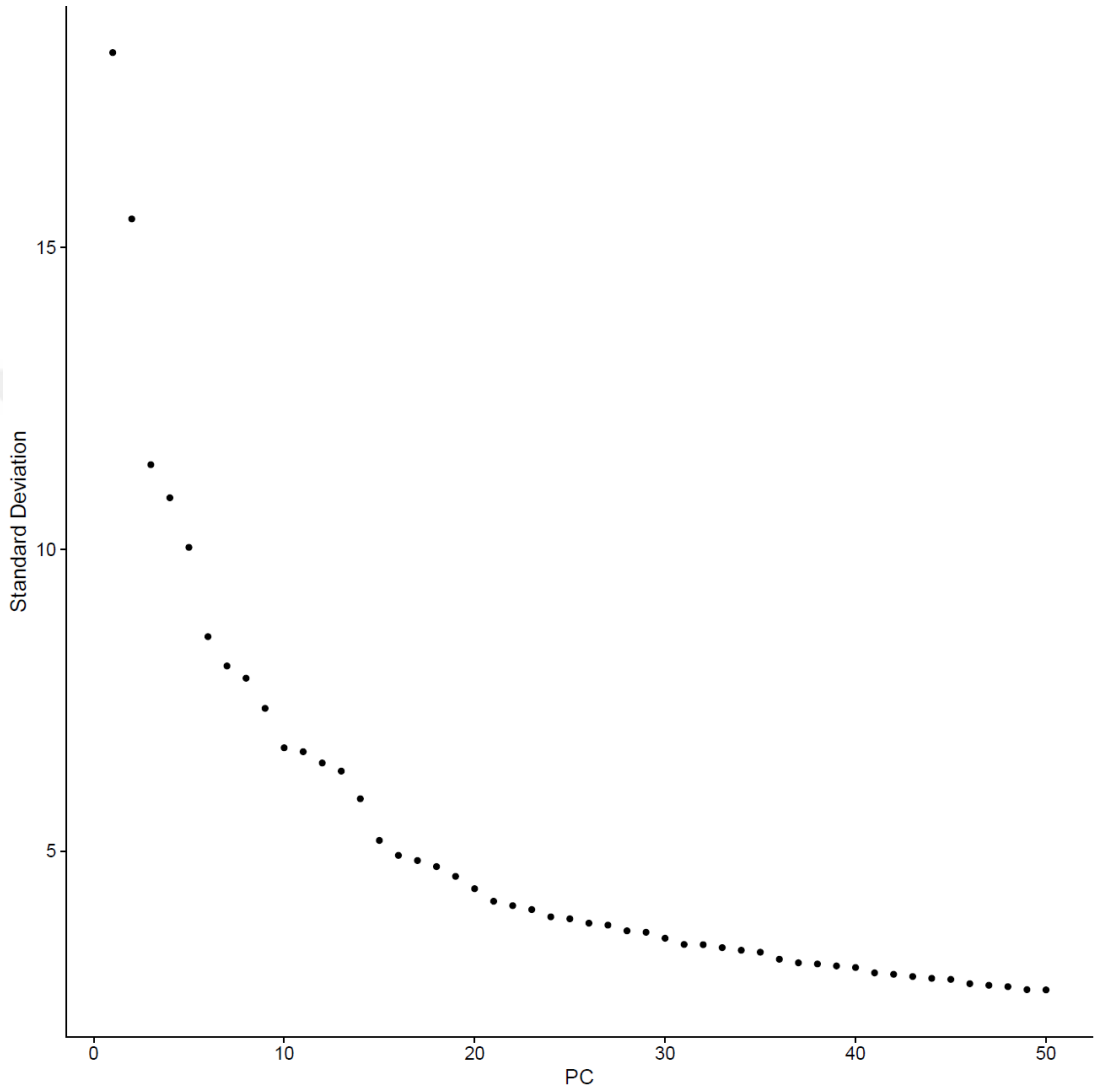

Figure 4. Zebrafish scRNA-seq PCs vs standard deviation elbow plot

After constructing shared nearest neighbor (SNN) graph, in order to identify clusters of similar cells, Louvain community detection method was applied, and resolution parameter was detected as 0.8 using clustree. 24 clusters were identified in total (Figure 5).

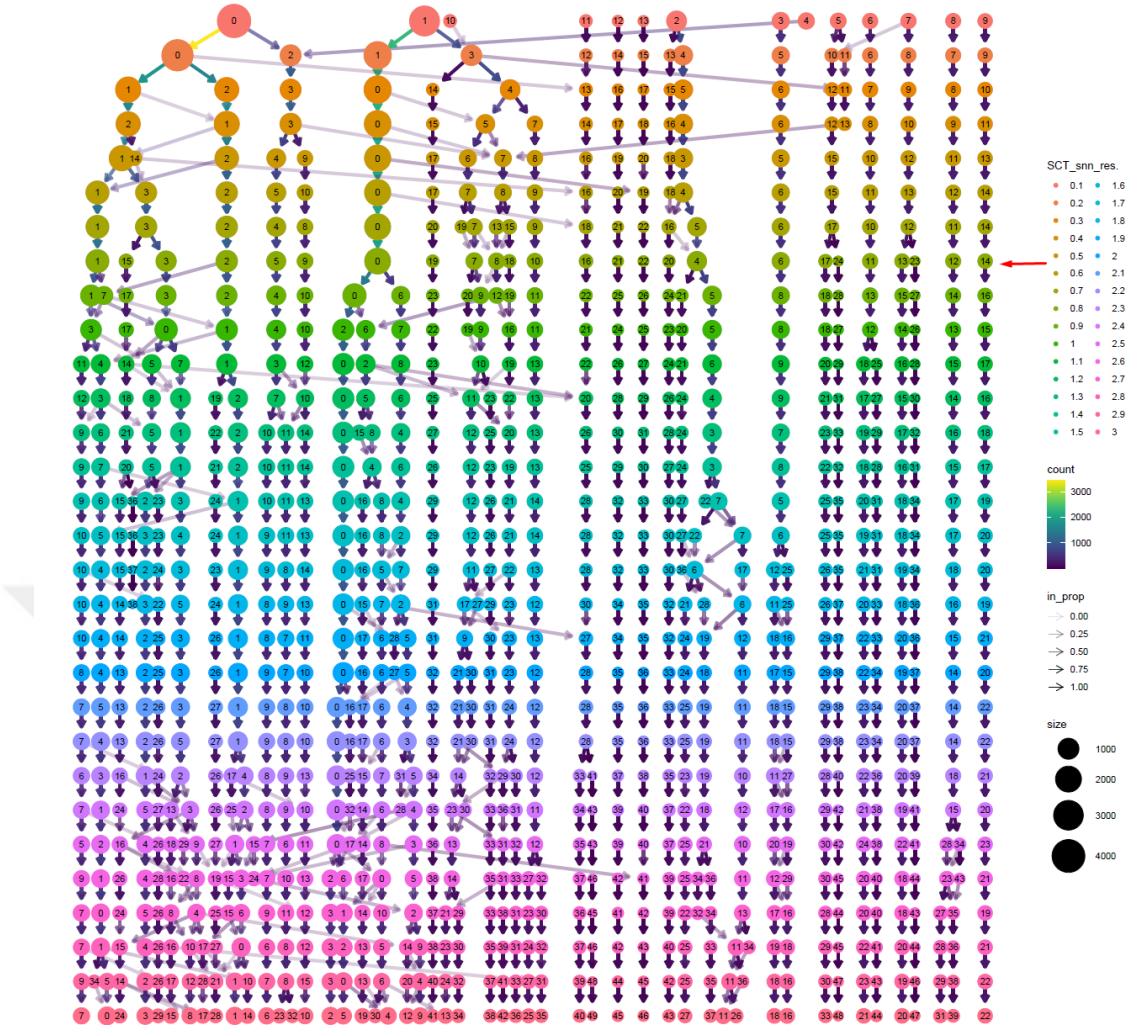

Figure 5. All zebrafish scRNA-seq clusters ranging from 0.1 until 3.0 resolution. Red arrow represents selected resolution which is 0.8. Circle size represents the largeness of the clusters. Gray arrows represent in what percentage one cluster merge or subdivide to another one in the next resolution.

UMAP was constructed, and the expressions of GABAergic marker gene (*slc32a1*), and glutamatergic marker genes (*slc17a6a*, *slc17a6b*, *slc17a7a*) were shown in FeaturePlot (Figure 6, 7).

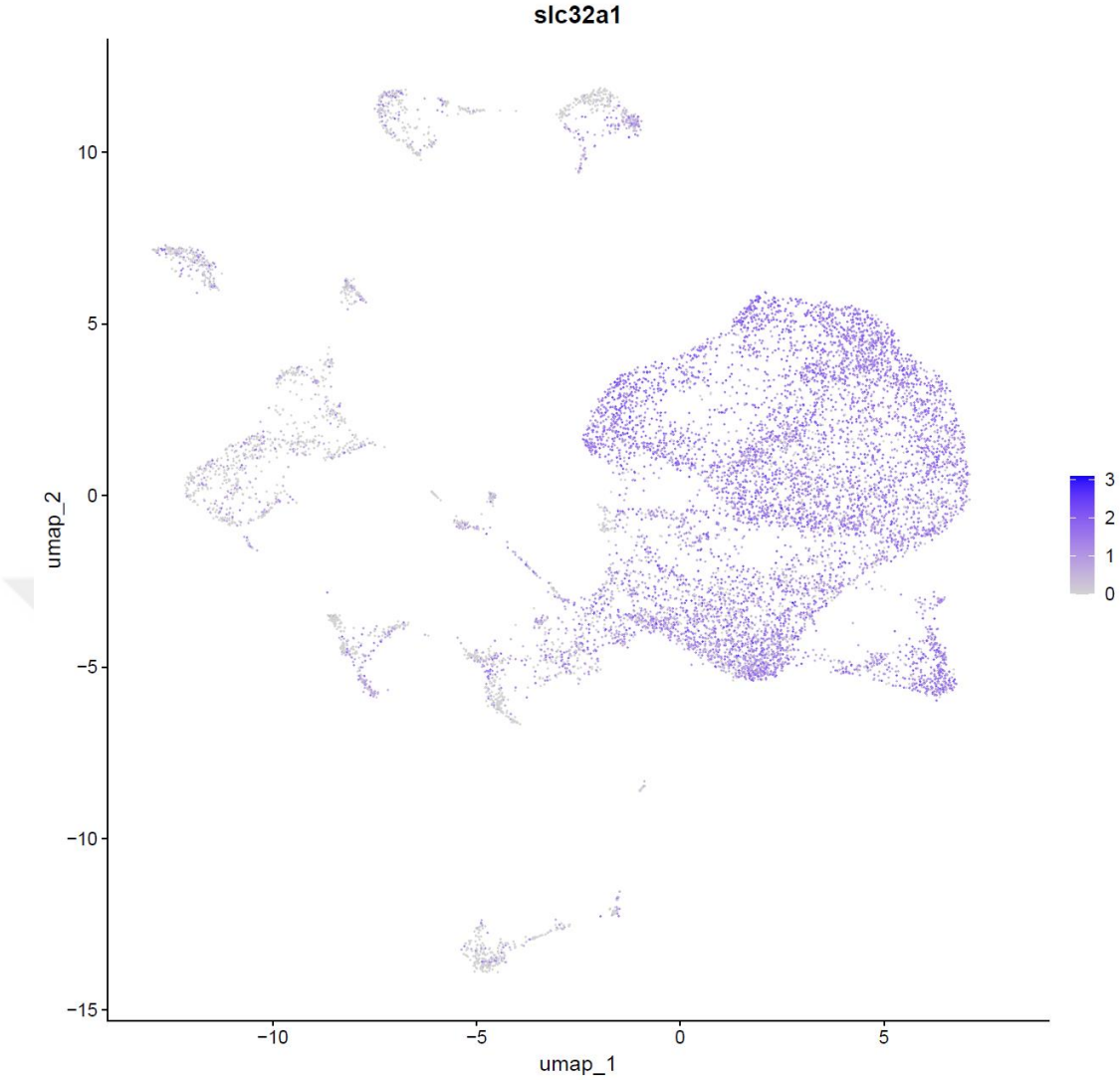

Figure 6. The zebrafish scRNA-seq gene expression of *slc32a1* gene in FeaturePlot.

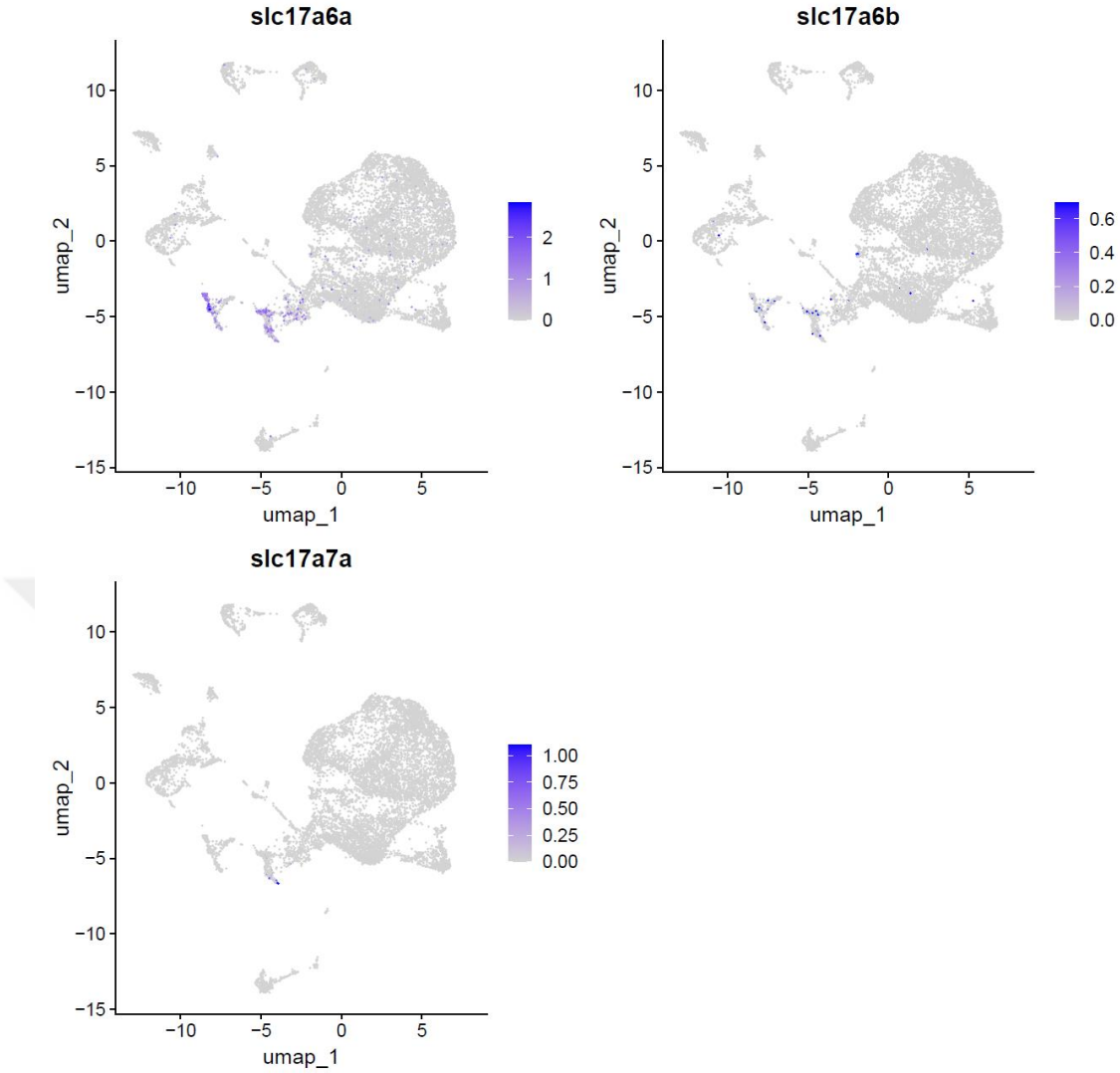

Figure 7. The zebrafish scRNA-seq gene expressions of *slc17a6a, slc17a6b, slc17a7a* genes in FeaturePlots.

Clusters were sorted based on their genetic similarity using CoCy R toolkit. Sorted clusters then represented in UMAP. Thus, 9 non-neuronal, 13 GABAergic, 1 GABA-Glut, and 2 glutamatergic clusters were identified (Figure 8). Since non-neuronal cells are not in the scope of this project, they were excluded from the analysis.

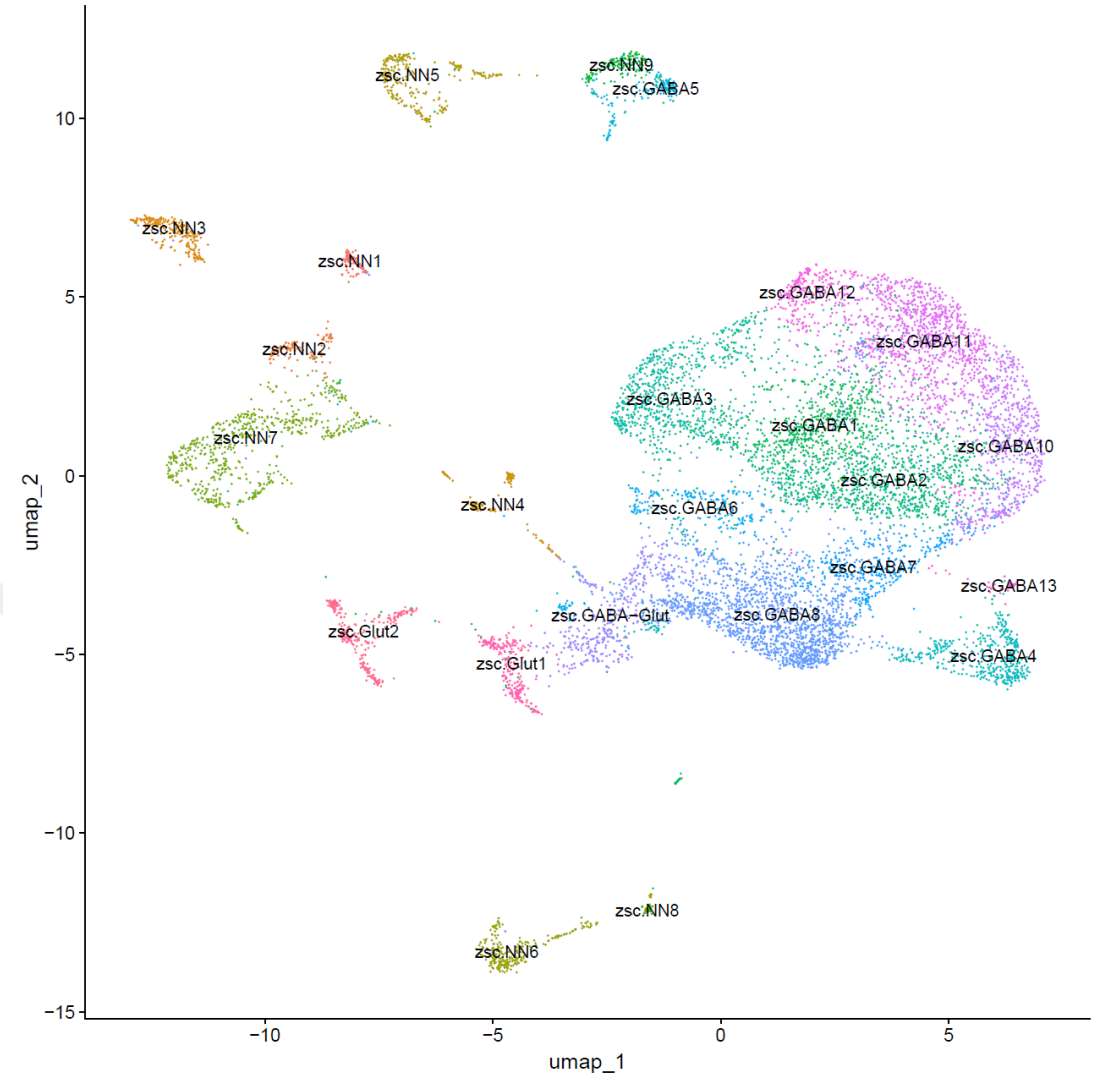

Figure 8. All zebrafish scRNA-seq clusters represented in UMAP; these clusters belong to zebrafish scRNA-seq therefore naming prefix is “zsc”.

Differential expression analysis was applied on genes using Wilcoxon rank-sum test. Top 10 most expressed genes for each cluster were shown in heatmap (Figure 9).

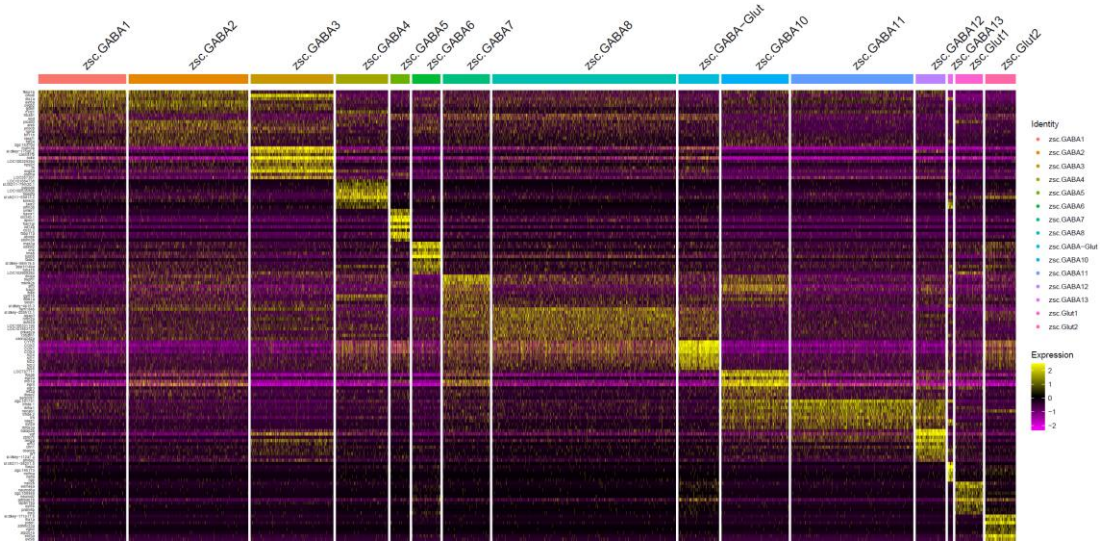

Figure 9. Top 10 most expressed zebrafish scRNA-seq genes for each cluster represented on Heatmap.

4.1.1.2 Analysis of zebrafish spatial transcriptomics (AZTEC) dataset

Firstly, SCTranform normalization was applied, and variable genes across the dataset were detected, and all 99 genes were selected as variable genes. Top 10 most variable genes were highlighted which are *npy*, *slc17a6a*, *cldn5a*, *glula*, *id1*, *cx43*, *tac1*, *snap25a*, *smox*, *gad1b* (Figure 10).

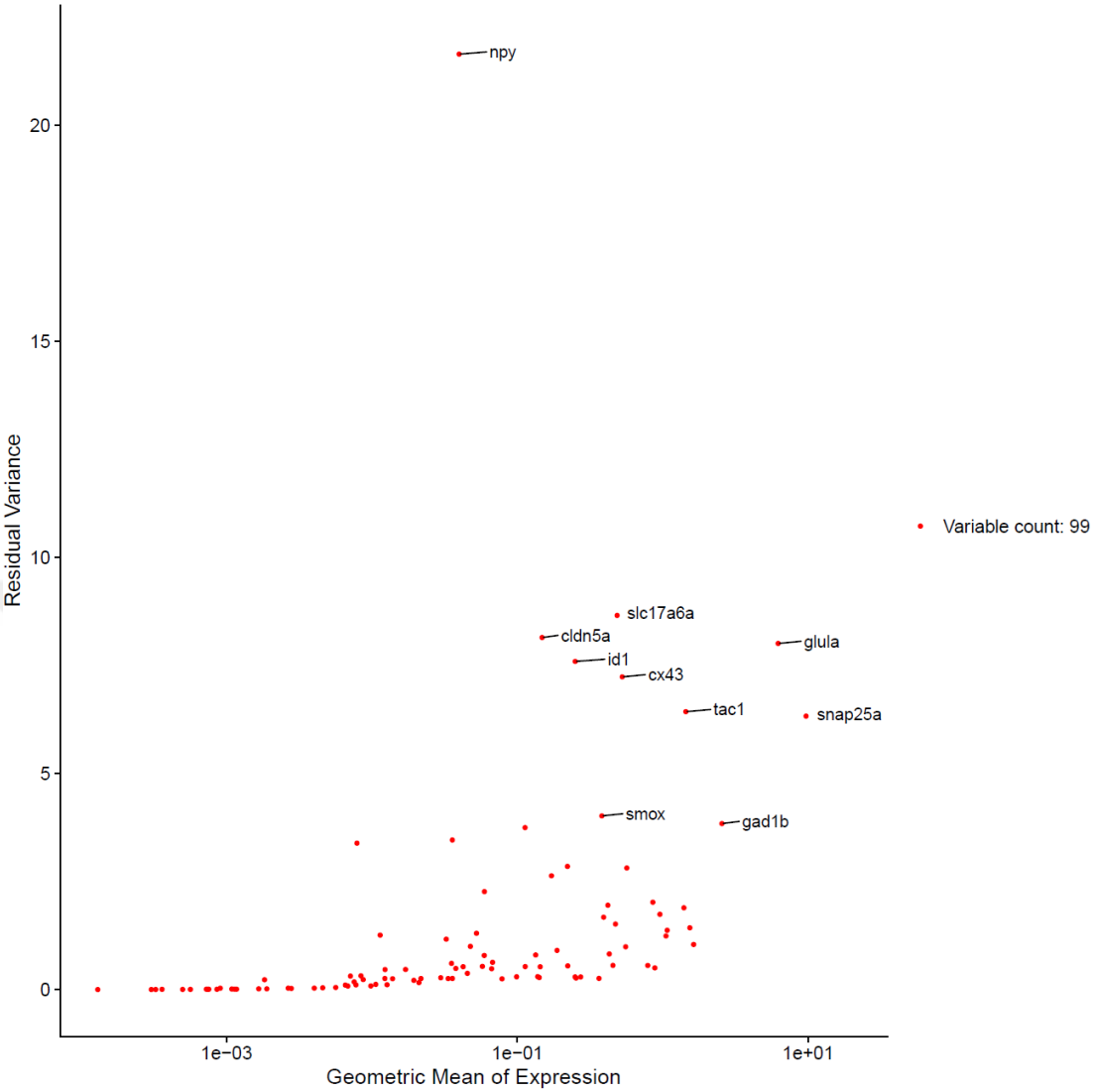

Figure 10. Analysis of the AZTEC variable genes. Red dots representing variable features.

Principal Component Analysis (PCA) was not applied to AZTEC data due to low number of genes. After constructing shared nearest neighbor (SNN) graph, in order to identify clusters of similar cells, Louvain community detection method was applied, and resolution parameter was detected as 0.9 using clustree. 17 clusters were identified in total (Figure 11).

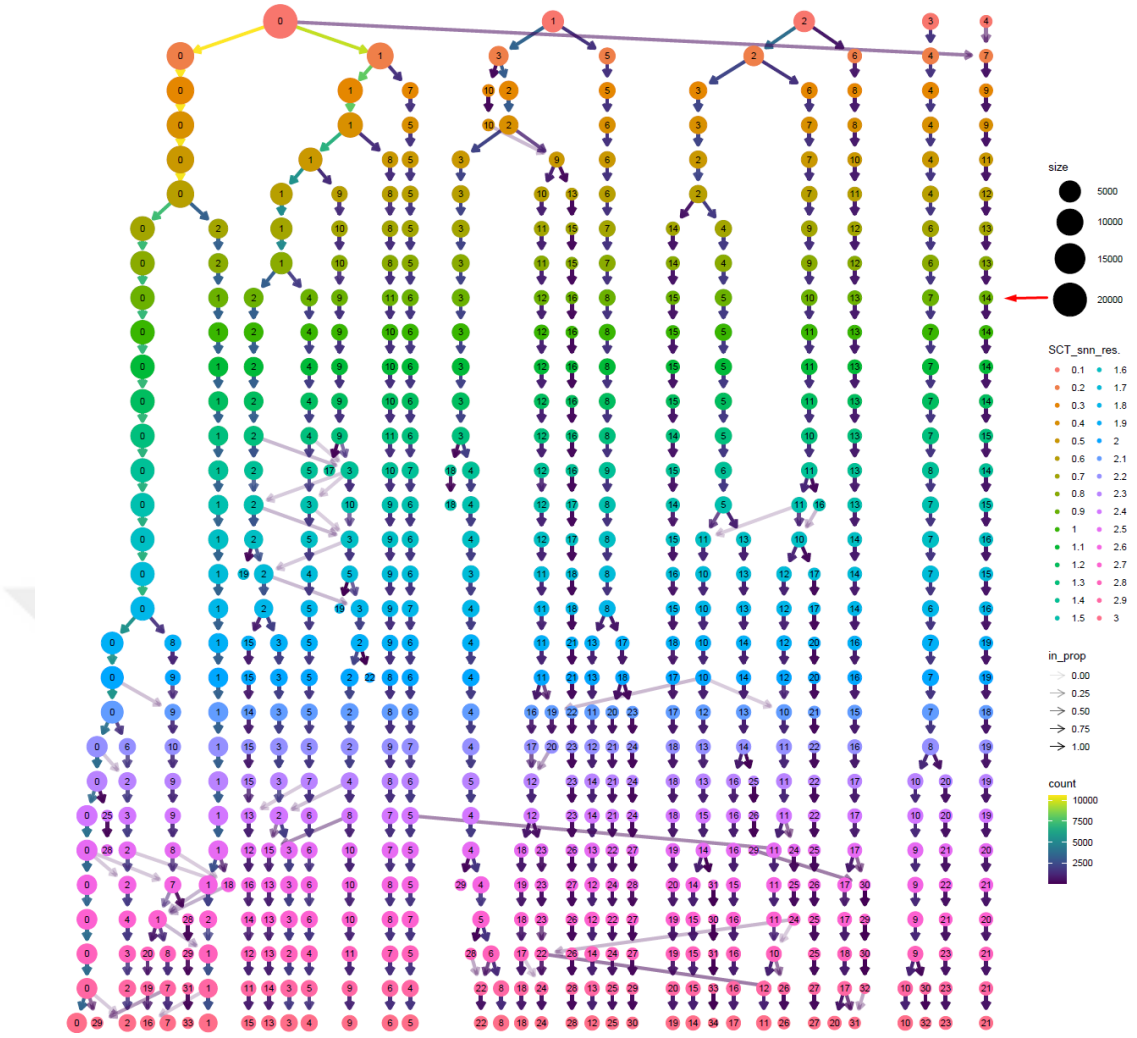

Figure 11. All AZTEC clusters ranging from 0.1 until 3.0 resolution. Red arrow represents selected resolution which is 0.9. Circle size represents the largeness of the clusters. Gray arrows represent in what percentage one cluster merge or subdivide to another one in the next resolution.

UMAP was constructed, integration of two animals was observed (Figure 12). Afterwards, the expressions of GABAergic marker gene (*slc32a1*), and glutamatergic marker genes (*slc17a6a, slc17a6b*) were shown in FeaturePlot (Figure 13, 14).

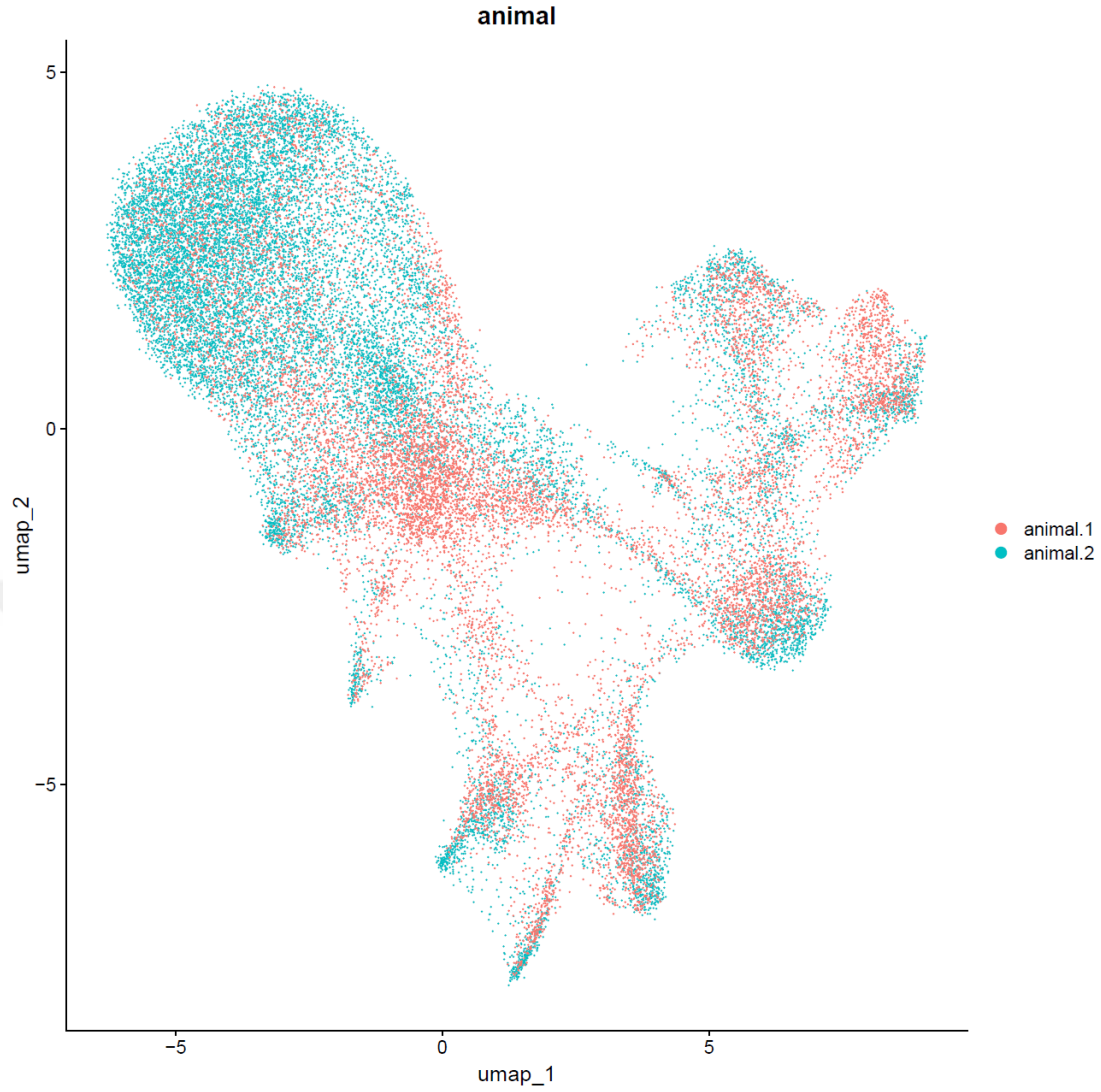

Figure 12. Integration of AZTEC animals.

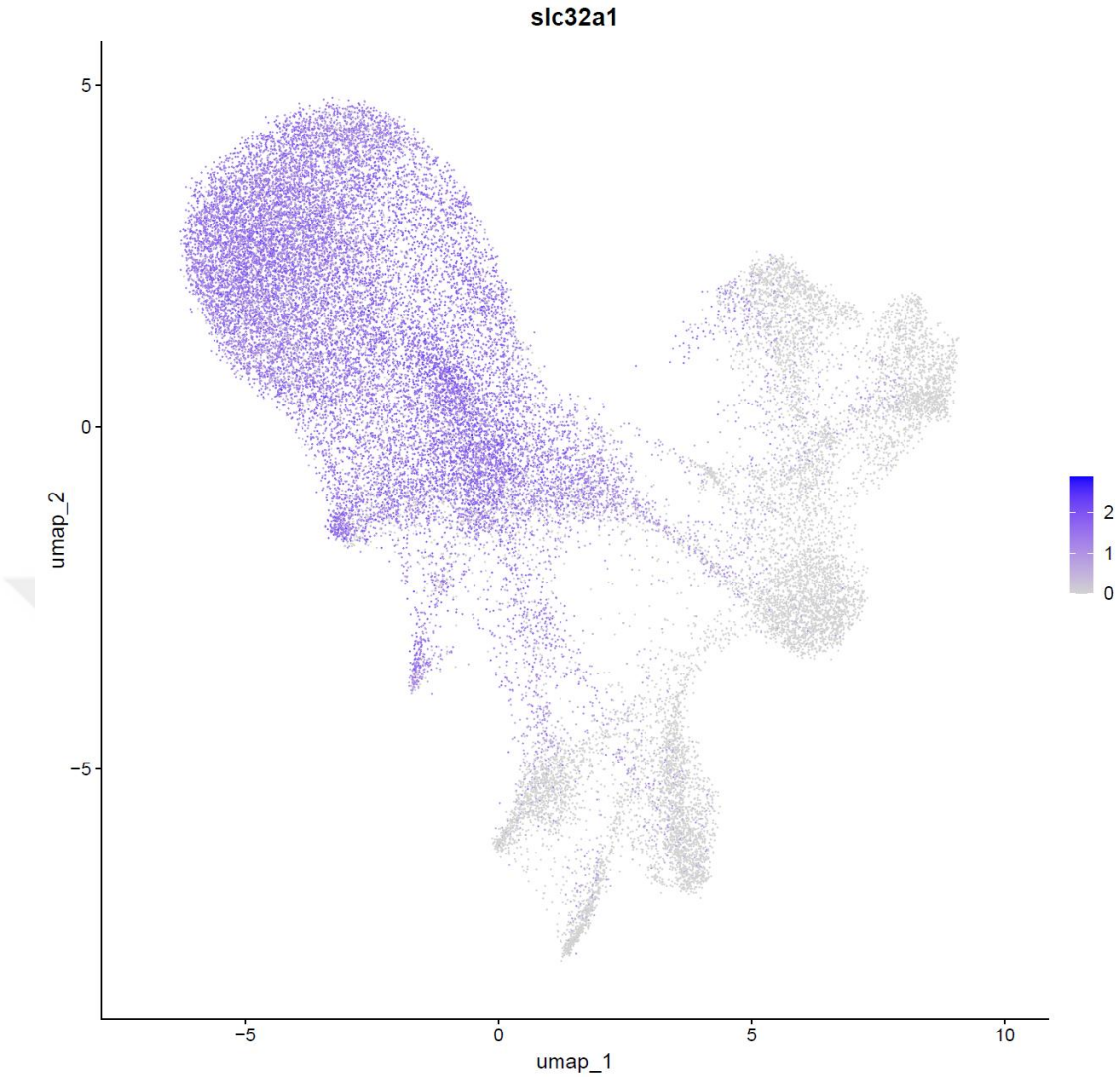

Figure 13. The AZTEC gene expression of *slc32a1* gene in FeaturePlot.

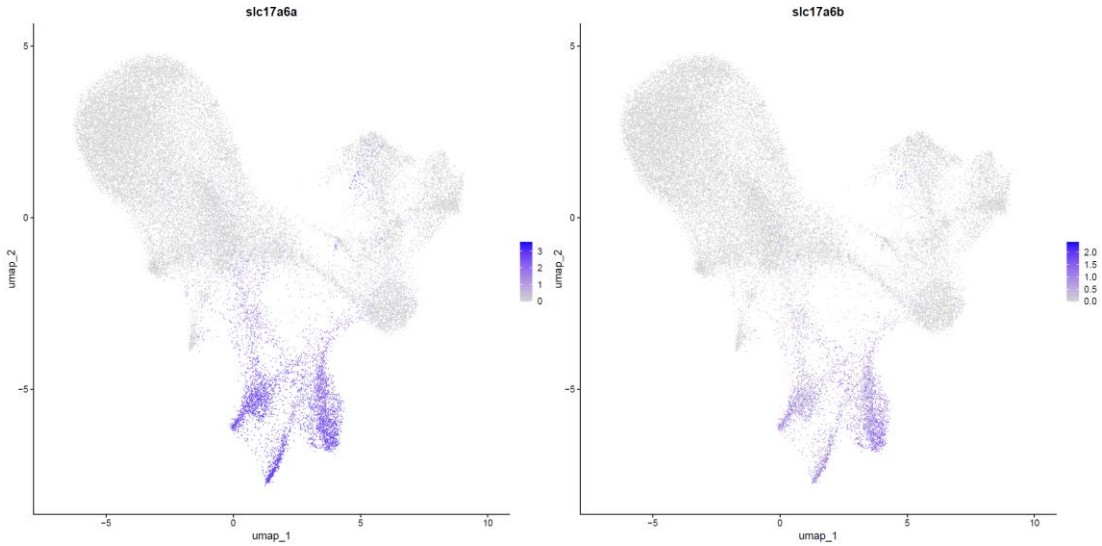

Figure 14. The AZTEC gene expressions of *slc17a6a, slc17a6b* genes in FeaturePlots.

Clusters were sorted based on their genetic similarity using CoCy R toolkit. Sorted clusters then represented in UMAP. Before annotation of the clusters, their 2D spatial plots were generated in order to classify them according to both spatial and molecular information. According to this spatial information, cluster that abundantly express GABAergic marker genes, and present at outer layer of olfactory bulb were named as periglomerular cells (PGCs) whereas inner cells labeled as granule cells (GCs). Moreover, every cluster that abundantly expressed glutamatergic marker genes were named as mitral cells (MCs). Thus, 5 non-neuronal, 8 GABAergic (2 granule clusters, 5 periglomerular clusters, and 1 immature GABA cluster), 1 GABA-Glut, and 3 glutamatergic clusters (named as mitral cell clusters for now) were identified (Figure 15,16.17). Since non-neuronal cells are not in the scope of this project, they were excluded from the analysis.

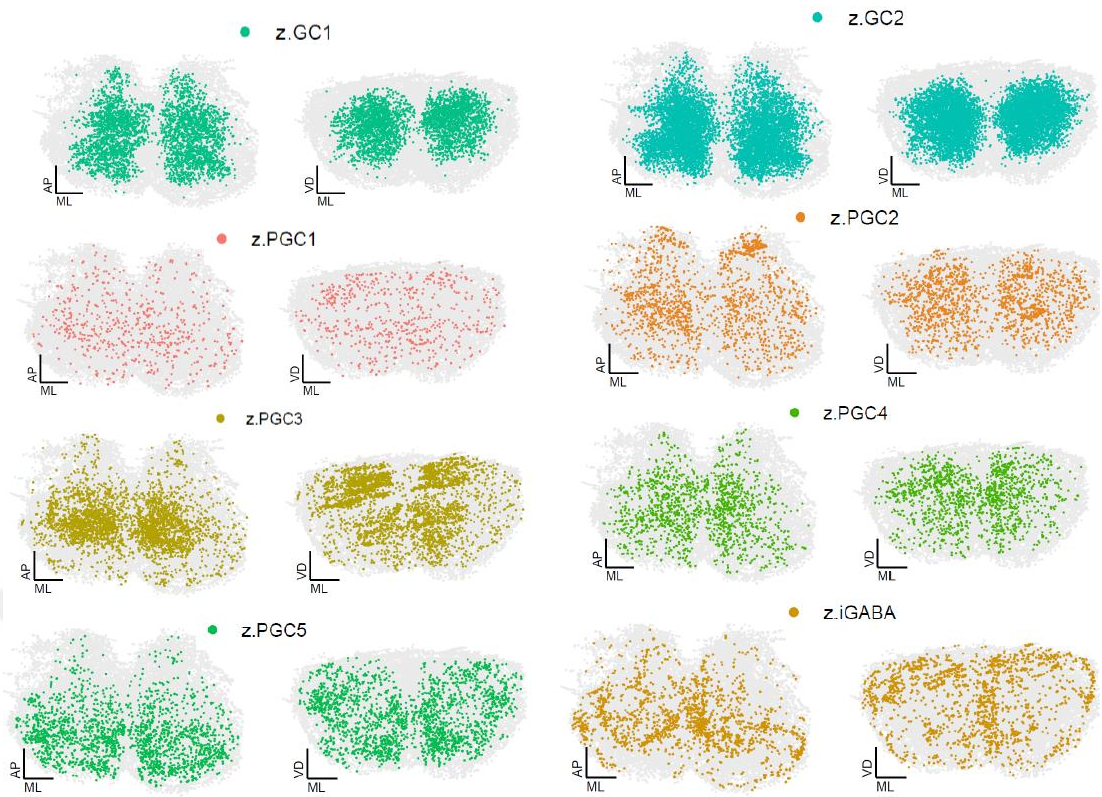

Figure 15. 2D spatial locations of AZTEC GABAergic clusters. GC: Granule Cells, PGC: Periglomerular Cells, iGABA: Immature GABAergic Cells. Scale is 100-micron length. AP: Anterior Posterior, ML: Medial Lateral, VD: Ventral Dorsal.

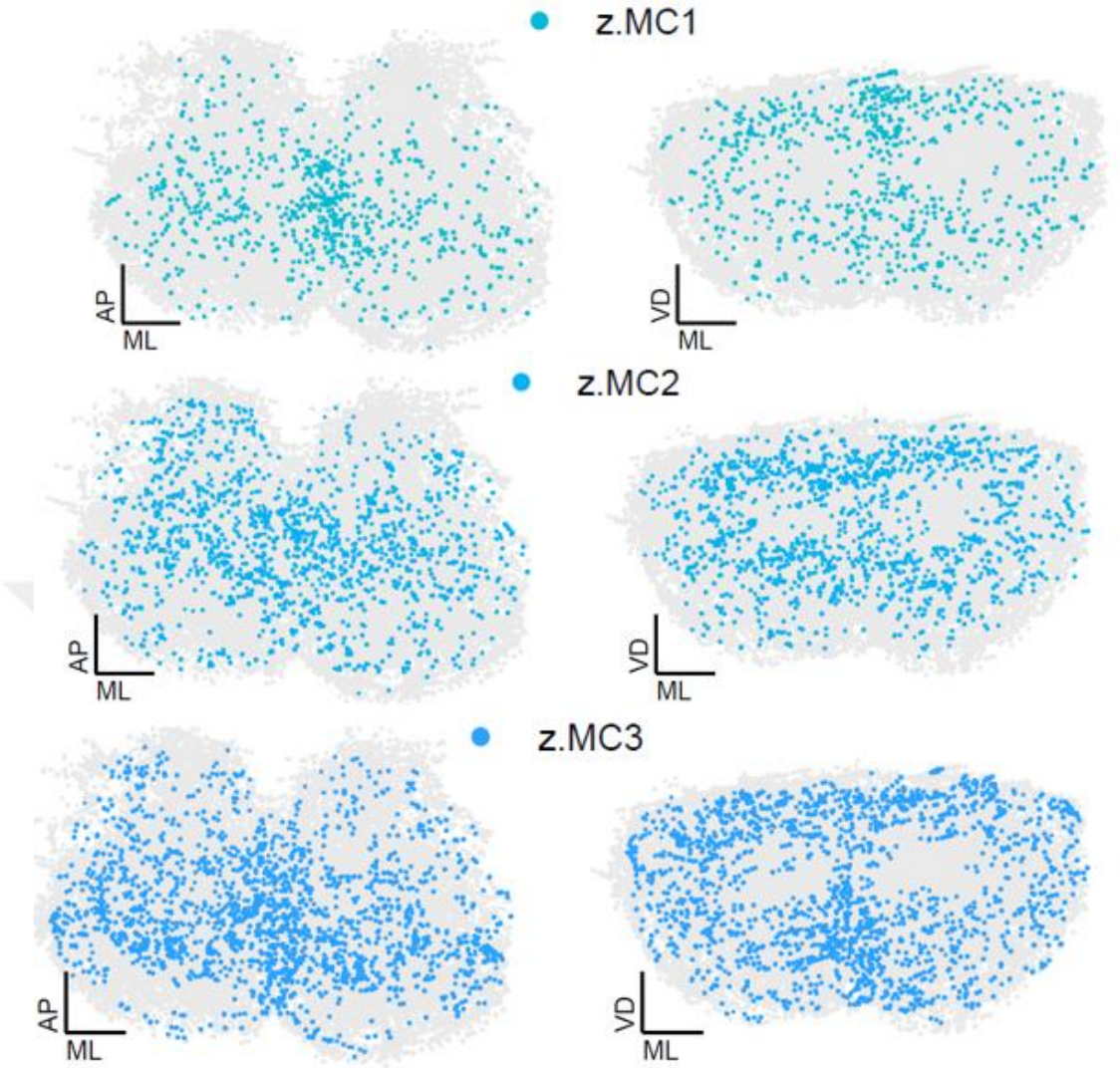

Figure 16. 2D spatial locations of AZTEC glutamatergic clusters. MC: Mitral Cells. Scale is 100-micron length. AP: Anterior Posterior, ML: Medial Lateral, VD: Ventral Dorsal.

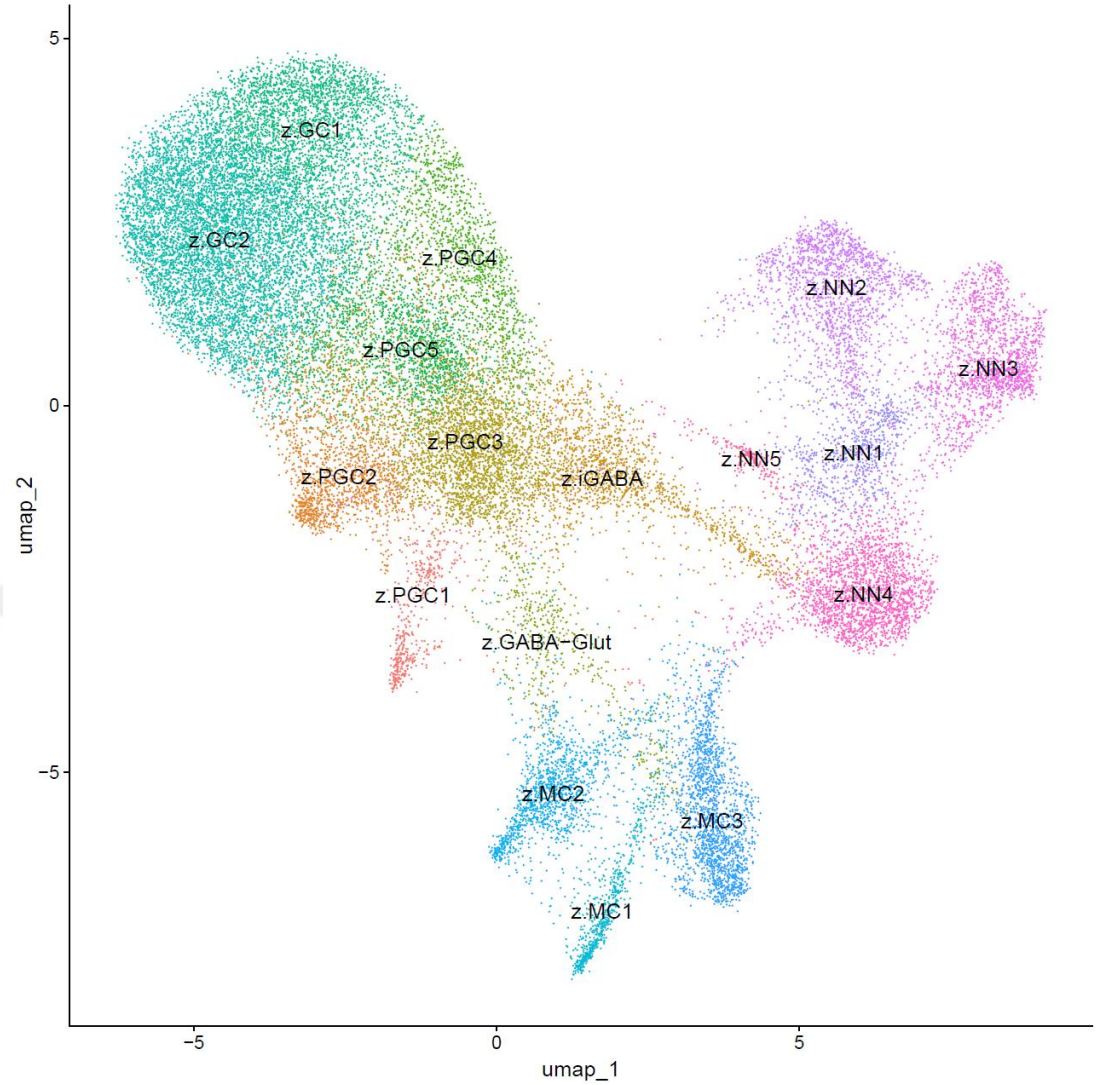

Figure 17. All AZTEC clusters represented in UMAP; these clusters belong to zebrafish spatial transcriptomics (AZTEC) dataset therefore naming prefix is “z”.

Differential expression analysis was applied on genes using Wilcoxon rank-sum test. Top 10 most expressed genes for each cluster were shown in heatmap (Figure 18).

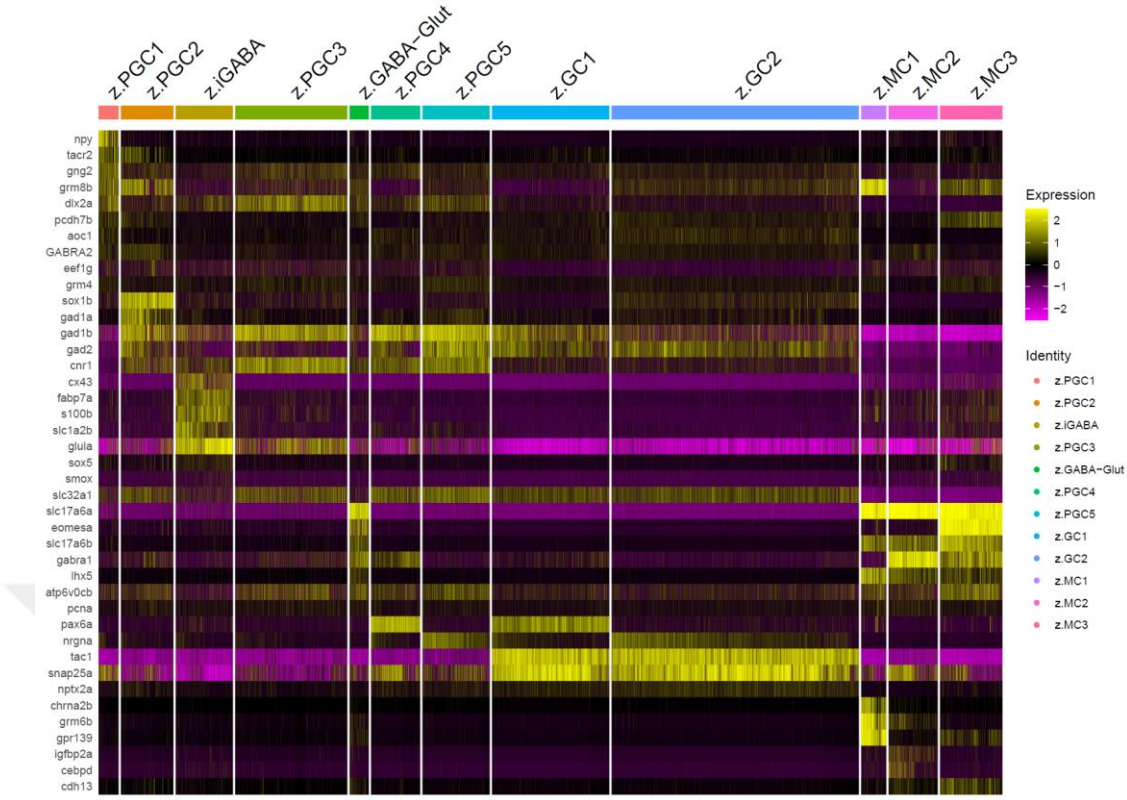

Figure 18. Top 10 most expressed AZTEC genes for each cluster represented on Heatmap.

Eventually, by considering differential expression analysis, marker genes for AZTEC clusters were selected. While no marker genes detected for each GABAergic cluster individually, *tac1* was detected as a marker gene for granule cells (Figure 19). The presence of other GABAergic clusters was combinatorial in terms of gene expression. Clusters z.MC1, z.MC3 are characterized by distinct marker genes, with z.MC1 expressing genes *gpr139*, *grm6b*, z.MC3 expressing gene *eomesa* (Figure 20,21). No marker gene is detected for z.MC2 cluster.

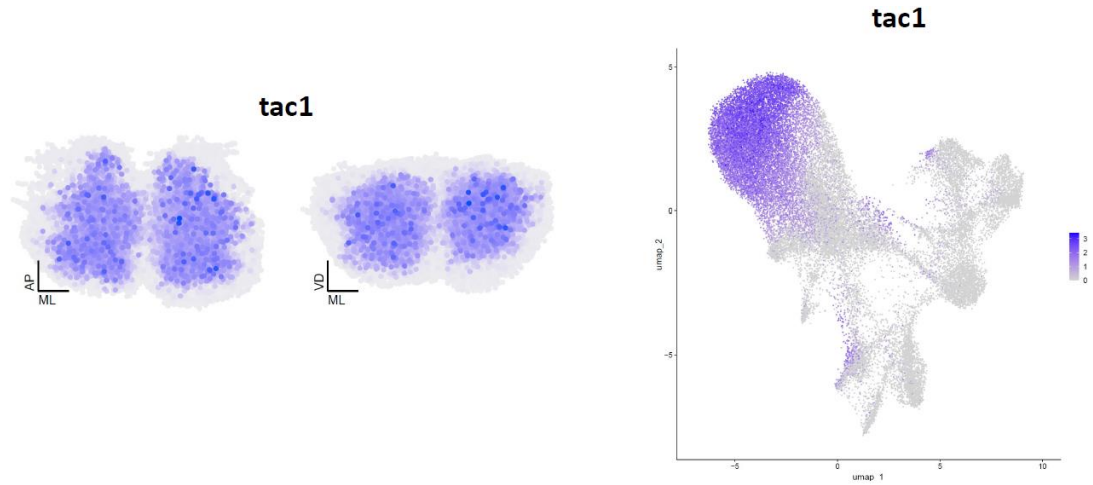

Figure 19. Zebrafish Granule marker gene expression: *tac1* in spatial 2D plot, and in UMAP. AP: Anterior Posterior, ML: Medial Lateral, VD: Ventral Dorsal.

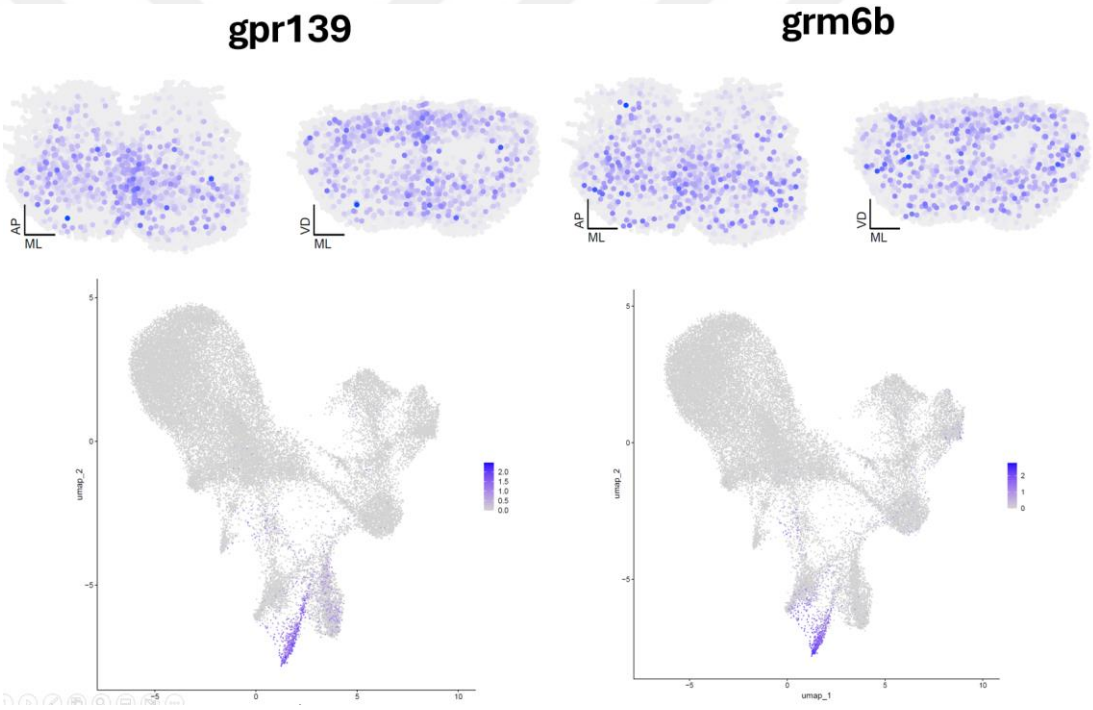

Figure 20. Cluster z.MC1 marker genes expression: *gpr139, and grm6b* in spatial 2D plot, and in UMAP. AP: Anterior Posterior, ML: Medial Lateral, VD: Ventral Dorsal.

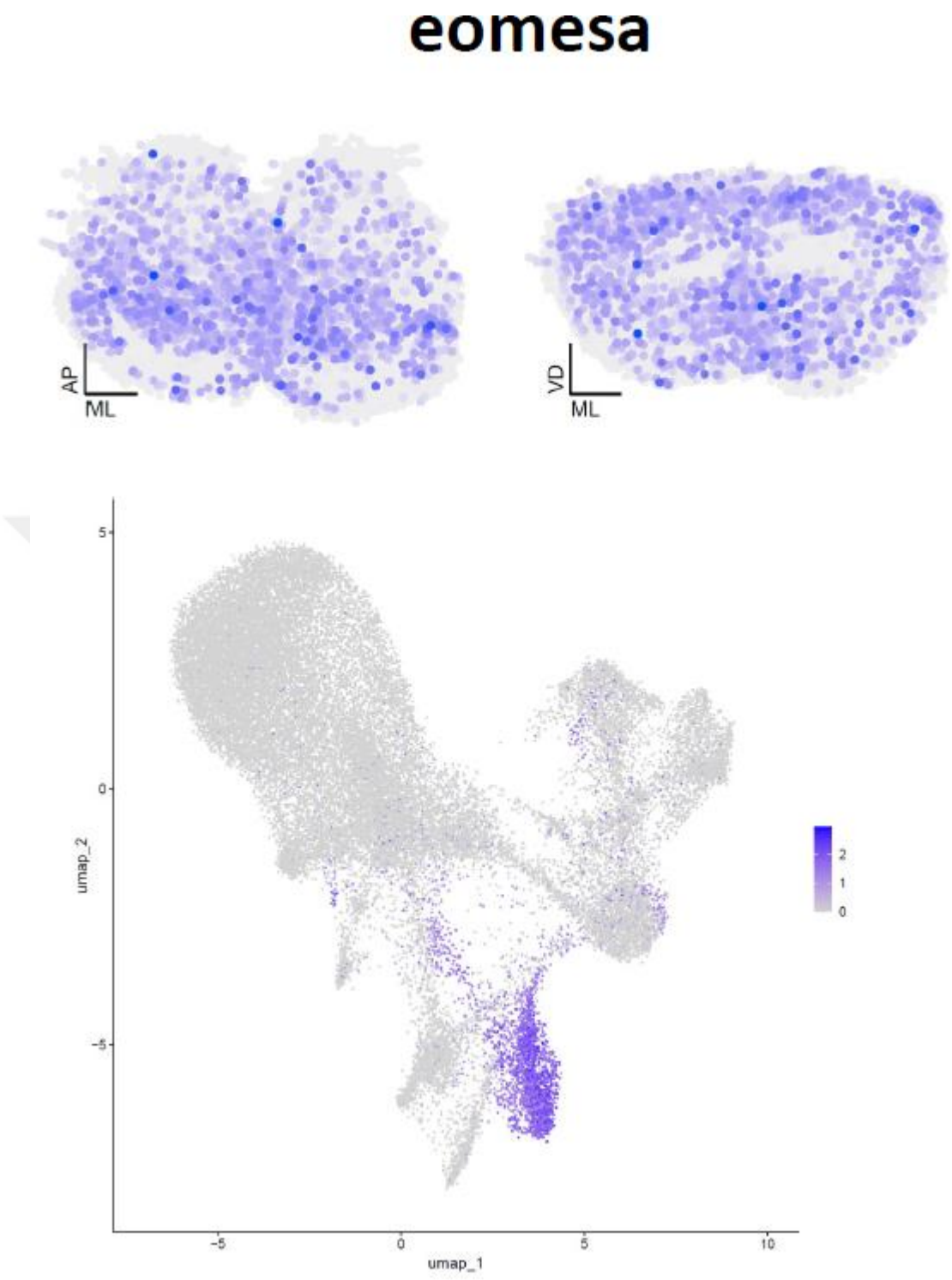

Figure 21. Cluster z.MC3 marker genes expression: *eomesa* in spatial 2D plot, and in UMAP. AP: Anterior Posterior, ML: Medial Lateral, VD: Ventral Dorsal.

4.1.1.3 Verification of AZTEC clusters with zebrafish scRNA-seq clusters

In order to validate our AZTEC clusters, AZTEC dataset, and zebrafish scRNA-seq dataset were integrated Canonical Correlation Analysis (CCA) (Figure 22). After integration, KMR matrix, and alluvial plot (along with a dendrogram attached) were created (Figure 23,24). It was observed that KMR matrix (created from harmonic means of KMR values.) follows a diagonal correlation which means AZTEC clusters are able to find their corresponding clusters in zebrafish scRNA-seq data.

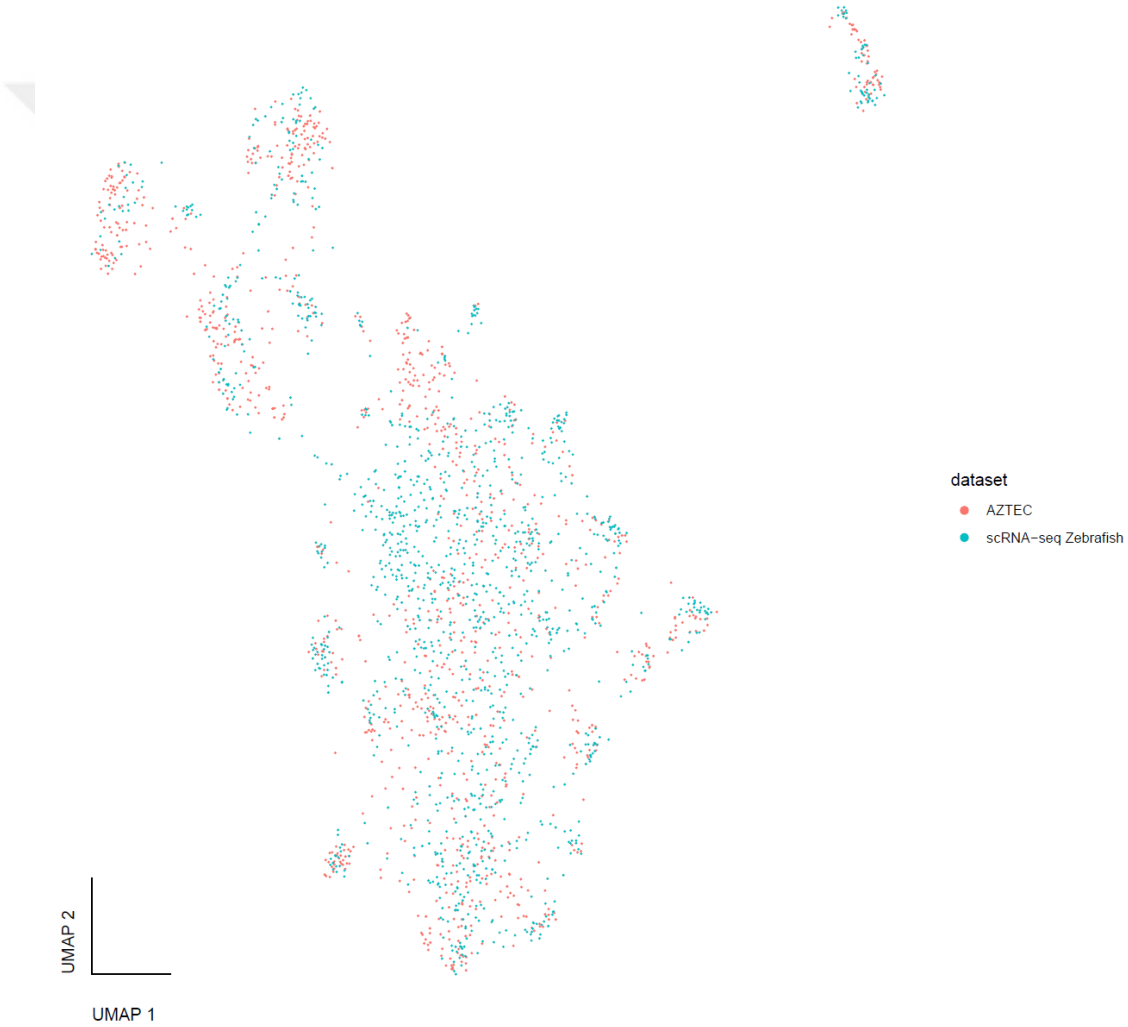

Figure 22. Integration of AZTEC, and zebrafish scRNA-seq datasets in UMAP.

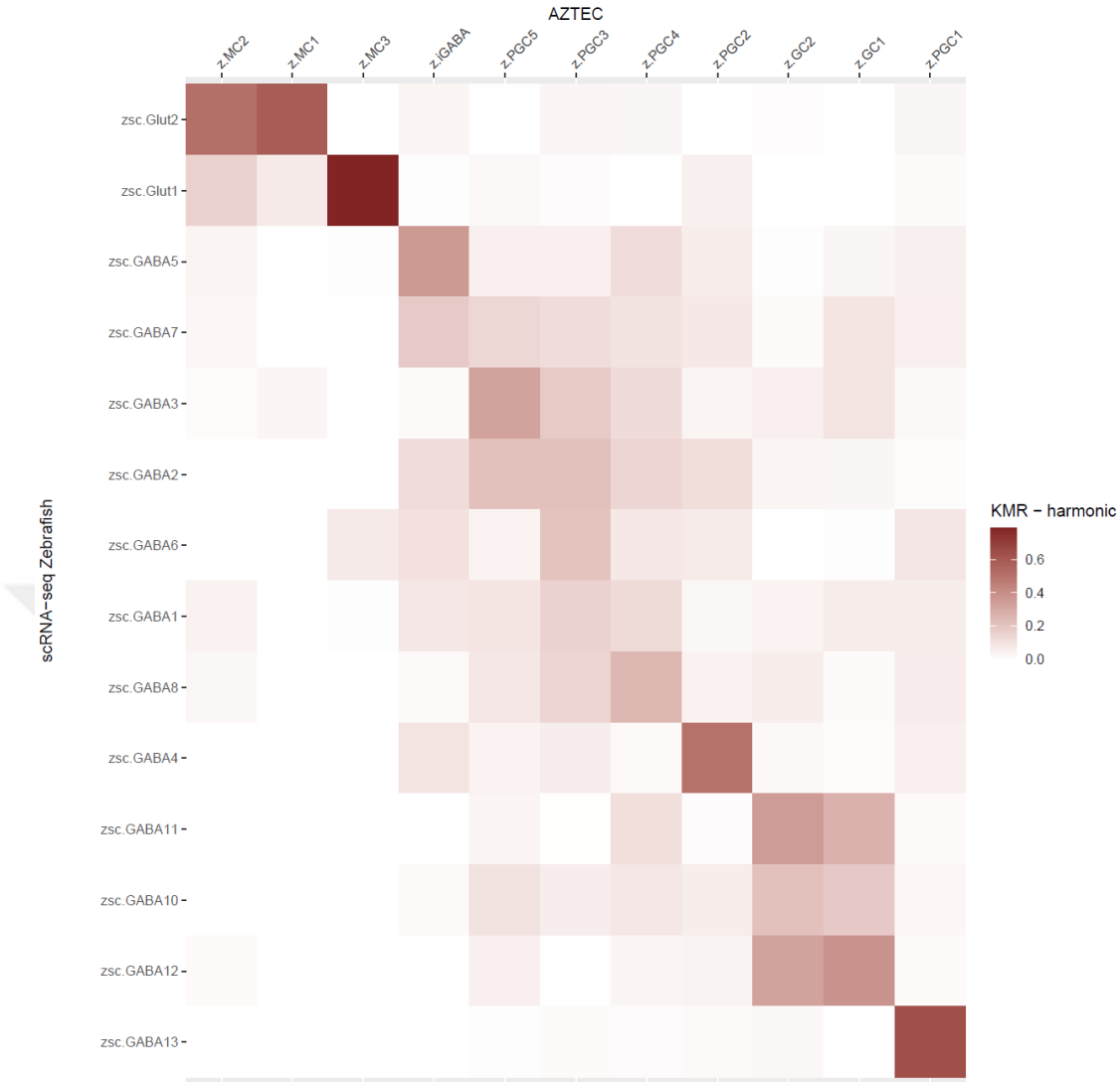

Figure 23. KMR matrix of AZTEC, and zebrafish scRNA-seq datasets clusters.

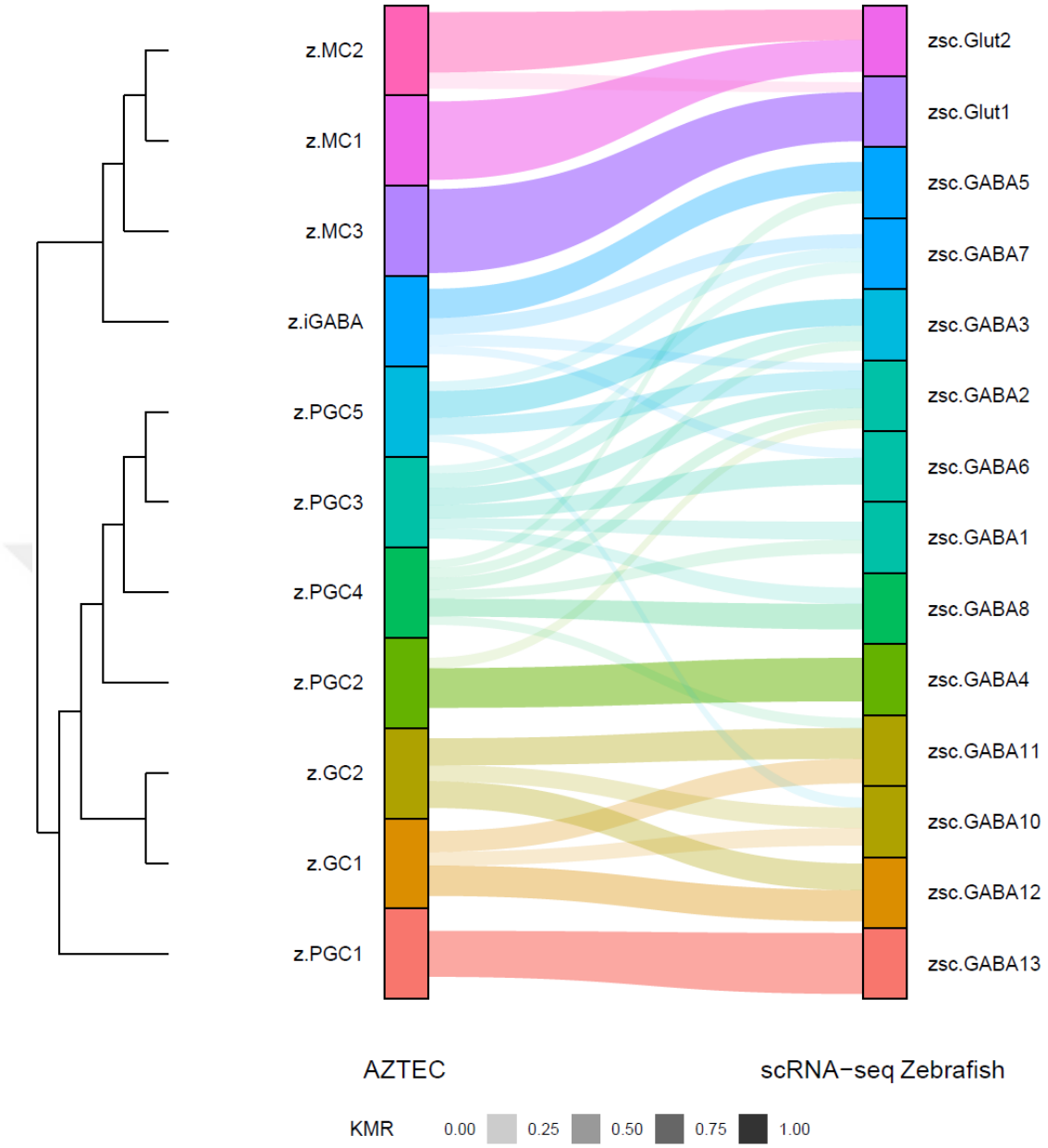

Figure 24. Alluvial plot of AZTEC, and zebrafish scRNA-seq datasets. Dendrogram was created based on genetic similarity of AZTEC clusters.

4.1.2 Analysis of mouse datasets

4.1.2.1 Analysis of mouse single-cell RNA sequencing dataset

Firstly, SCTranform normalization was applied, and variable genes across the dataset were detected, and first 3000 genes were selected as variable genes. Top 10 most variable genes were highlighted which are *Plp1*, *Flt1*, *Ptgds*, *Gpc5*, *Tmem132d*, *Mal*, *Slc1a2*, *Cx3cr1*, *Kcnip4*, *Rgs5* (Figure 25).

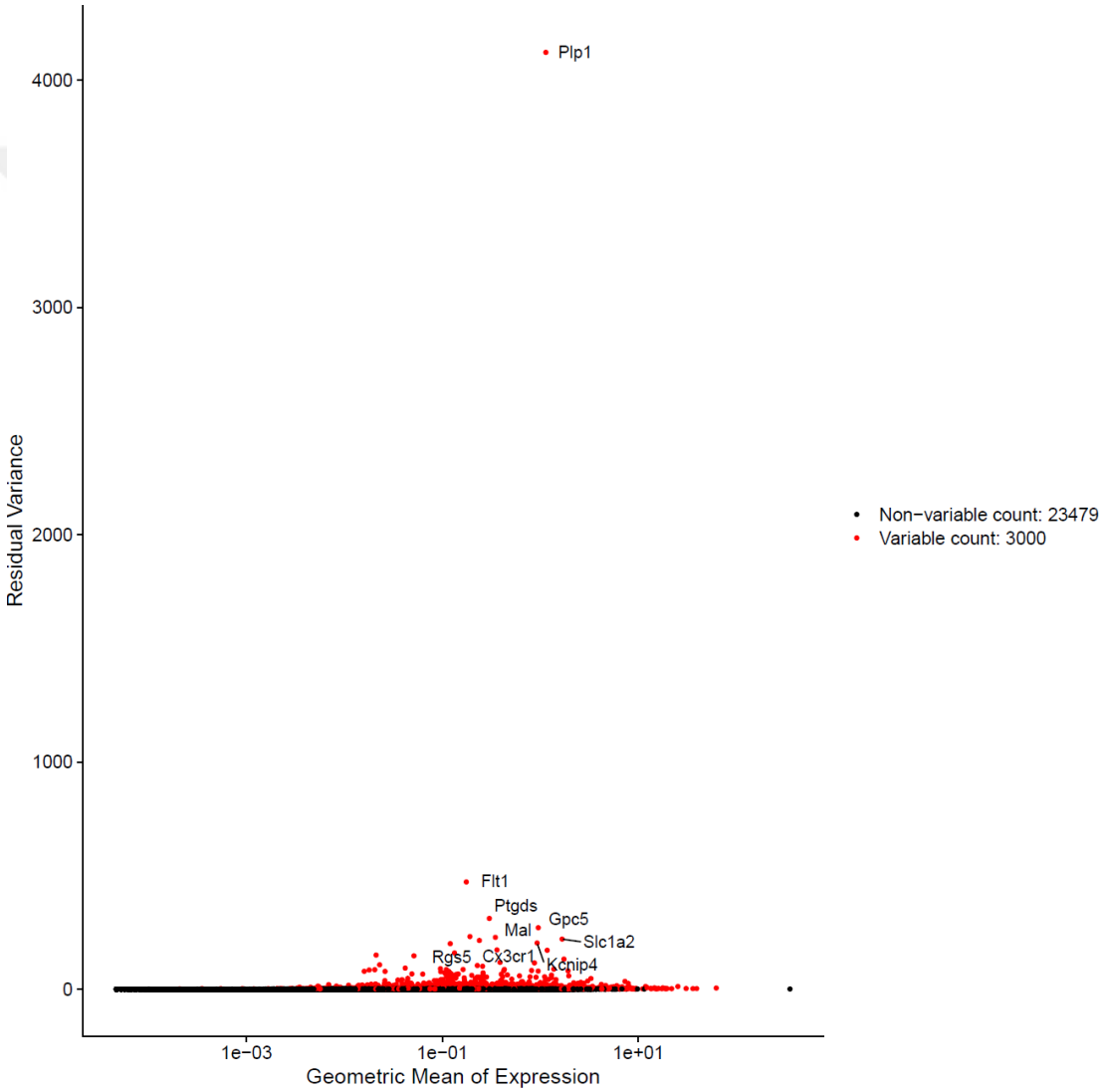

Figure 25. Analysis of the mouse scRNA-seq variable genes. Red dots representing variable features, and black dots representing non-variable genes.

After normalization step, Principal Component Analysis (PCA) was applied to the dataset. The number of Principle Components (PCs) used was determined considering the elbow plot, and first 30 PCs were selected (Figure 26).

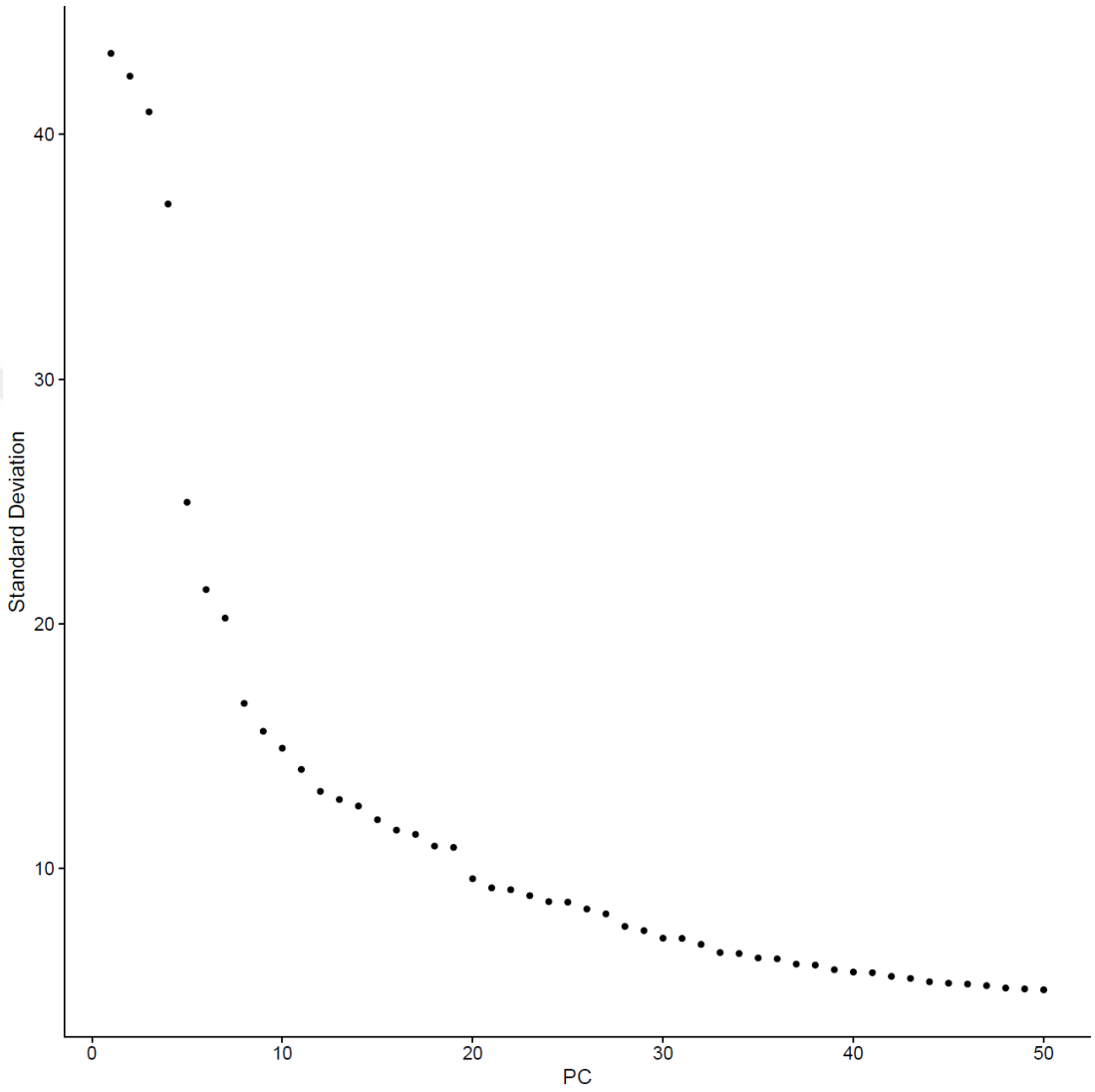

Figure 26. Mouse scRNA-seq PCs vs standard deviation elbow plot

After constructing shared nearest neighbor (SNN) graph, in order to identify clusters of similar cells, Louvain community detection method was applied, and resolution parameter was detected as 0.6 using clustree. 31 clusters were identified in total (Figure 27).

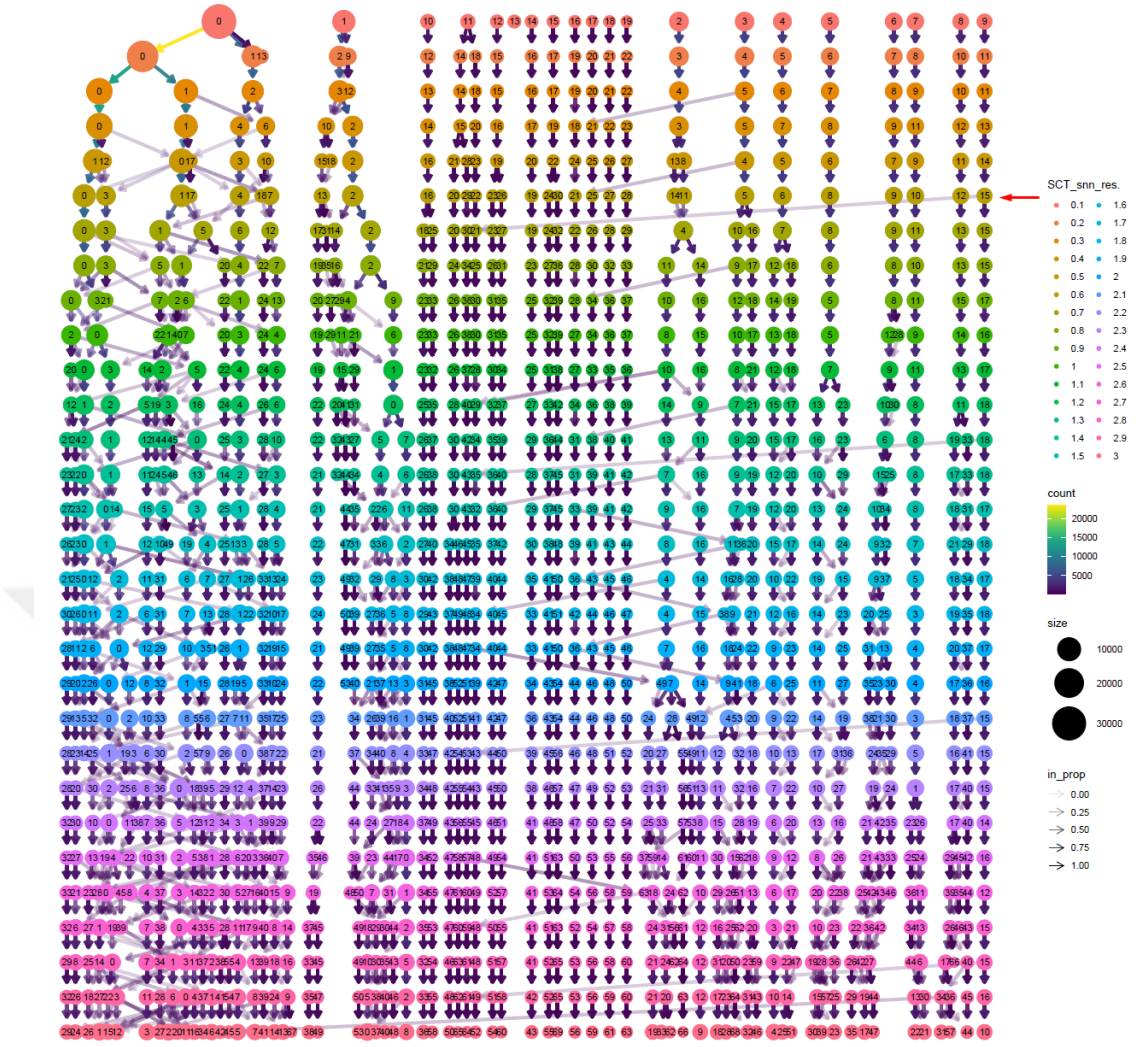

Figure 27. All mouse scRNA-seq clusters ranging from 0.1 until 3.0 resolution. Red arrow represents selected resolution which is 0.6. Circle size represents the largeness of the clusters. Gray arrows represent in what percentage one cluster merge or subdivide to another one in the next resolution.

UMAP was constructed, integration of two animals was observed (Figure 28). Afterwards, the expressions of GABAergic marker gene (*Slc32a1*), and glutamatergic marker genes (*Slc17a7, Slc17a6*) were shown in FeaturePlot (Figure 29, 30).

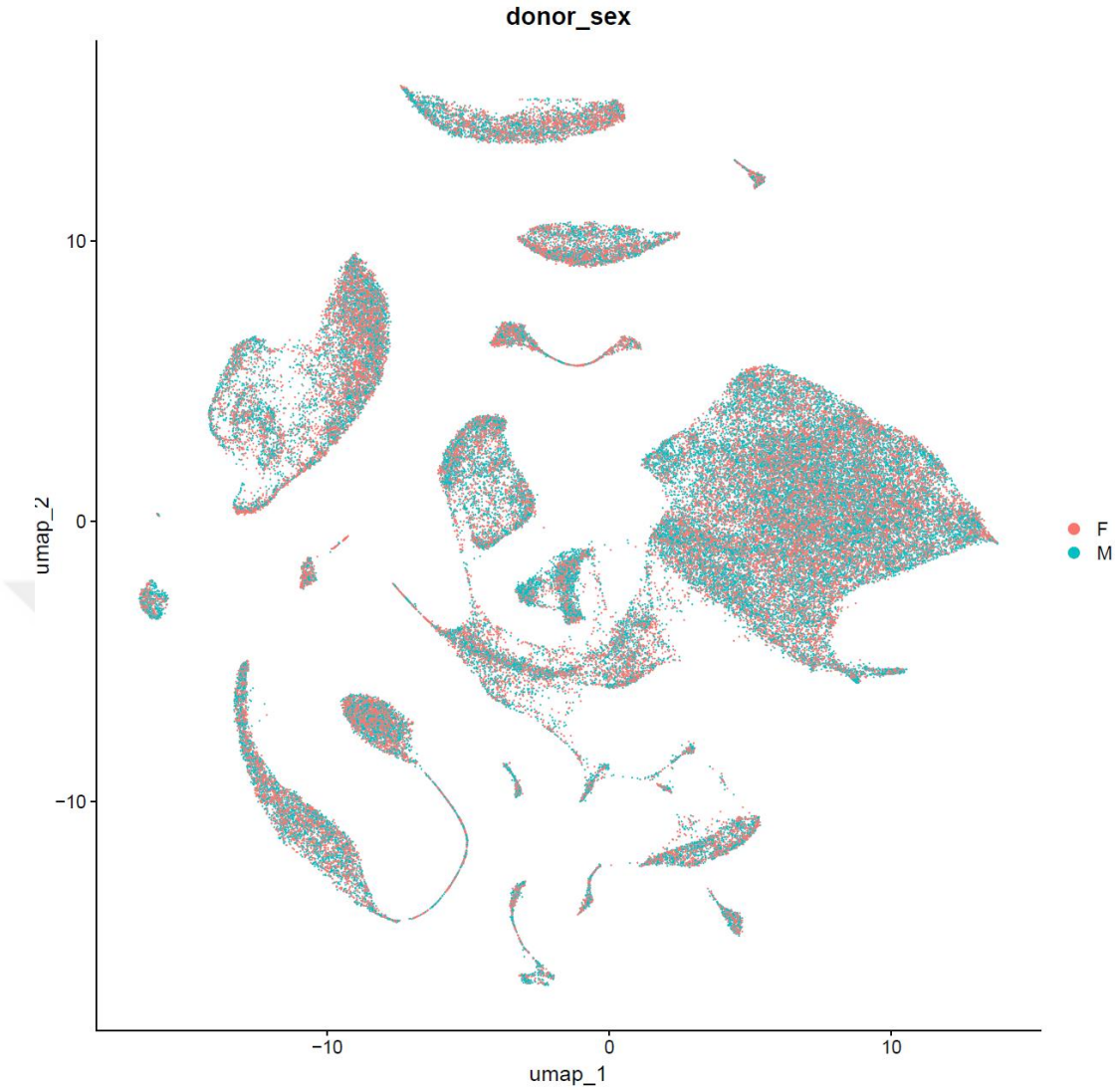

Figure 28. Integration of mouse scRNA-seq animals. 1 female and 1 male, F: Female, M: Male.

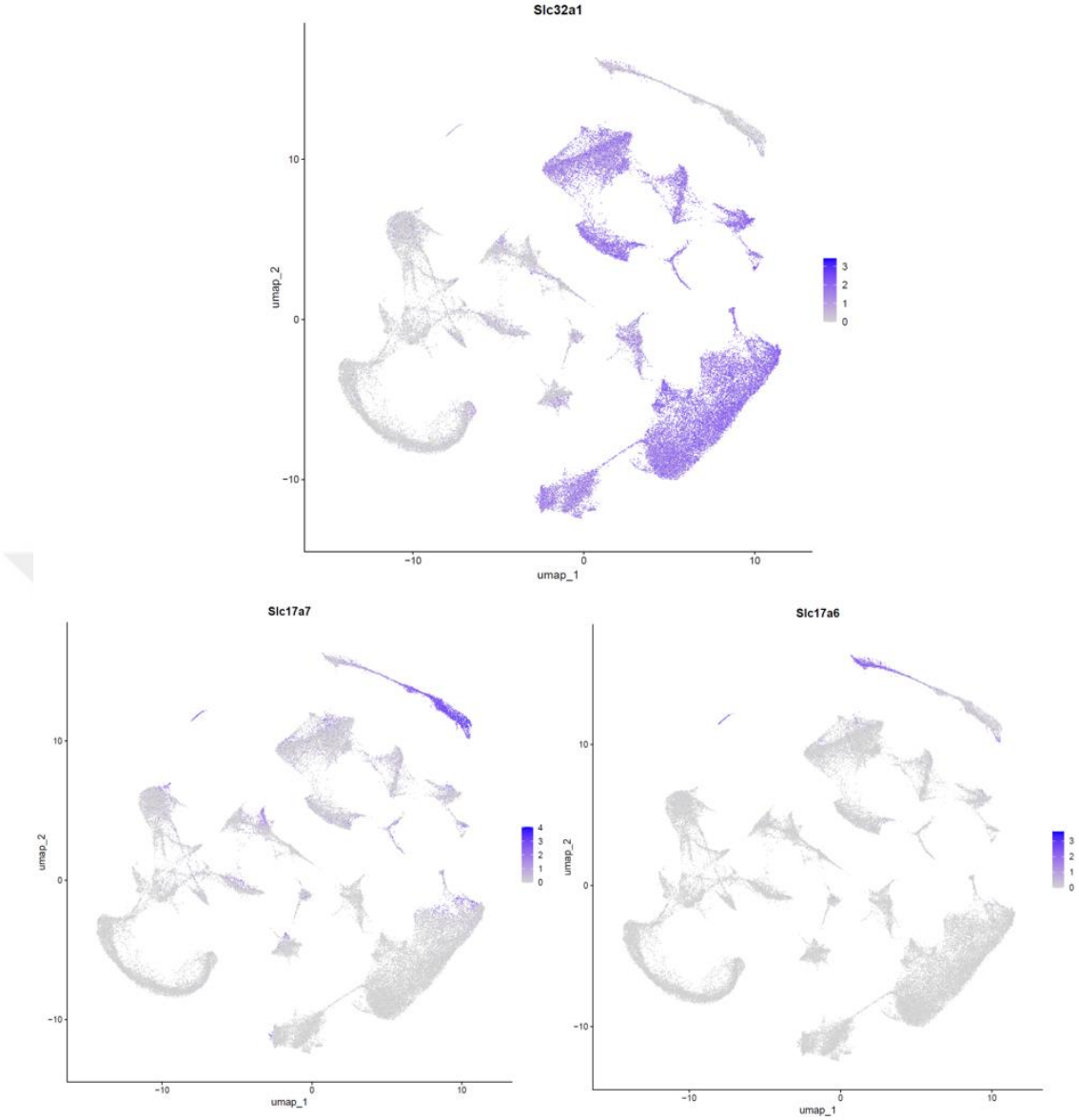

Figure 29. The mouse scRNA-seq gene expression of *Slc32a, Slc17a7, Slc17a6* genes in FeaturePlots.

Clusters were sorted based on their genetic similarity using CoCy R toolkit. Sorted clusters then represented in UMAP. Thus, 12 non-neuronal, 17 GABAergic, and 2 glutamatergic clusters were identified (Figure 30). Since non-neuronal cells are not in the scope of this project, they were excluded from the analysis.

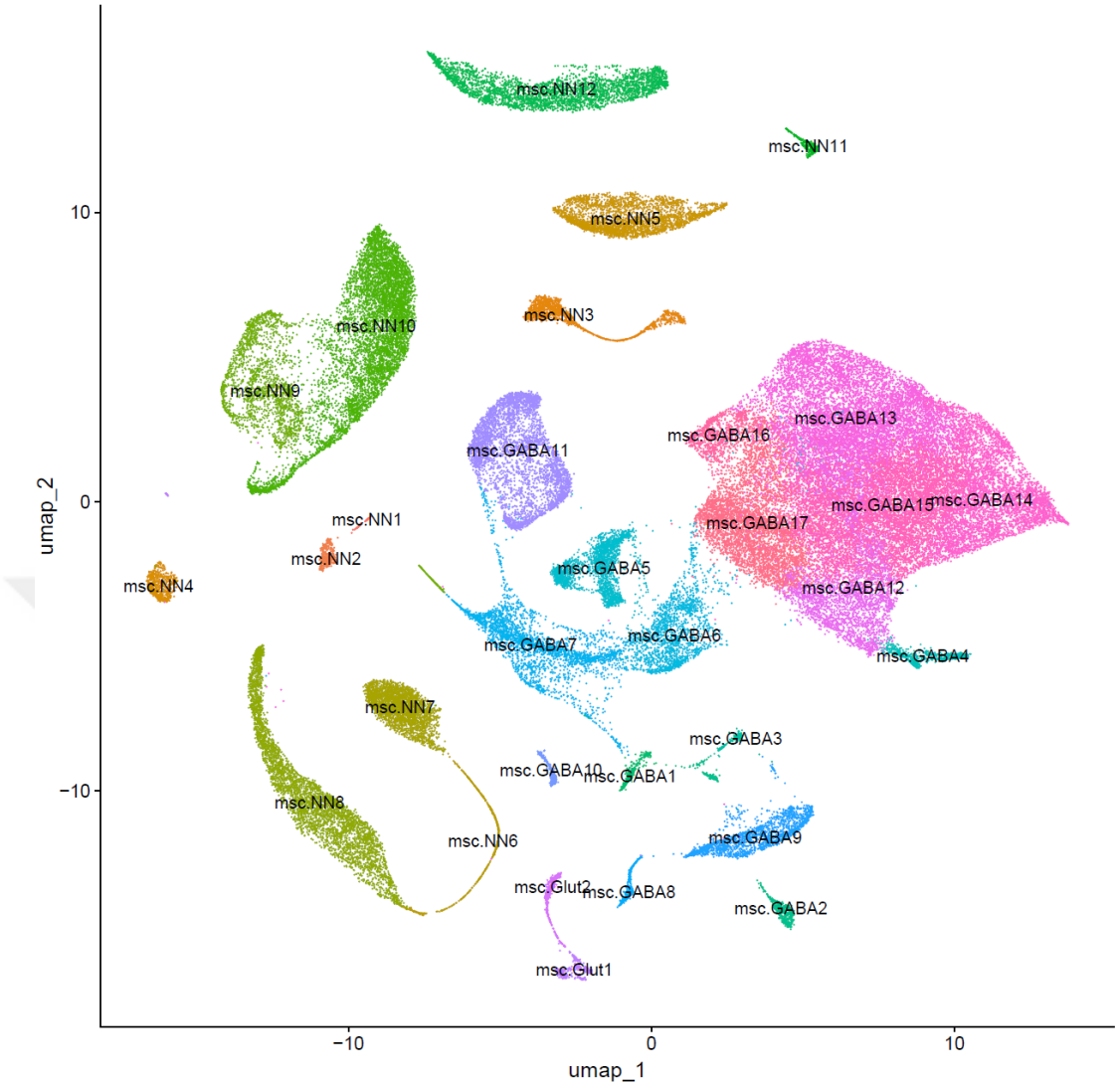

Figure 30. All mouse scRNA-seq clusters represented in UMAP. These clusters belong to mouse scRNA-seq therefore naming prefix is “msc”.

Differential expression analysis was applied on genes using Wilcoxon rank-sum test. Top 10 most expressed genes for each cluster were shown in heatmap (Figure 31).

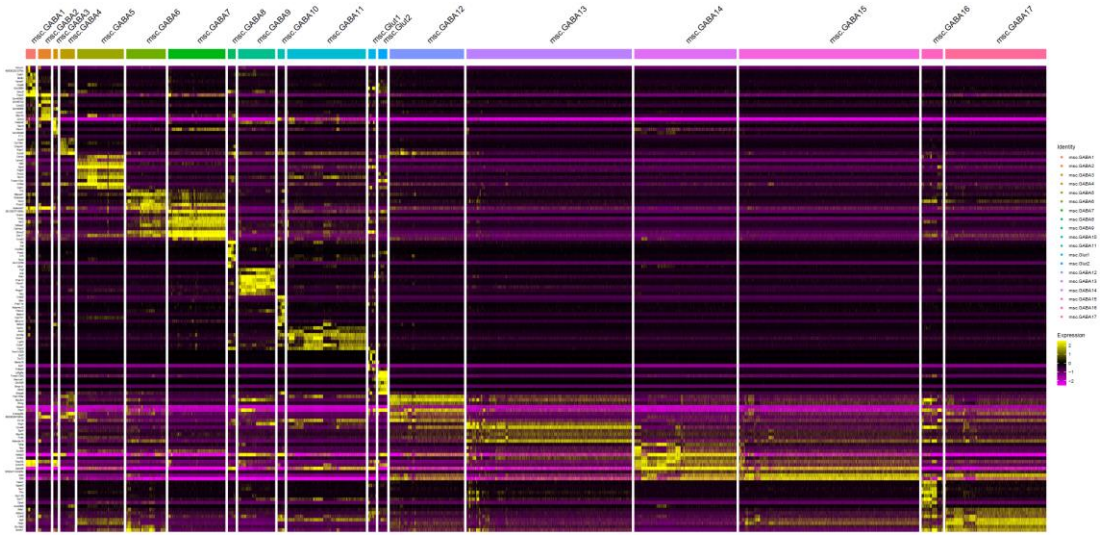

Figure 31. Top 10 most expressed mouse scRNA-seq genes for each cluster represented on Heatmap.

4.1.2.2 Analysis of mouse spatial transcriptomics (MERFISH) dataset

Firstly, SCTranform normalization was applied, and variable genes across the dataset were detected, and all 1122 genes were selected as variable genes. Top 10 most variable genes were highlighted which are *Acta2*, *Cldn11*, *Ctss*, *Mog*, *Slco1c1*, *Abcc9*, *Cdhr1*, *Cldn5*, *Igf2*, *Trh* (Figure 31).

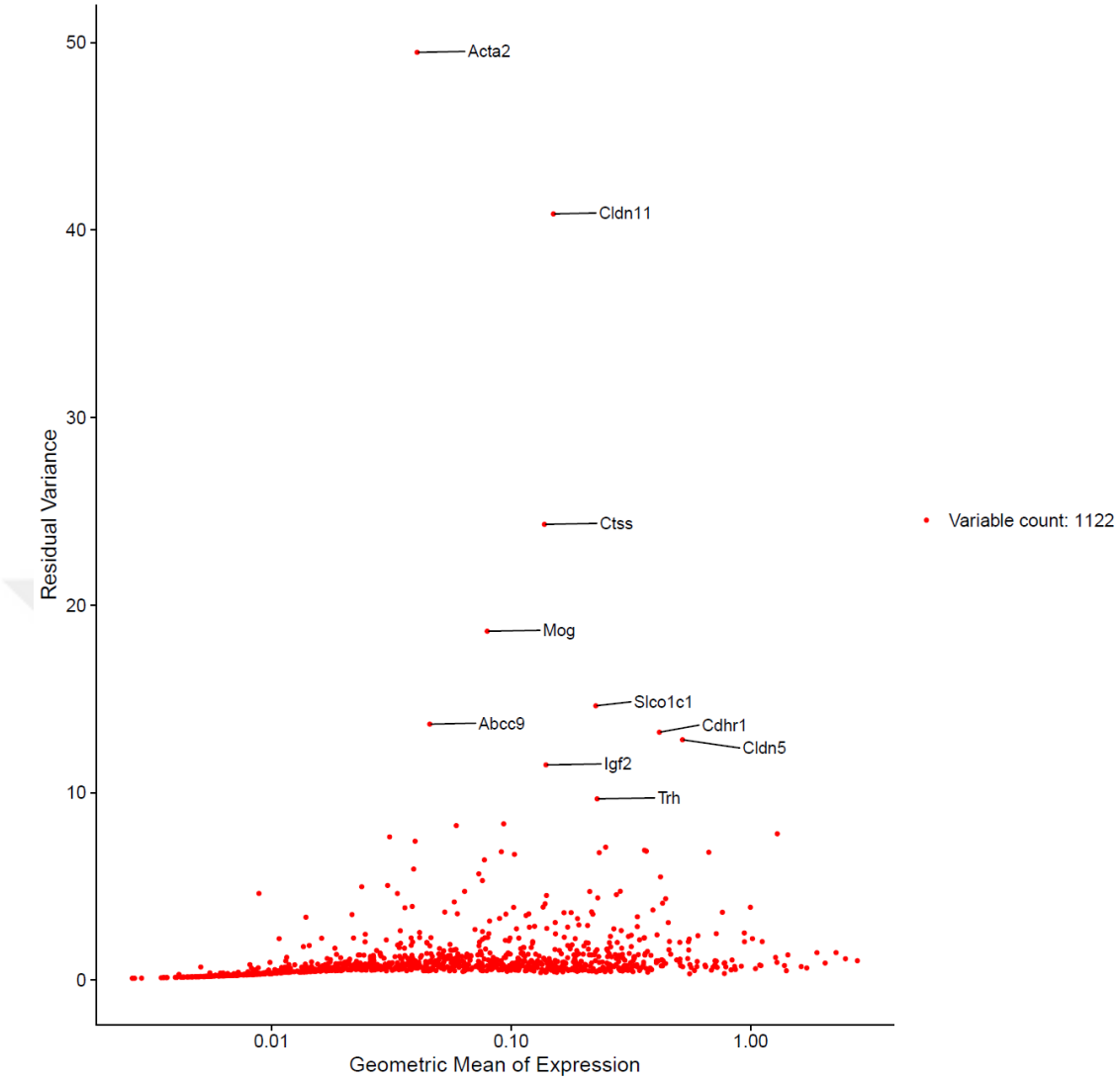

Figure 31. Analysis of the MERFISH variable genes. Red dots representing variable features, and black dots representing non-variable genes.

After normalization step, Principal Component Analysis (PCA) was applied to the dataset. The number of Principle Components (PCs) used was determined considering the elbow plot, and first 30 PCs were selected (Figure 32).

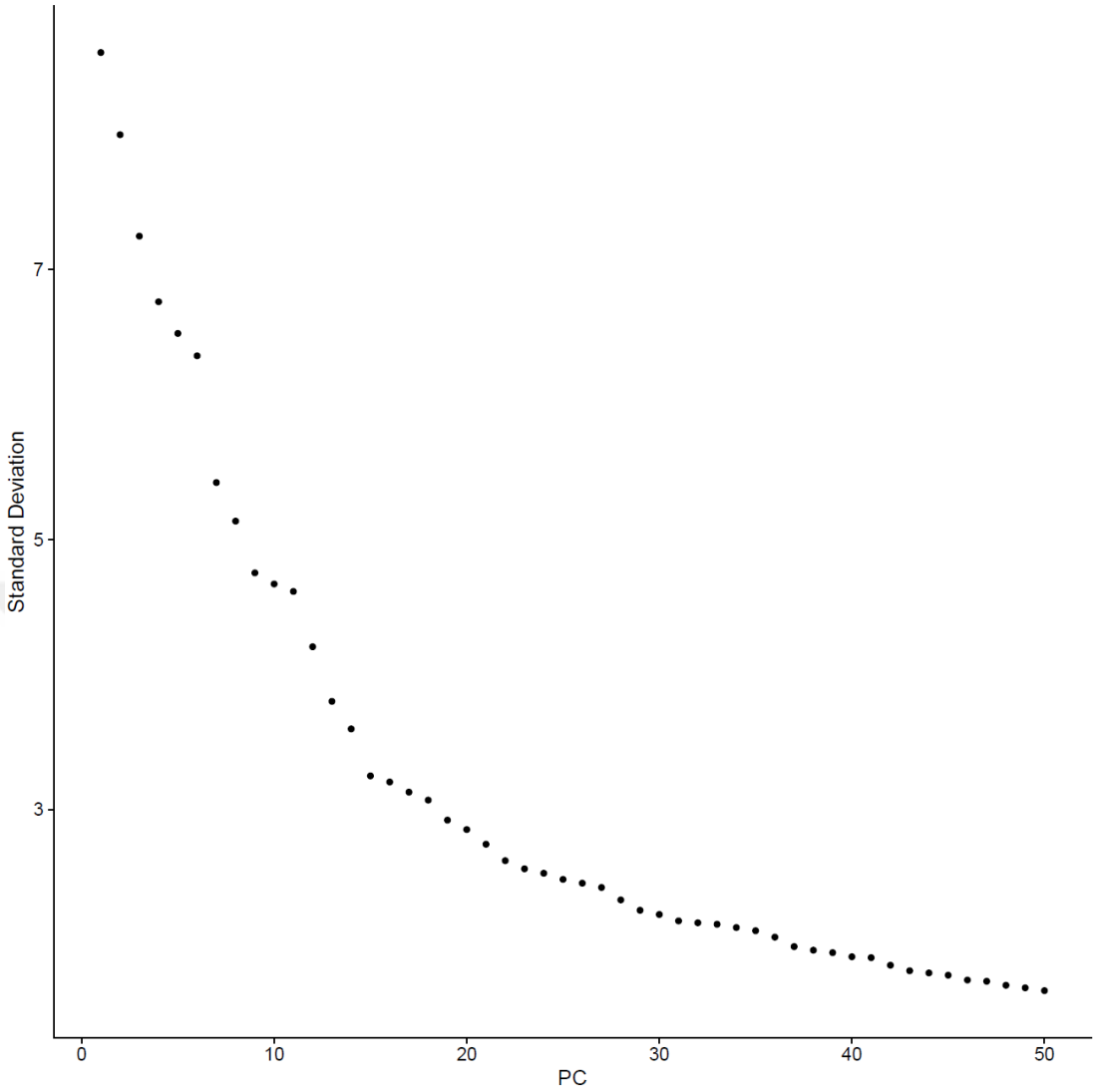

Figure 32. MERFISH PCs vs standard deviation elbow plot

After constructing shared nearest neighbor (SNN) graph, in order to identify clusters of similar cells, Louvain community detection method was applied, and resolution parameter was detected as 0.6 using clustree. 26 clusters were identified in total (Figure 33).

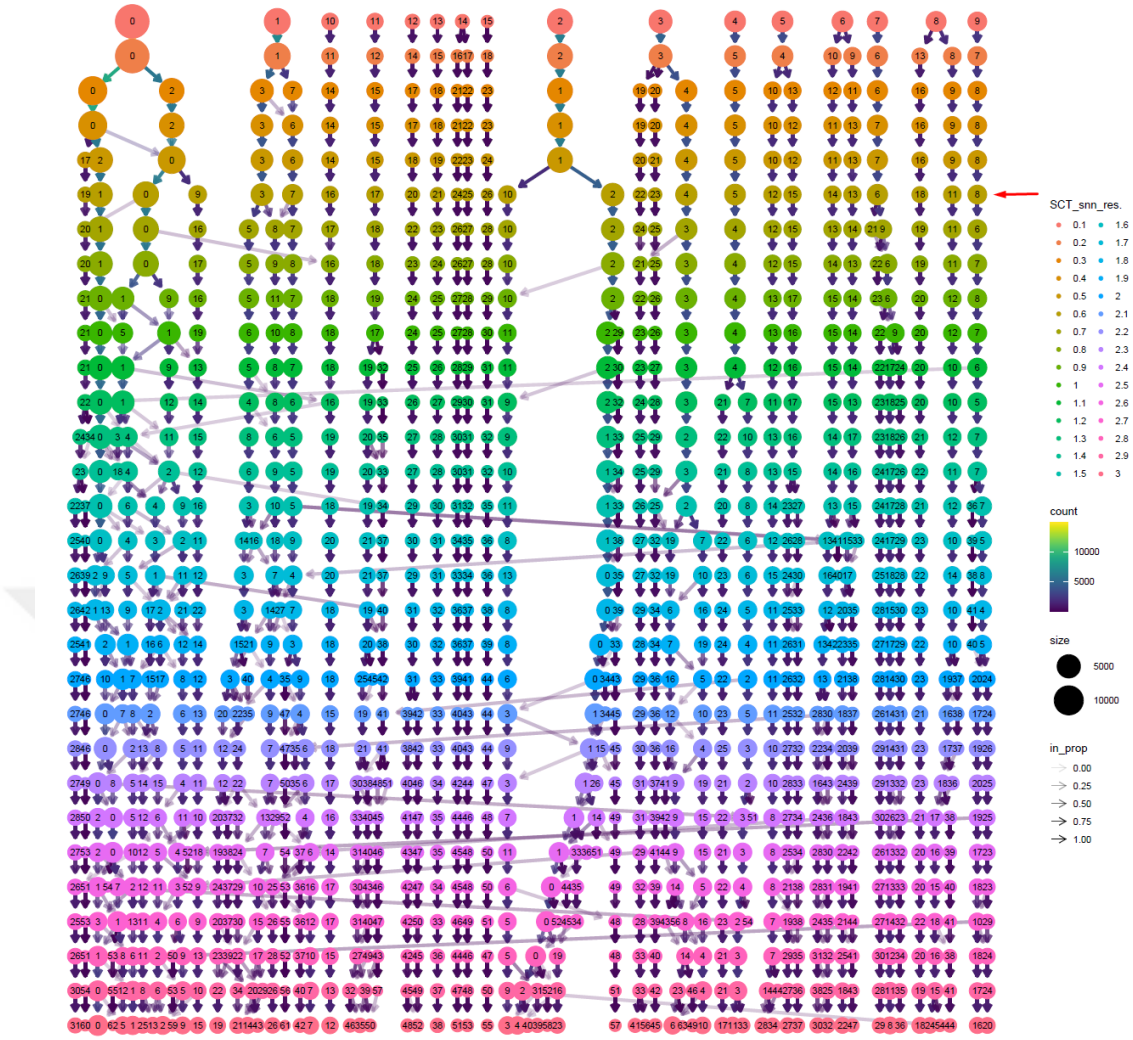

Figure 33. All MERFISH clusters ranging from 0.1 until 3.0 resolution. Red arrow represents selected resolution which is 0.6. Circle size represents the largeness of the clusters. Gray arrows represent in what percentage one cluster merge or subdivide to another one in the next resolution.

UMAP was constructed, and the expressions of GABAergic marker gene (*Slc32a1*), and glutamatergic marker genes (*Slc17a7, Slc17a6*) were shown in FeaturePlot (Figure 34, 35).

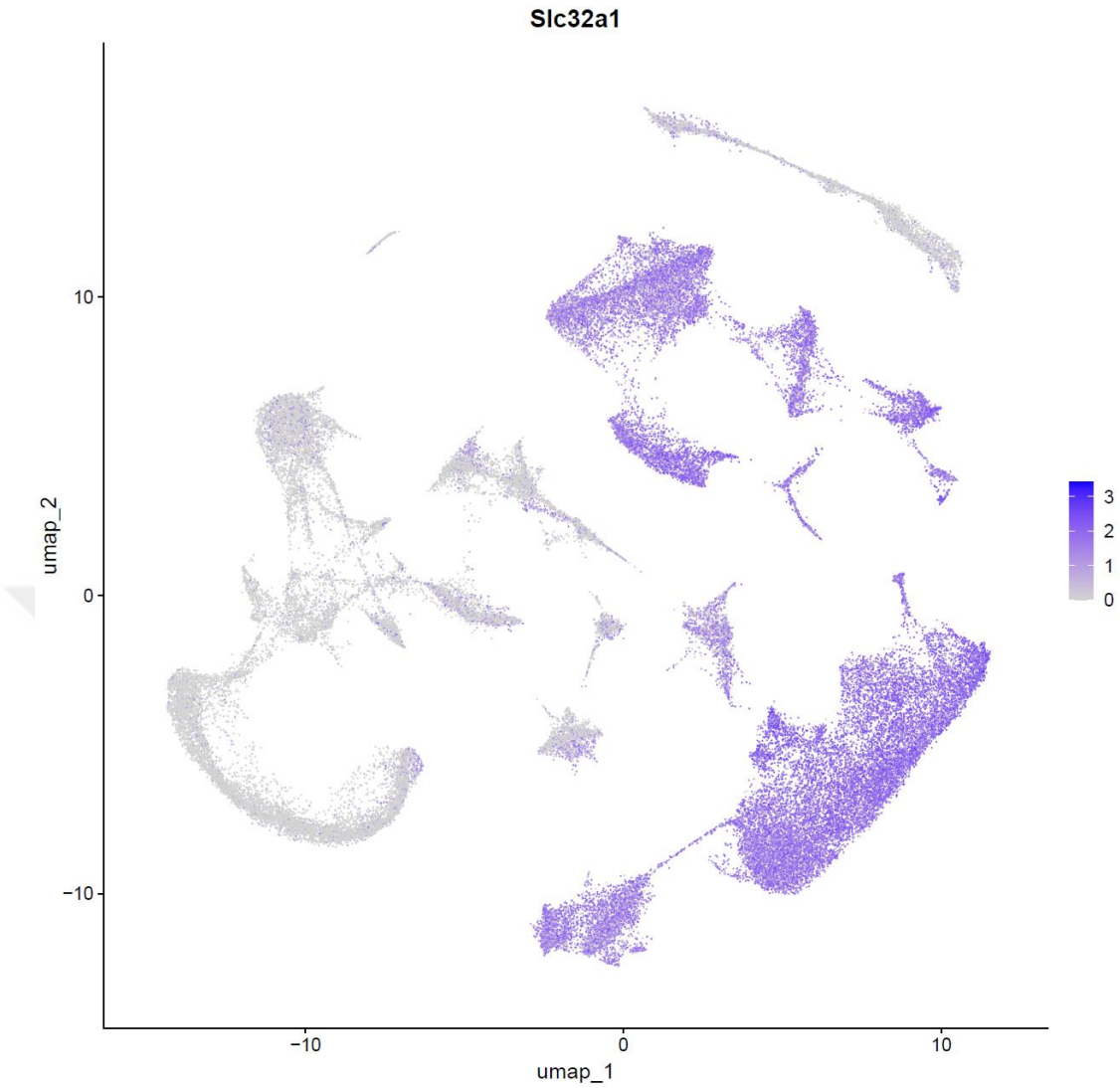

Figure 34. MERFISH gene expression of *Slc32a1* gene in FeaturePlot.

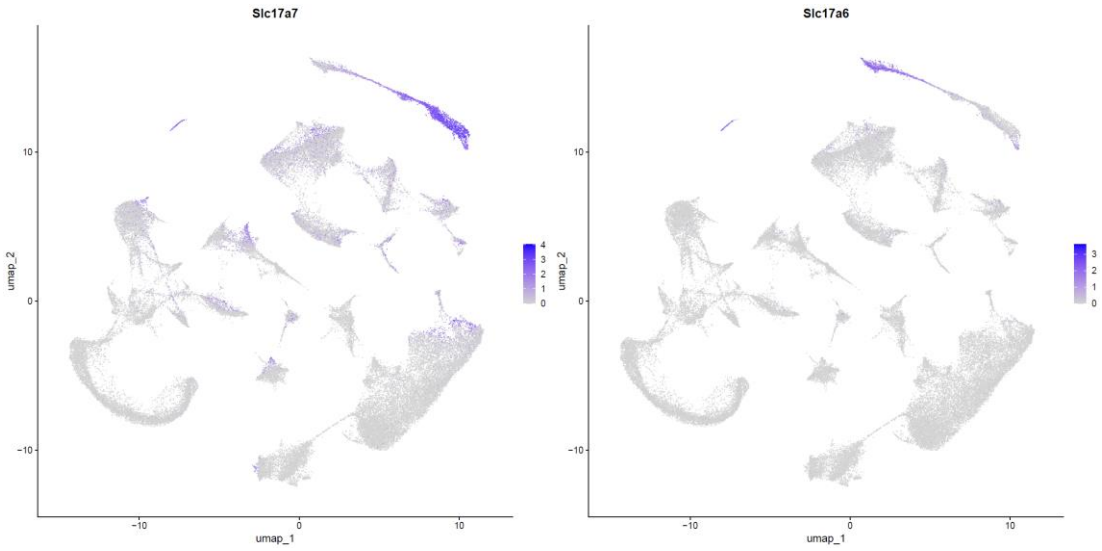

Figure 35. MERFISH gene expression of *Slc17a7, Slc17a6* genes in FeaturePlot.

Clusters were sorted based on their genetic similarity using CoCy R toolkit. Sorted clusters then represented in UMAP. Before annotation of the clusters, their 2D spatial plots were generated in order to classify them according to both spatial and molecular information. According to this spatial information, cluster that abundantly express GABAergic marker genes, and present at outer layer of olfactory bulb were named as periglomerular cells (PGCs) whereas inner cells labeled as granule cells (GCs) (Figure 36, 37. 39).

Moreover, glutamatergic clusters that are present at the outer layer of olfactory bulb were named as external tufted cells (ETCs). Although tufted cells, and mitral cells are different spatially, and genetically, for now they are labeled as tufted cells/ mitral cells (TC/MC) due to resolution limitation. P.S: This cluster will later be subdivided. Thus, 10 non-neuronal, 13 GABAergic (5 granule clusters, 7 periglomerular clusters, and 1 immature GABA cluster), and 3 glutamatergic clusters (2 external tufted clusters, 1 TC/MC cluster) were identified (Figure 38, 39). Since non-neuronal cells are not in the scope of this project, they were excluded from the analysis.

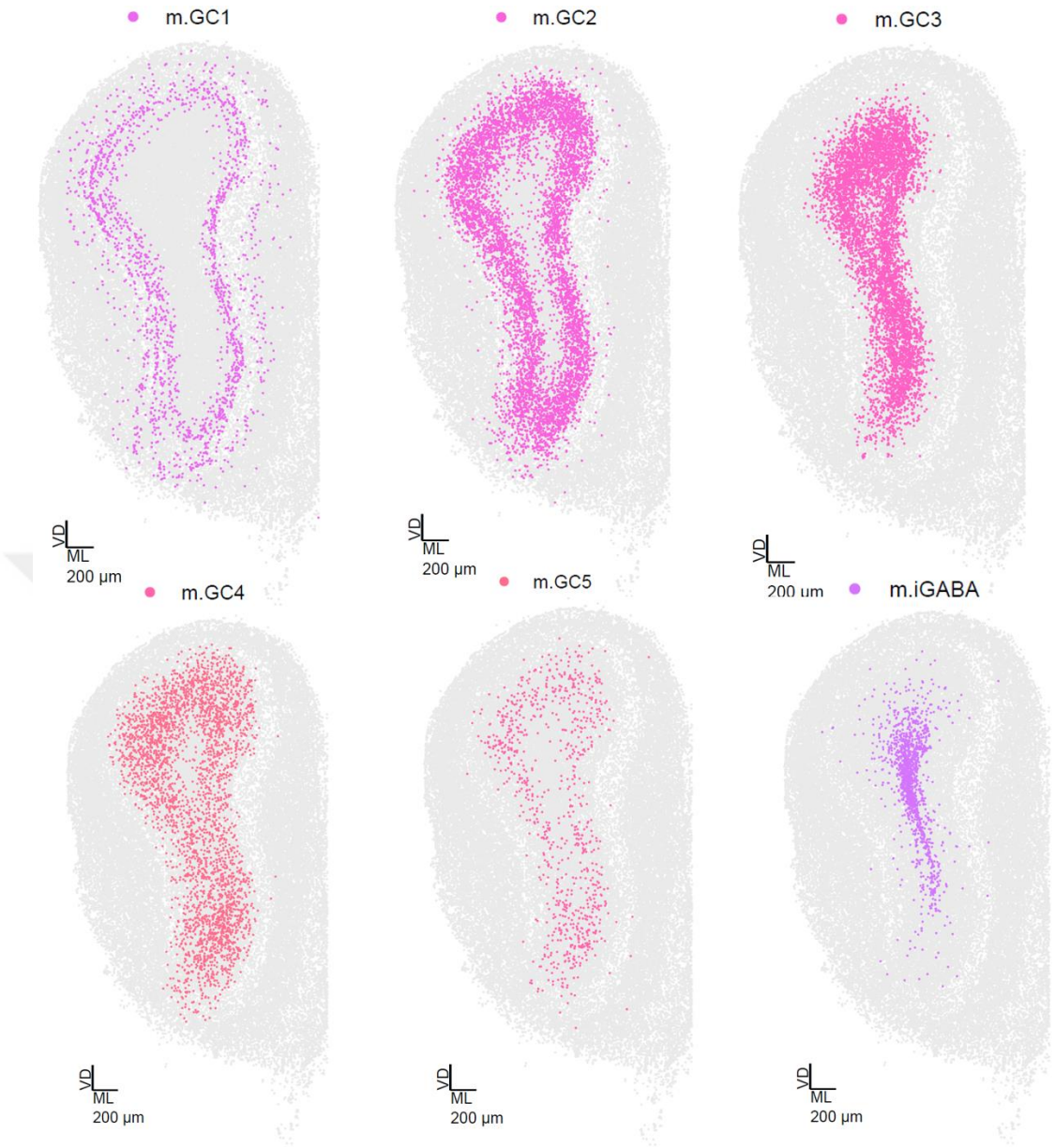

Figure 36. 2D spatial locations of MERFISH GABAergic granule, and immature clusters. GC: Granule Cells, iGABA: Immature GABAergic Cells. Scale is 200-micron length. VD: Ventral Dorsal, ML: Medial Lateral.

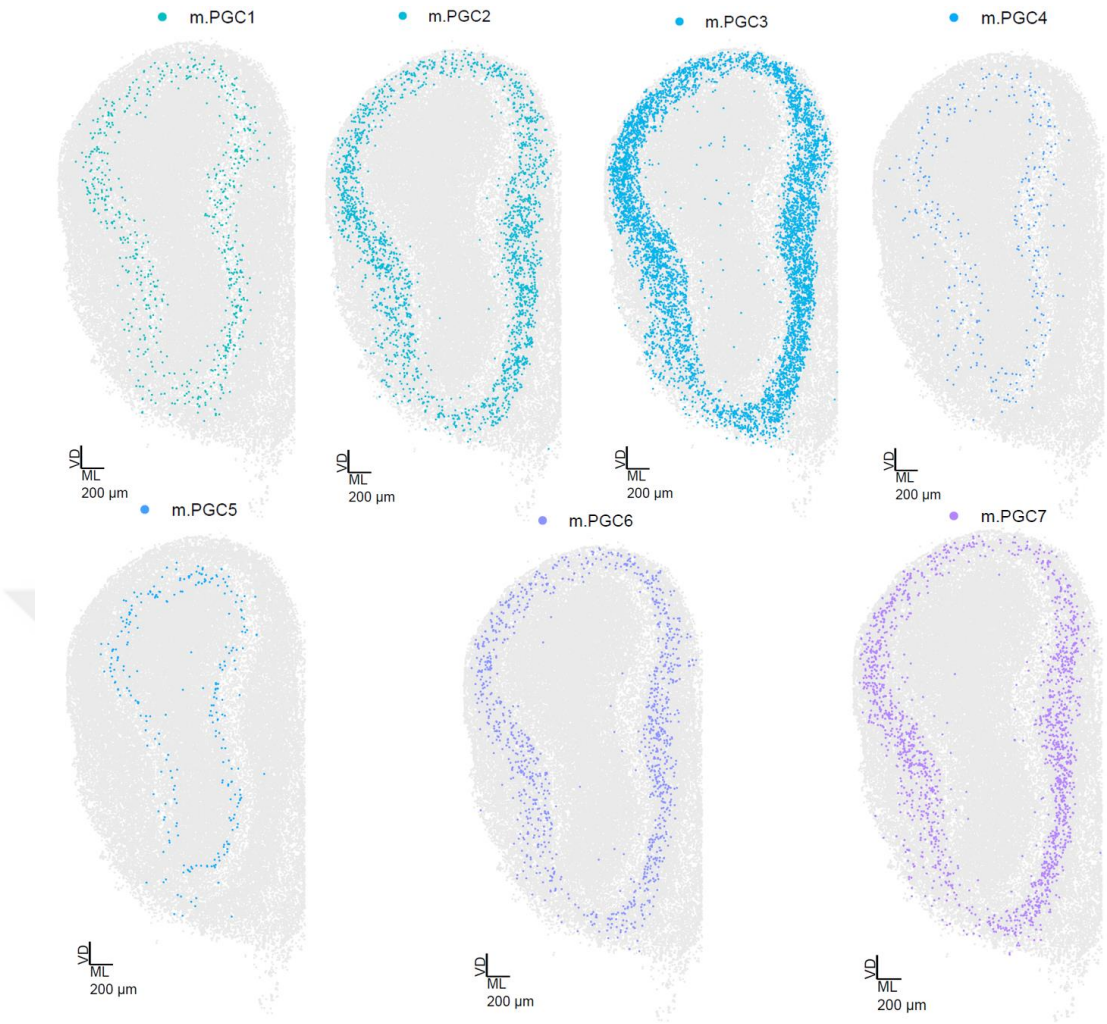

Figure 37. 2D spatial locations of MERFISH GABAergic periglomerular clusters. PGC: Periglomerular Cells. Scale is 200-micron length. VD: Ventral Dorsal, ML: Medial Lateral.

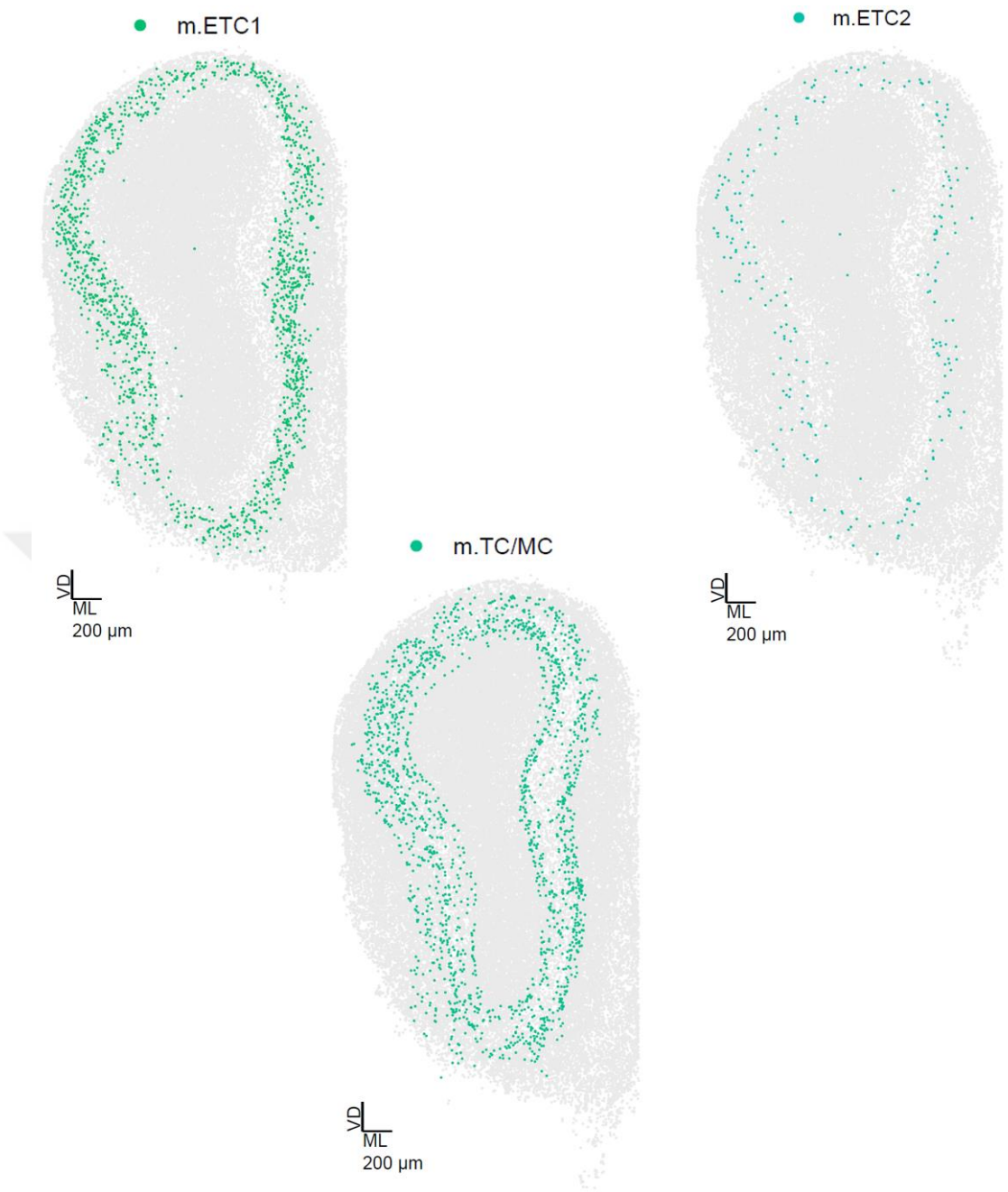

Figure 38. 2D spatial locations of MERFISH glutamatergic external tufted, tufted/mitral clusters. ETC: External Tufted Cells, TC/MC: Tufted Cells/Mitral Cells Scale is 200-micron length. VD: Ventral Dorsal, ML: Medial Lateral.

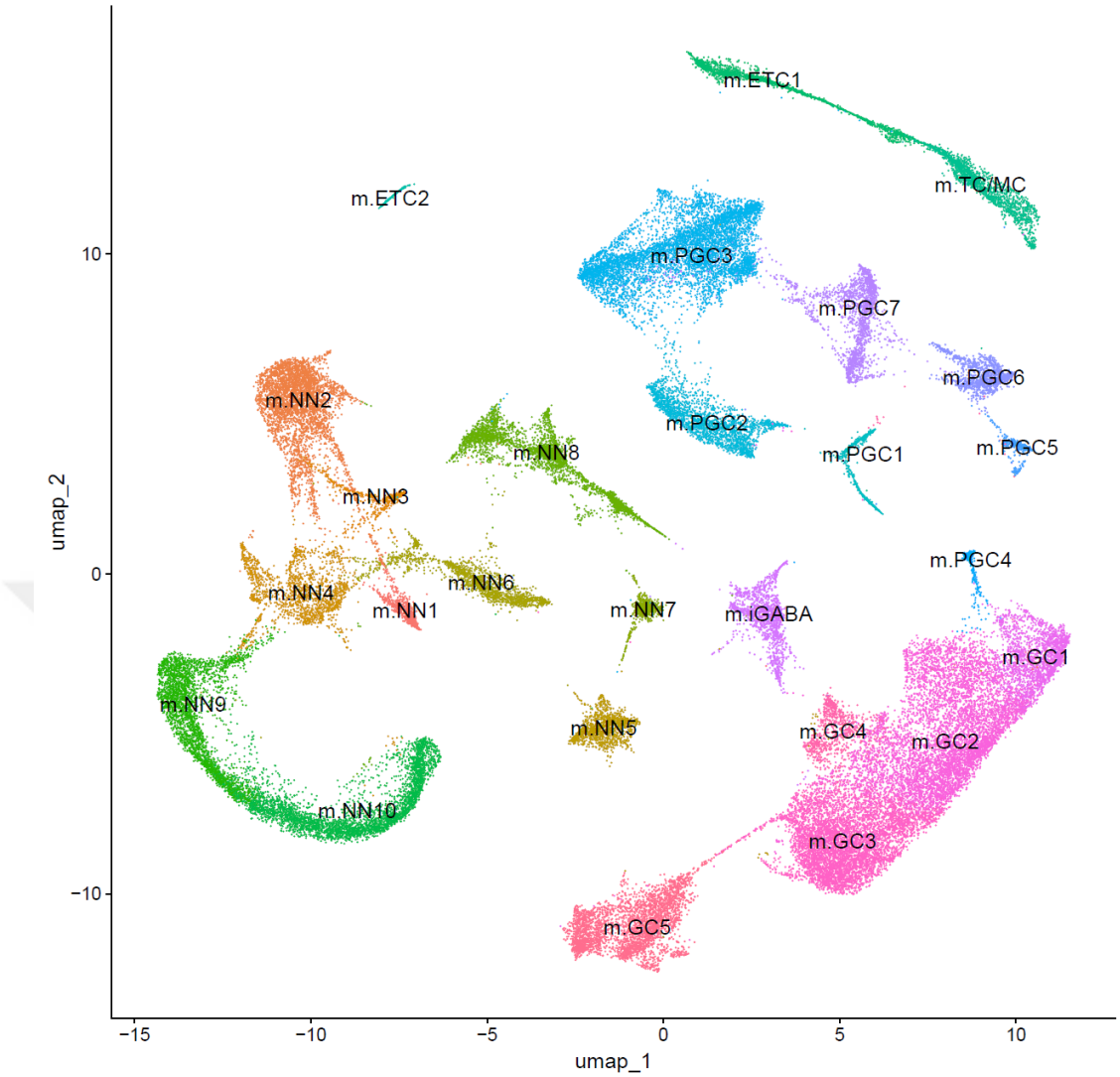

Figure 39. All MERFISH clusters represented in UMAP; these clusters belong to mouse spatial transcriptomics (MERFISH) dataset therefore naming prefix is “m”.

Differential expression analysis was applied on genes using Wilcoxon rank-sum test. Top 10 most expressed genes for each cluster were shown in heatmap (Figure 40).

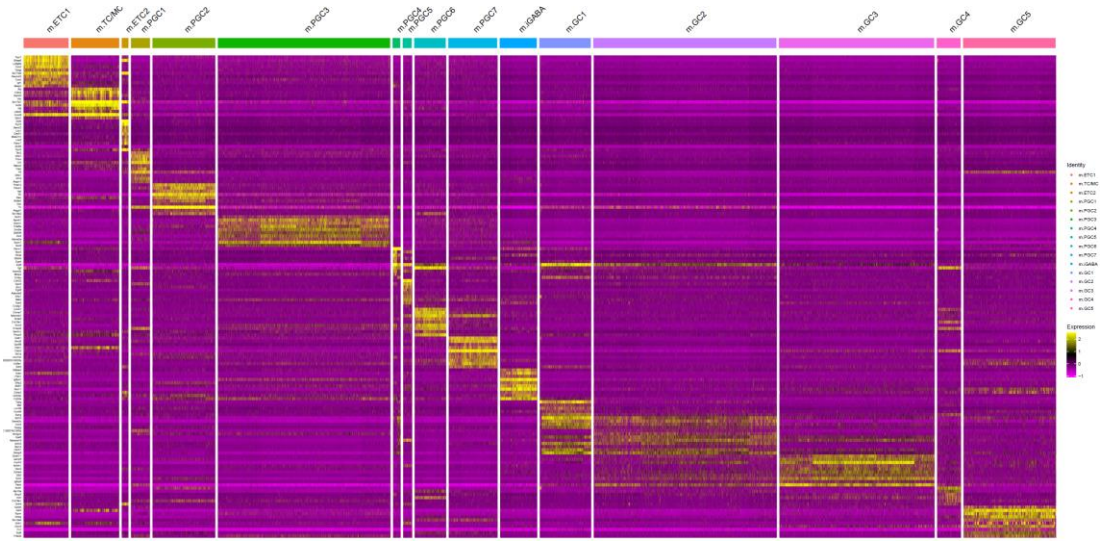

Figure 40. Top 10 most expressed MERFISH genes for each cluster represented on Heatmap.

Eventually, by considering differential expression analysis, marker genes for MERFISH clusters were selected. Periglomerular clusters had the following marker genes: m.PGC1(*Crh, Glra1, Pnoc)* (Figure 41), m.PGC2 (*Prss12, Th, Trh*) (Figure 42), m.PGC3 (*Apold1, Pde5a, Scn5a*) (Figure 43), m.PGC4 (*Parm1*) (Figure 44), m.PGC5 (*Cdh23*) (Figure 44), m.PGC6 (*Chrna7, Plcxd3*) (Figure 45), m.PGC7 (*Calb1, Vwc2l*) (Figure 46). Granule clusters had the following marker genes: m.GC1 (*Tpbg*) (Figure 47), m.GC3 (*Kcnh3*) (Figure 47), m.GC5 (*Rprm*) (Figure 48). While periglomerular clusters show distinct clusters, granule clusters demonstrate a continuum of several clusters that are very similar to each other in terms of gene expression. For example, *Ptk2b* gene is abundantly expressed in all granule clusters, and represent a common marker gene for granule cell types (Figure 49). Lastly for GABAergic clusters, immature GABAergic cluster (iGABA) has two potential marker genes which are *Igfbpl1, Sox11* (Figure 48).

For glutamatergic clusters, external tufted cell clusters had the following marker genes: m.ETC1 (*Ccn2, Ly6g6e, Trpc7*) (Figure 50) , m.ETC2 (*Ebf3, Trp73*) (Figure 51). Tufted and mitral cell combined cluster had 3 potential marker genes which are *Alk, Cbln4, Sv2b* (Figure 52).

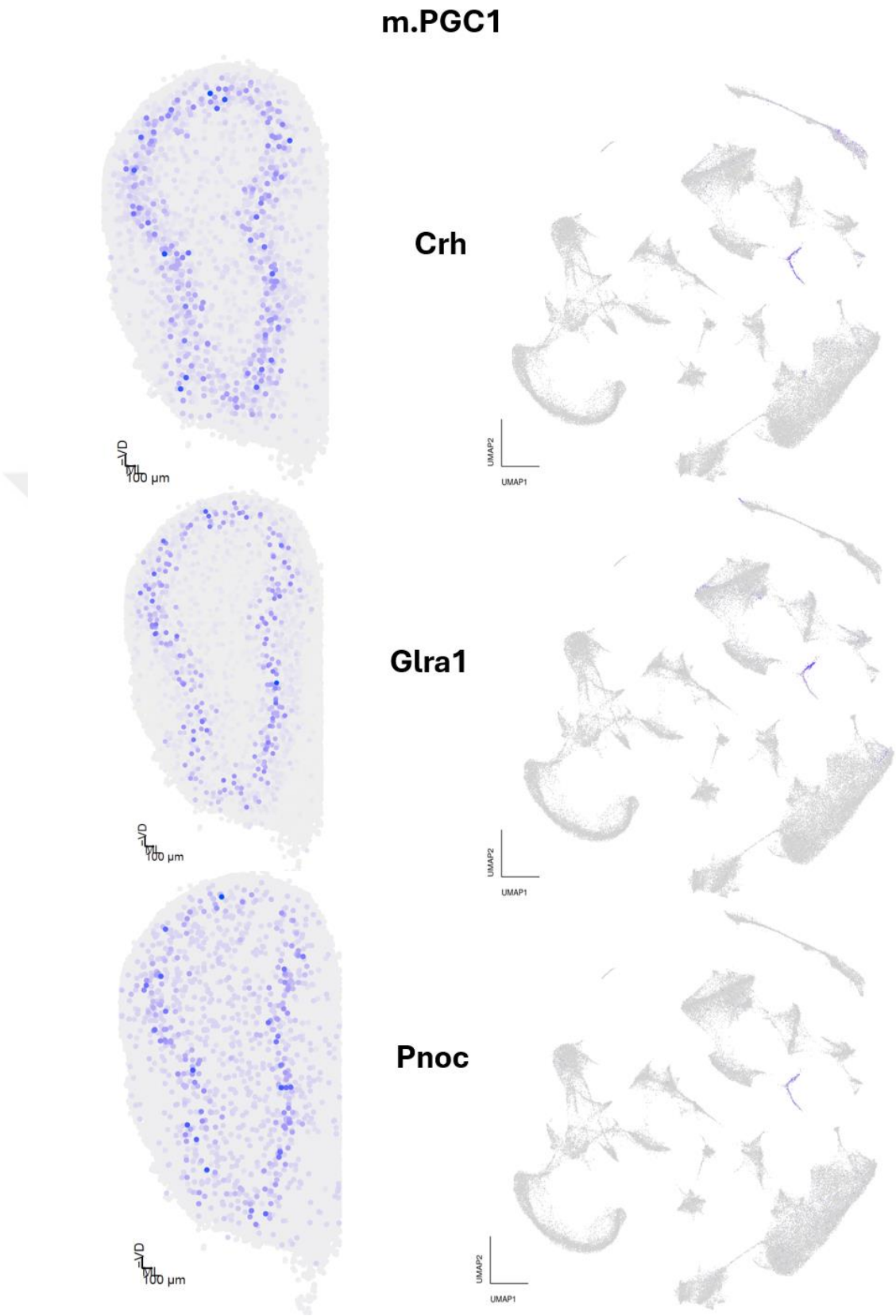

Figure 41. Cluster m.PGC1 marker genes: *Crh*, *Glra1*, *Pnoc* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

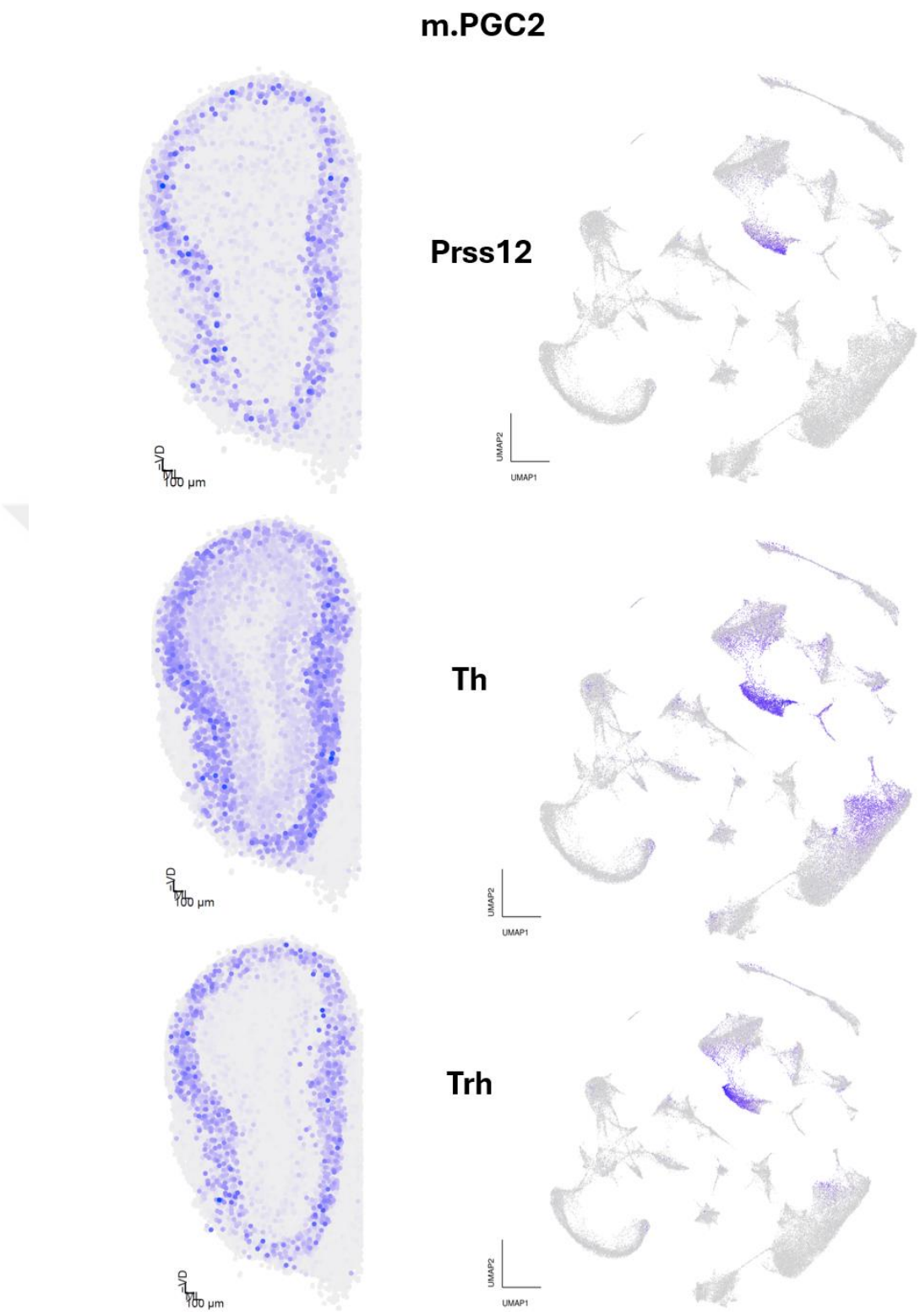

Figure 42. Cluster m.PGC2 marker genes: *Prss12*, *Th*, *Trh* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

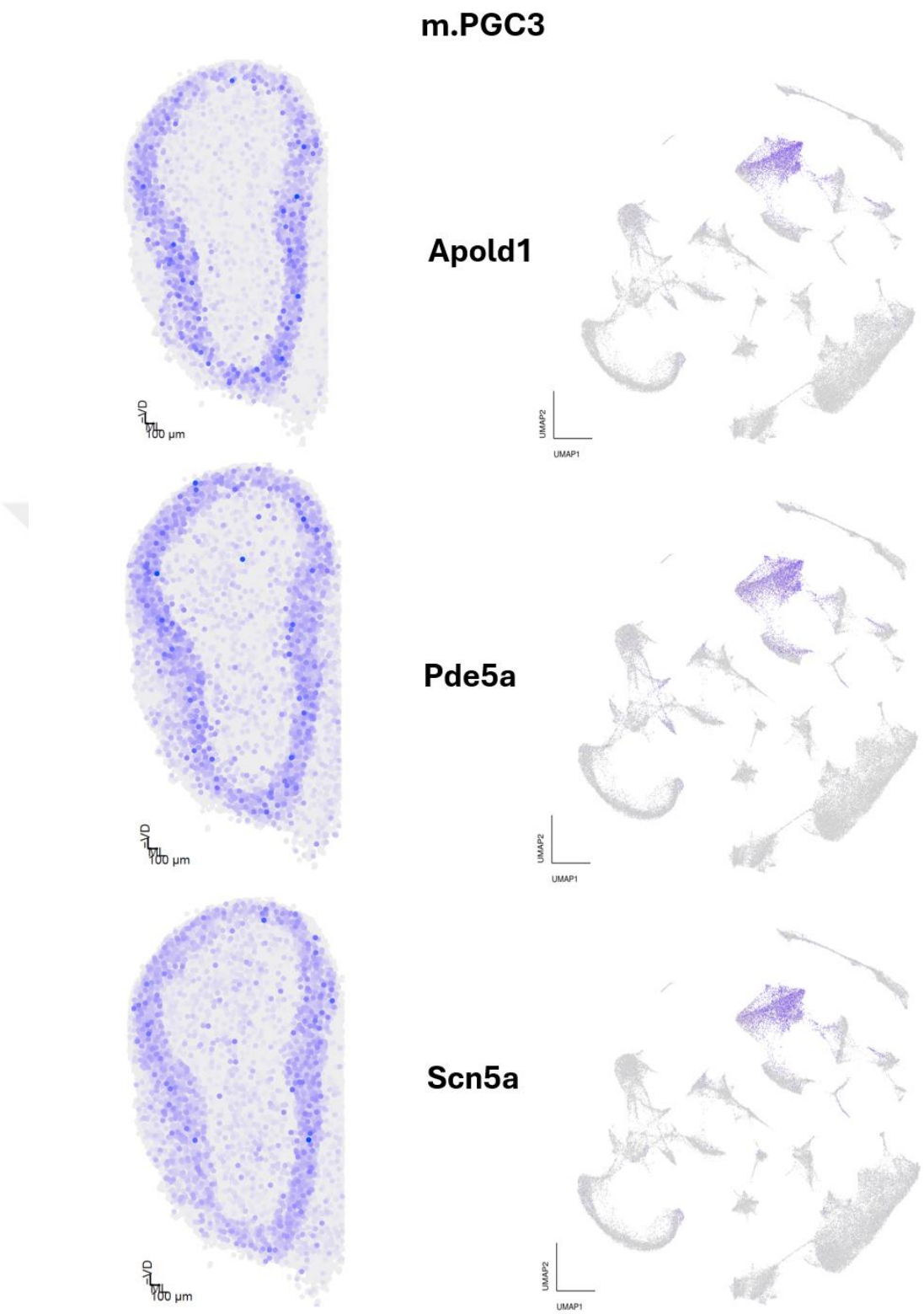

Figure 43. Cluster m.PGC3 marker genes: *Apold1*, *Pde5a*, *Scn5a* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

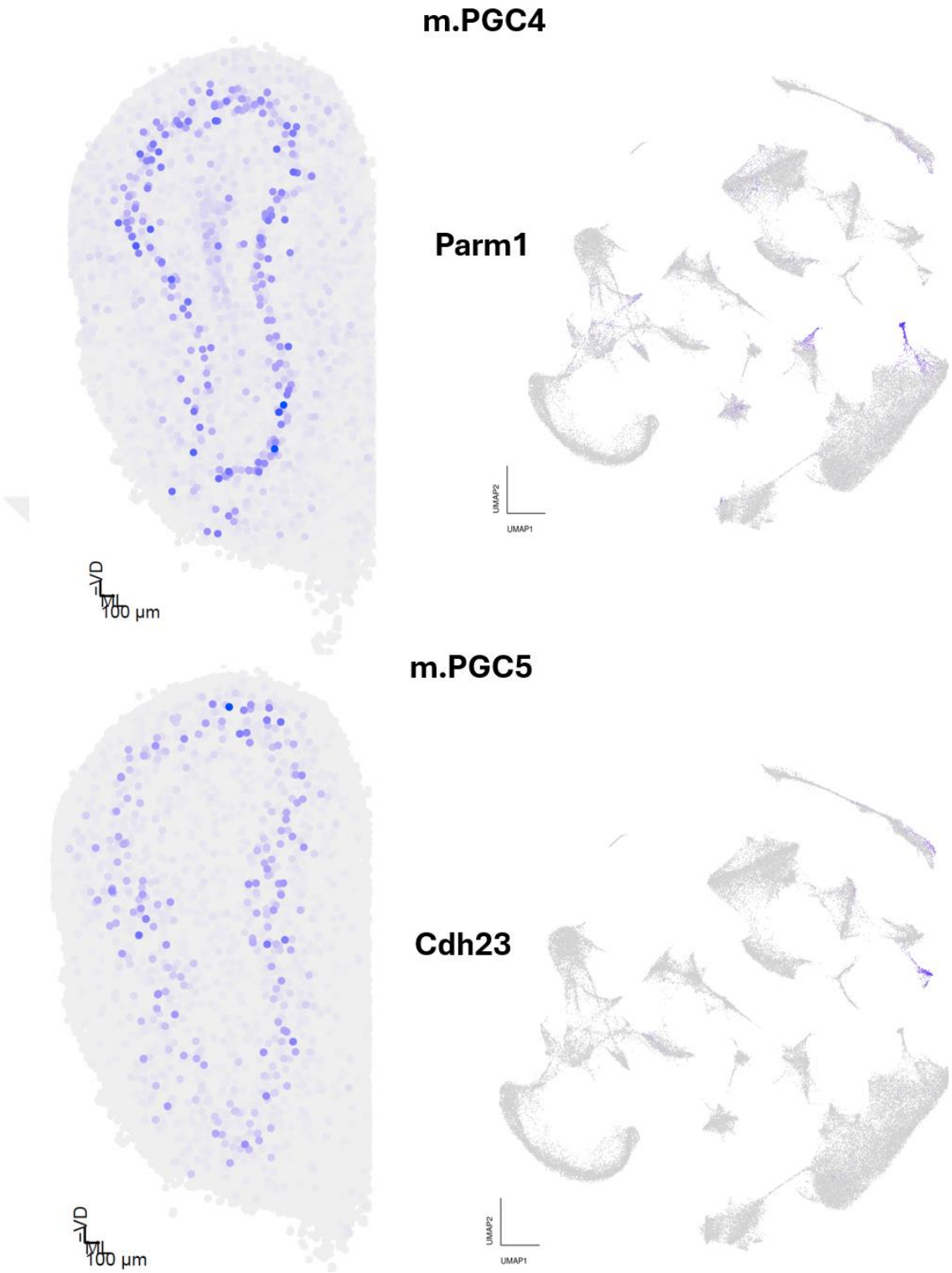

Figure 44. Cluster m.PGC4 marker gene: *Parm1*, and cluster m.PGC5 marker gene: *Cdh23* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

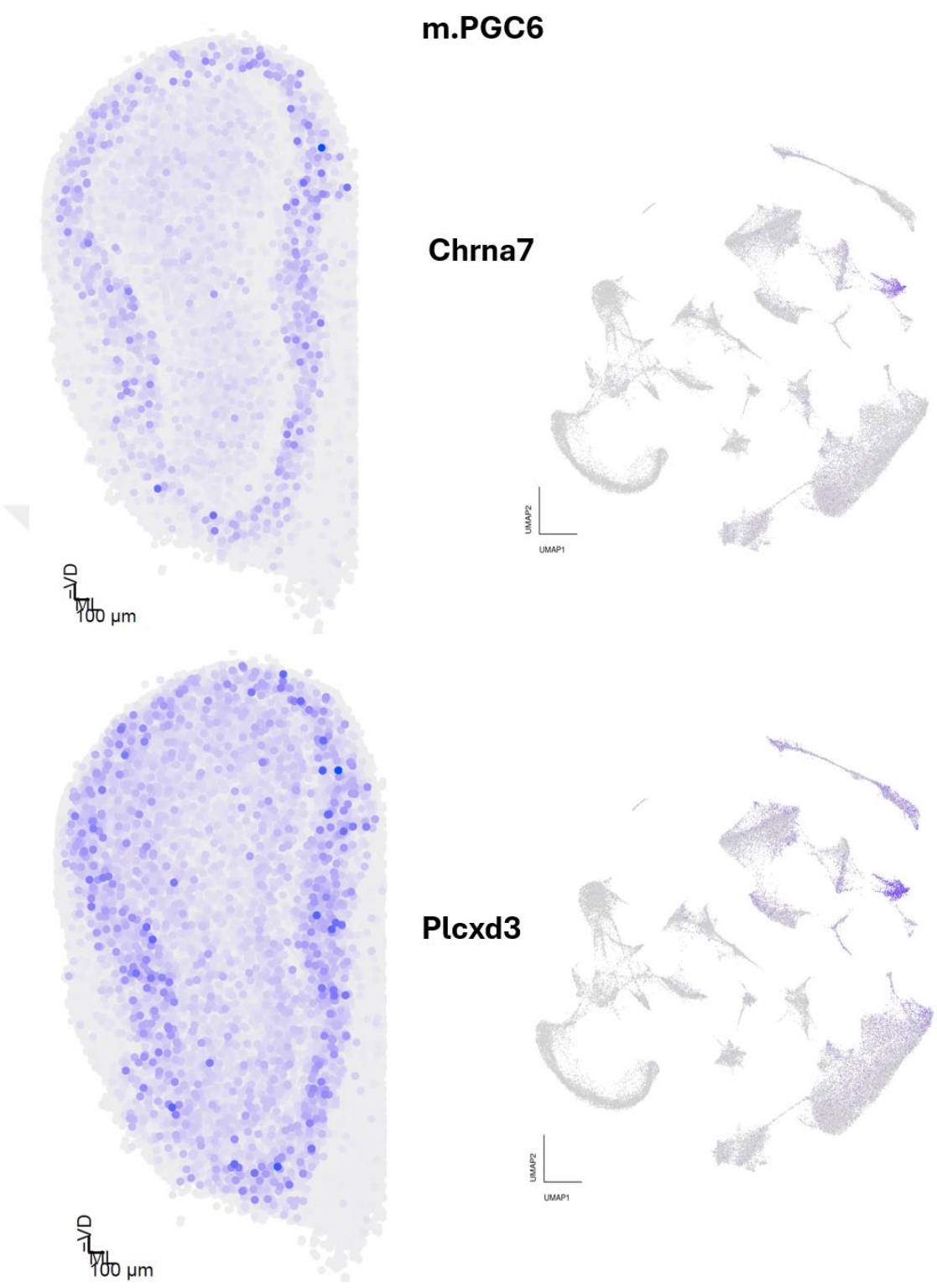

Figure 45. Cluster m.PGC6 marker genes: *Chrna7*, *Plcxd3* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

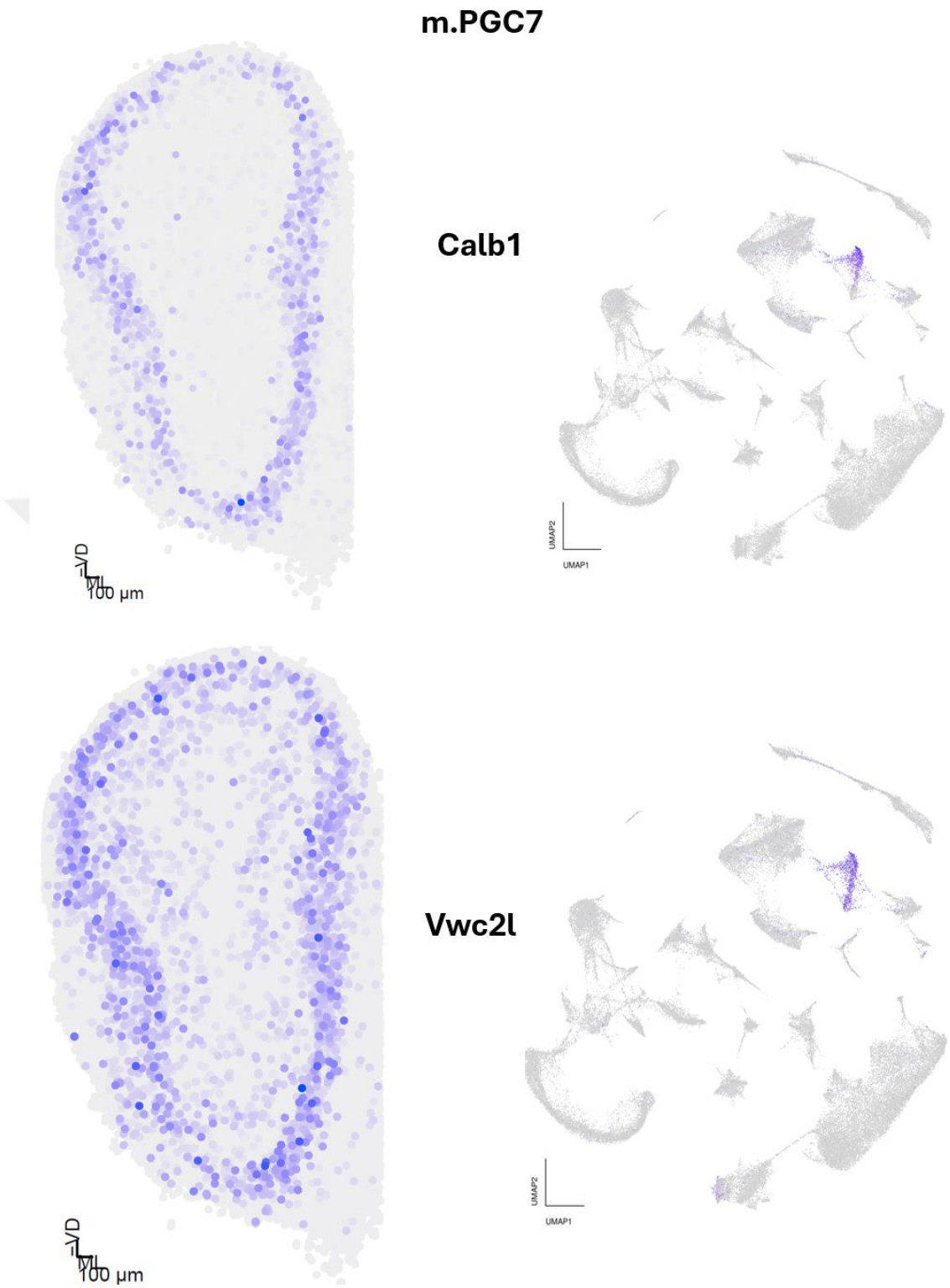

Figure 46. Cluster m.PGC7 marker genes: *Calb1, Vwc2l* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

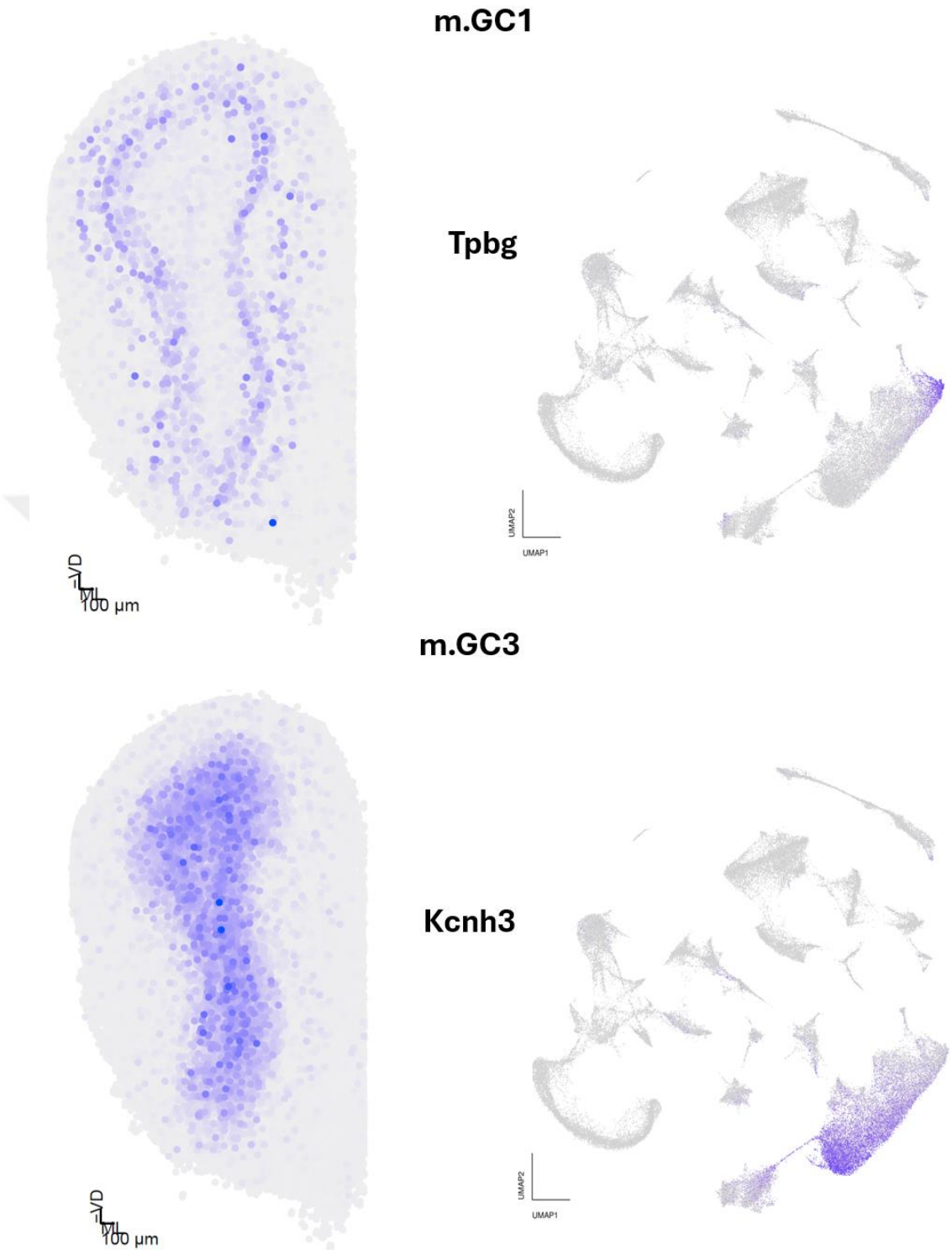

Figure 47. Cluster m.GC1 marker gene: *Tpbg*, and cluster m.GC3 marker gene: *Kcnh3* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

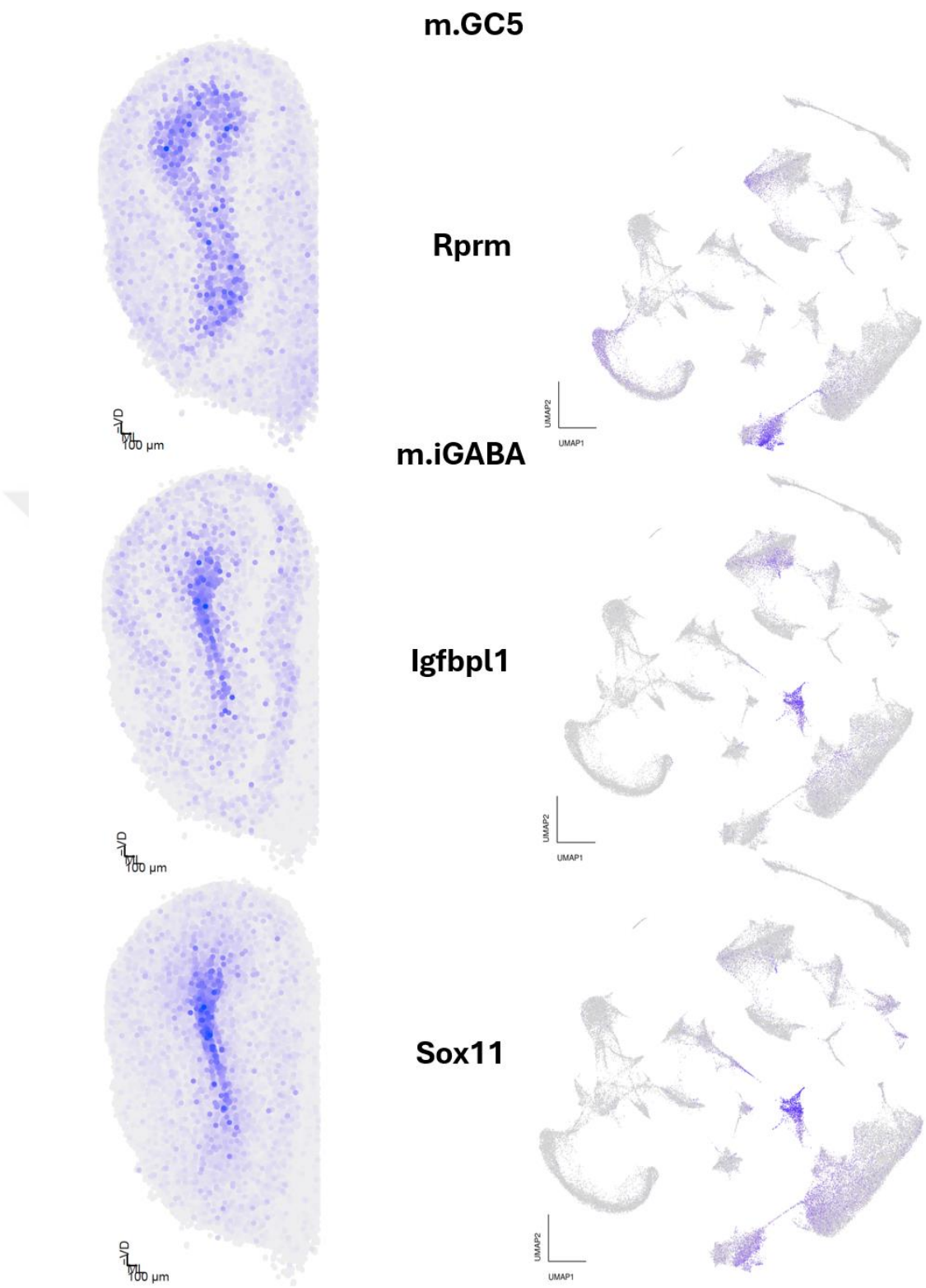

Figure 48. Cluster m.GC5 marker gene: *Rprm*, and cluster m.iGABA marker genes: *Igfbpl1*, *Sox11* in 2D spatial plot, and UMAP.

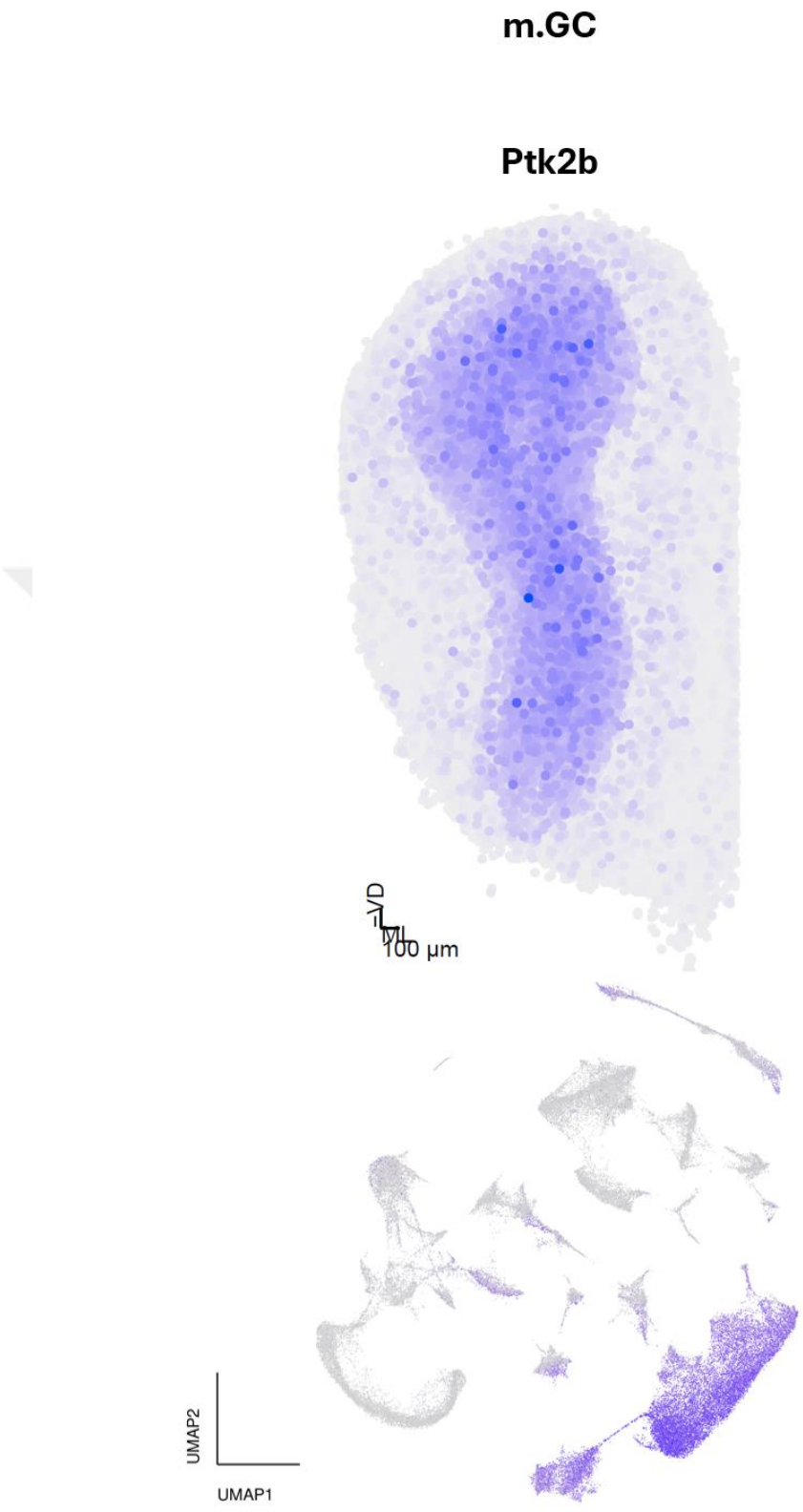

Figure 49. MERFISH granum cell common marker gene: *Ptk2b* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

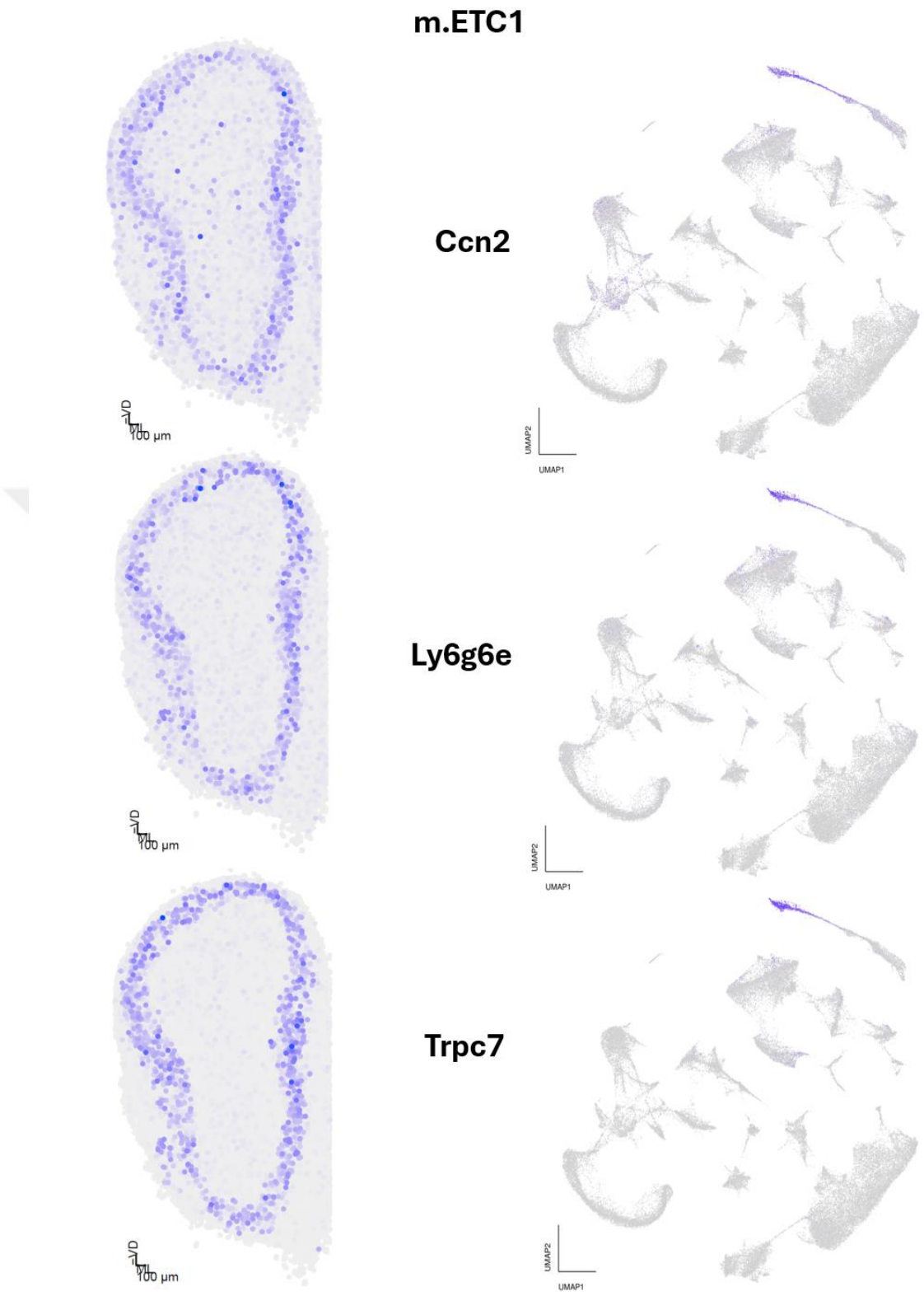

Figure 50. Cluster m.ETC1 marker genes: *Ccn2*, *Ly6g6e*, *Trpc7* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

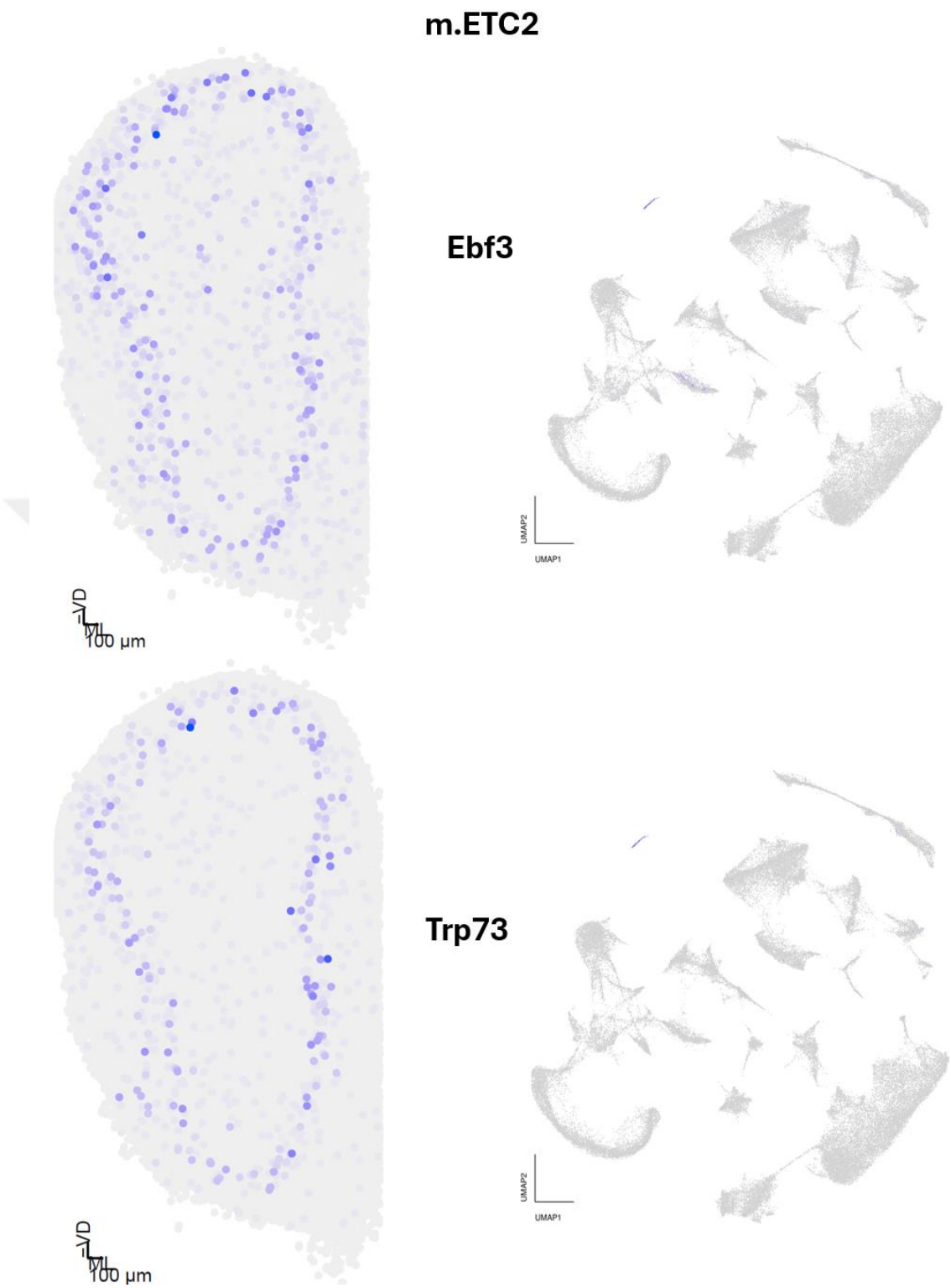

Figure 51. Cluster m.ETC2 marker genes: *Ebf3*, *Trp73* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

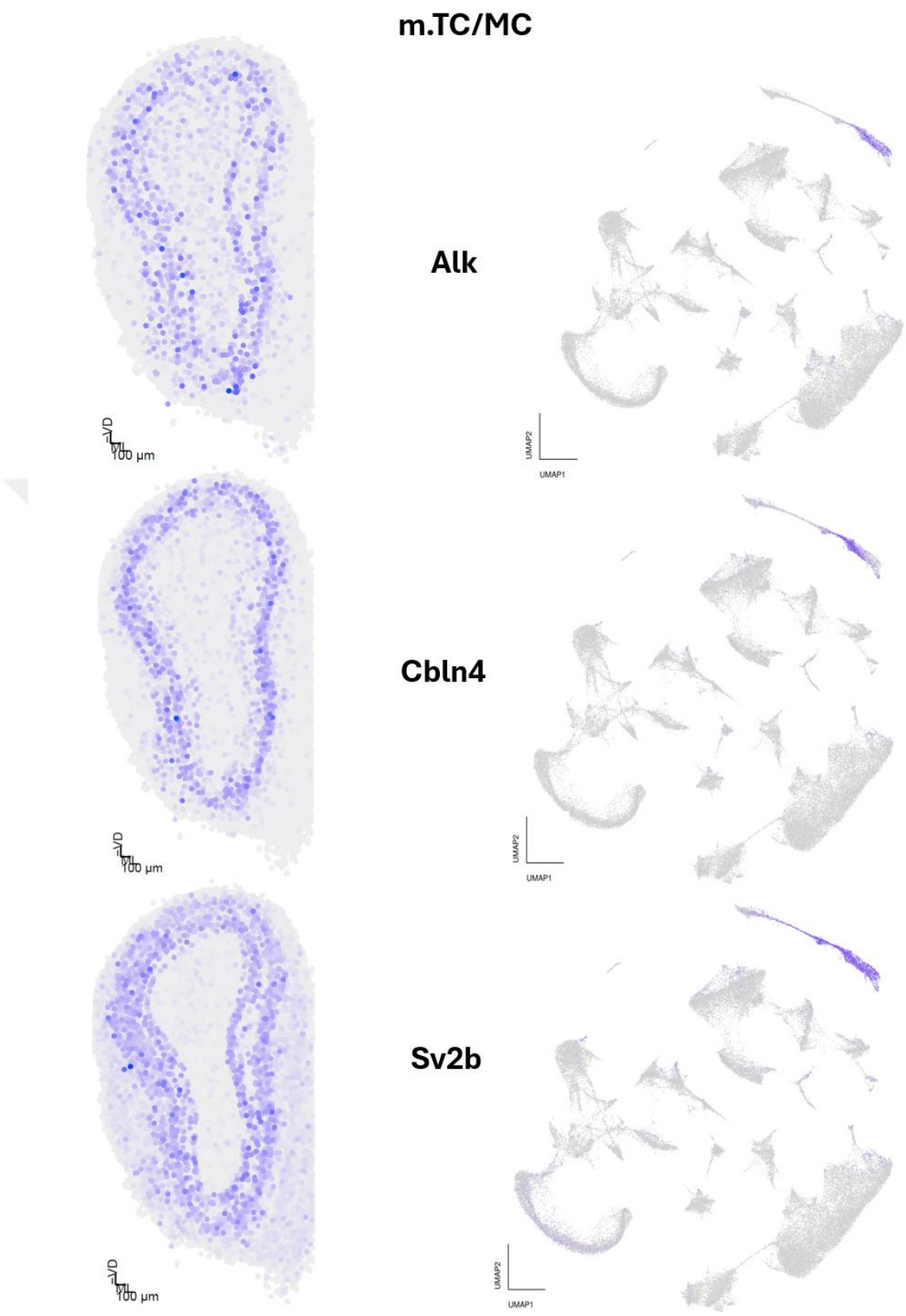

Figure 52. Cluster m.TC/MC marker genes: *Alk*, *Cbln4*, *Sv2b* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

4.1.2.3 Verification of MERFISH clusters with mouse scRNA-seq clusters

In order to validate our MERFISH clusters, MERFISH dataset, and mouse scRNA-seq dataset were integrated Canonical Correlation Analysis (CCA) (Figure 53). After integration, KMR matrix, and alluvial plot (along with a dendrogram attached) were created (Figure 54,55). It was observed that KMR matrix (created from harmonic means of KMR values.) follows a diagonal correlation which means MERFISH clusters are able to find their corresponding clusters in mouse scRNA-seq data.

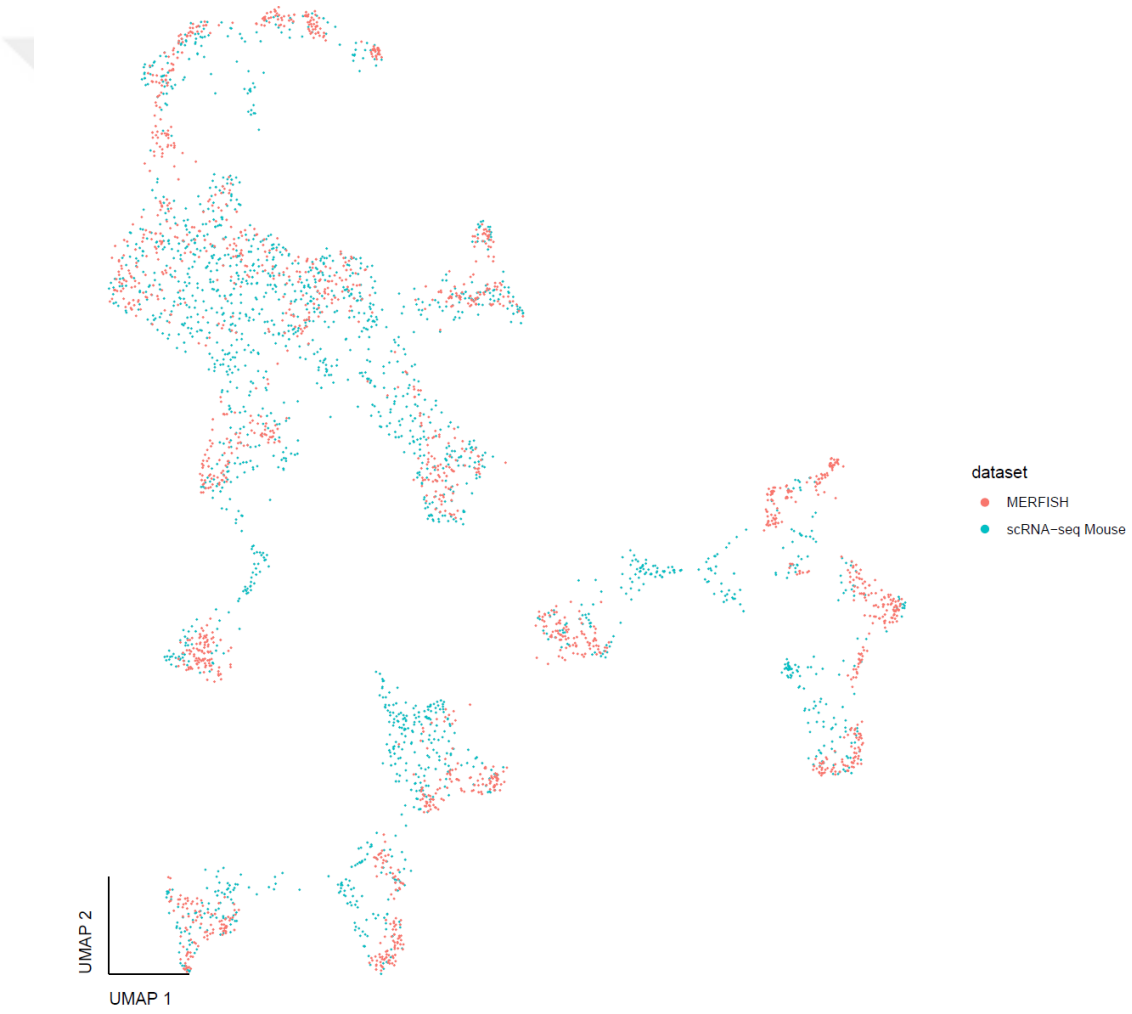

Figure 53. Integration of MERFISH, and mouse scRNA-seq datasets in UMAP.

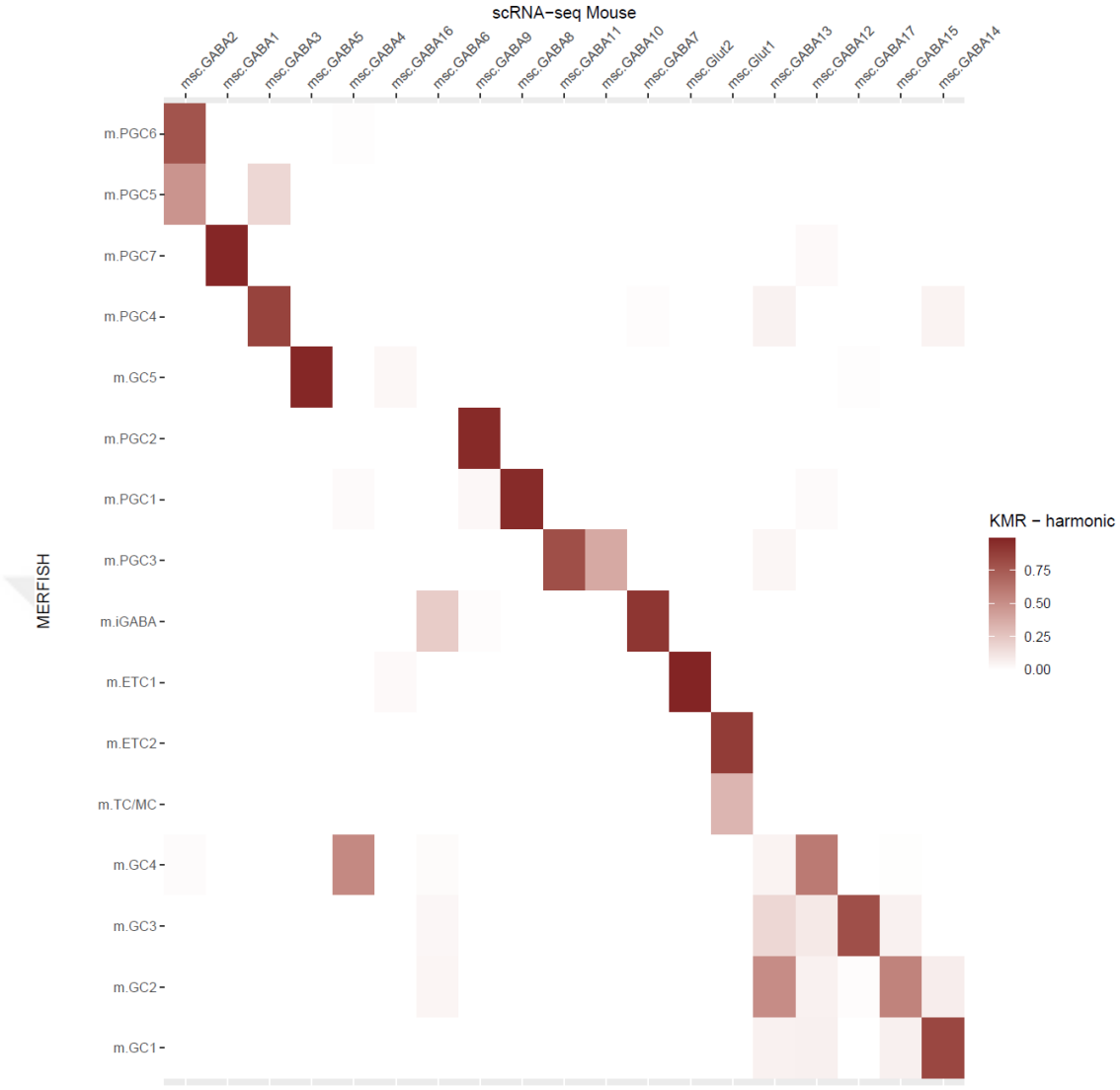

Figure 54. KMR matrix of MERFISH, and mouse scRNA-seq datasets clusters.

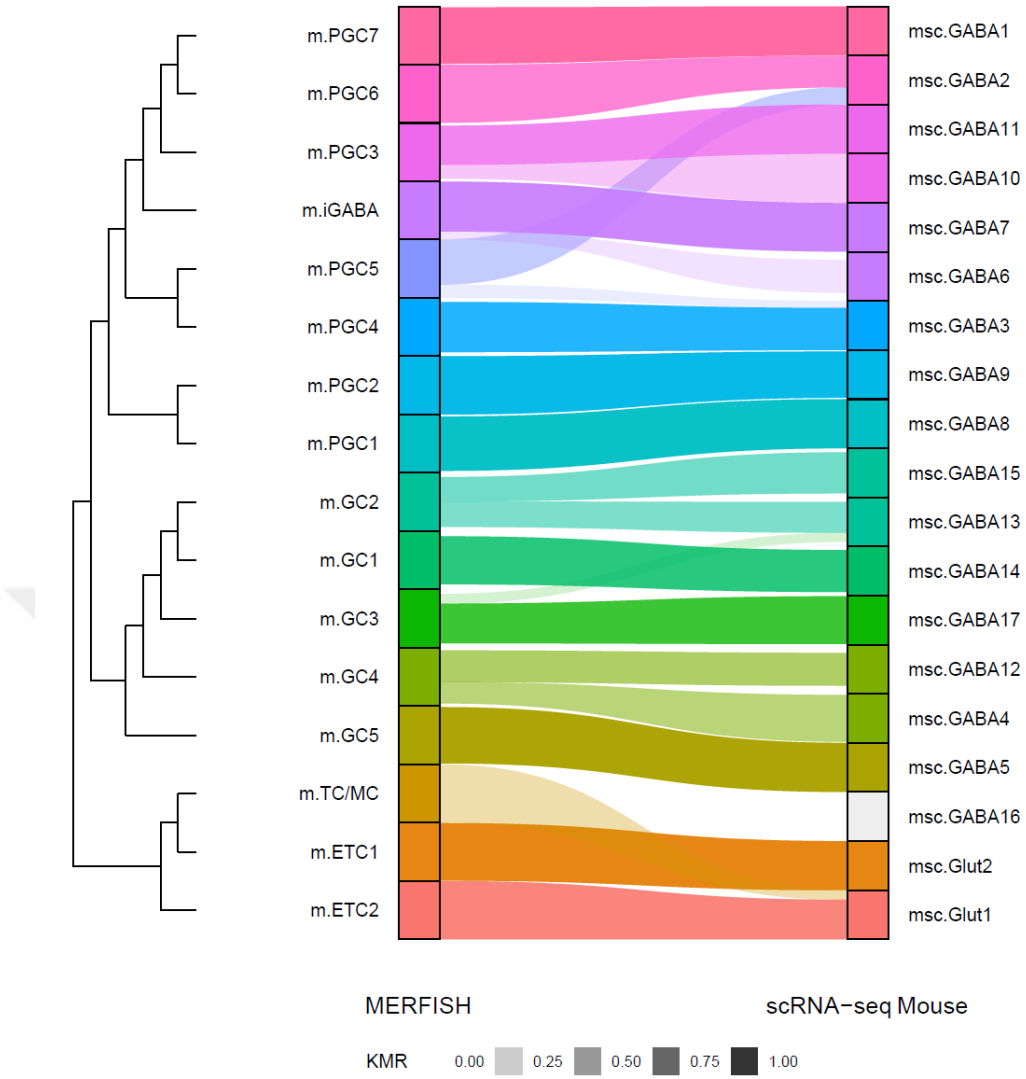

Figure 55. Alluvial plot of MERFISH, and mouse scRNA-seq datasets. Dendrogram was created based on genetic similarity of MERFISH clusters.

4.1.3 Integration of zebrafish and mouse scRNA-seq datasets revealed evolutionary connections between species

In order to explain evolutionary connections between mouse and zebrafish, mouse scRNA-seq data, and zebrafish scRNA-seq dataset were integrated Canonical Correlation Analysis (CCA). After integration, KMR matrix, and alluvial plot (along with a dendrogram attached) were created (Figure 56,57). It was observed that KMR matrix is able to connect multiple clusters however there are also the ones that remain to be orphan. Although GABAergic clusters mostly struggle to find their counterpart, glutamatergic clusters match to their corresponding clusters relatively (Figure 56,57). In addition to this analysis, an alluvial plot which has 4 columns (one for each dataset) was created in order to find the connection between spatial data clusters through scRNA-seq datasets (Figure 58).

Such comparison revealed that, both cell types in zebrafish and mouse are mostly conserved. Especially in glutamatergic clusters, z.MC3 cluster which is the glutamatergic cluster that is located in outer layer of the OB (Olfactory bulb), is able to find its correspondence in m. ETC1 which is the mouse glutamatergic cluster that is located in the most outer layer of mouse OB (Figure 58). Furthermore, other glutamatergic clusters also find their counterparts in alluvial plot (Figure 58).

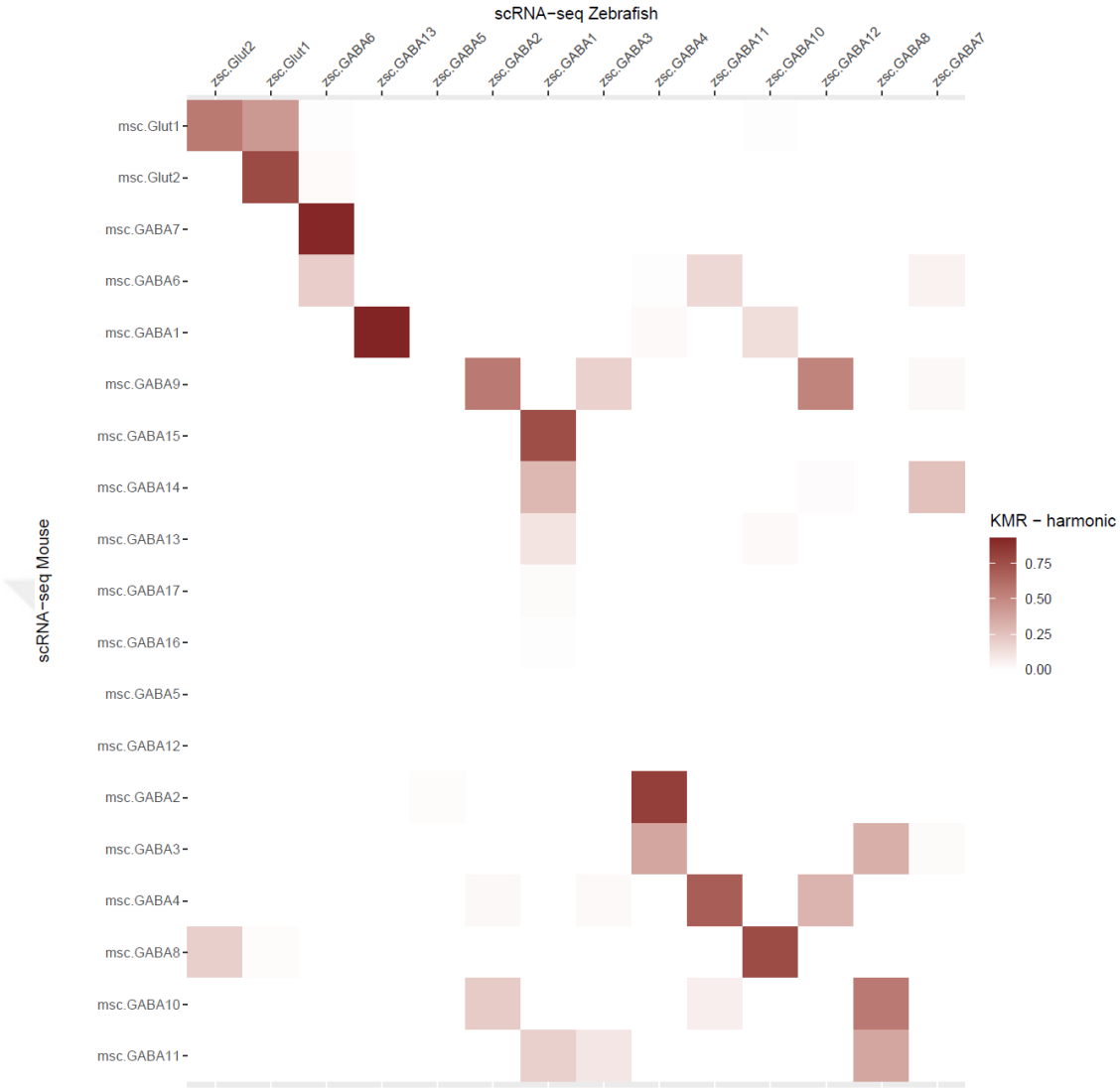

Figure 56. KMR matrix of zebrafish, and mouse scRNA-seq datasets clusters.

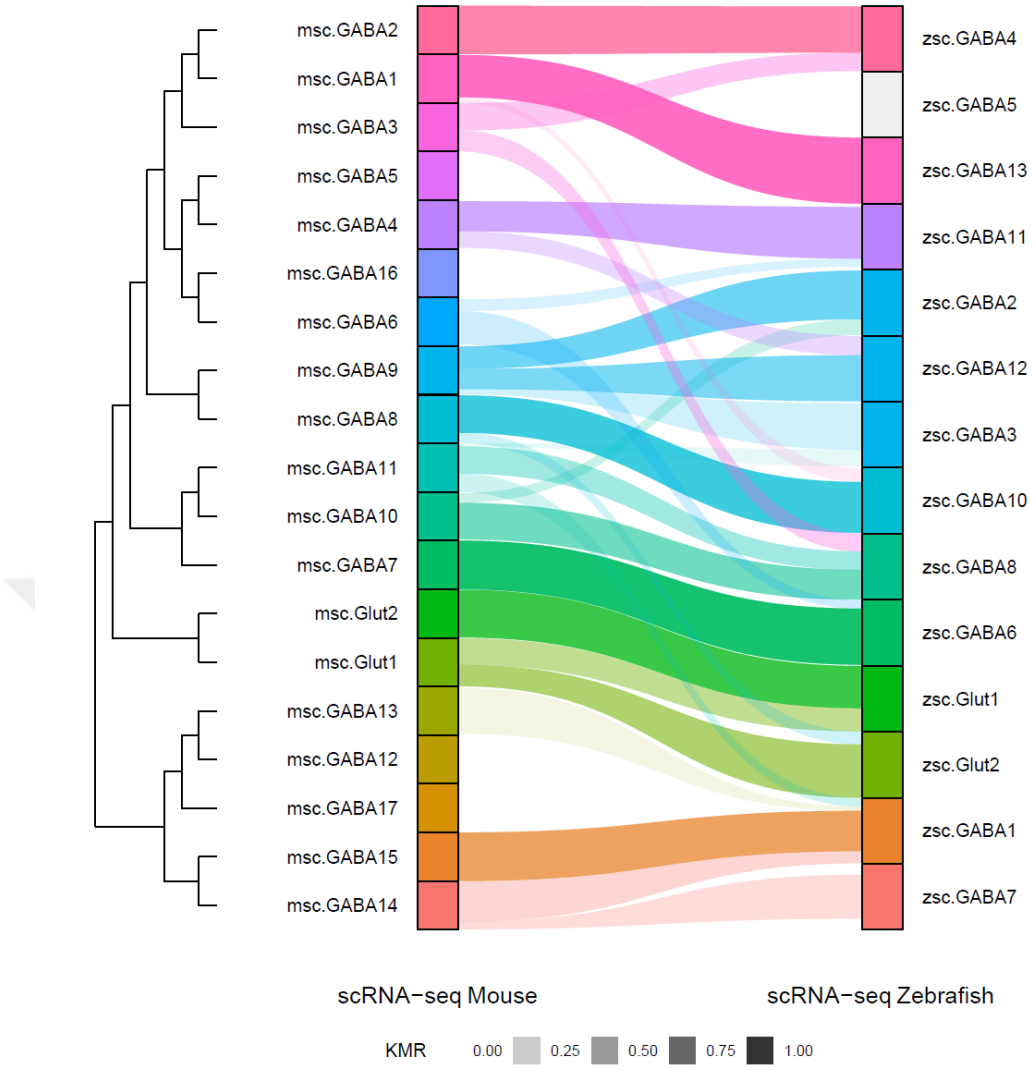

Figure 57. Alluvial plot of zebrafish, and mouse scRNA-seq datasets. Dendrogram was created based on genetic similarity of mouse scRNA-seq clusters.

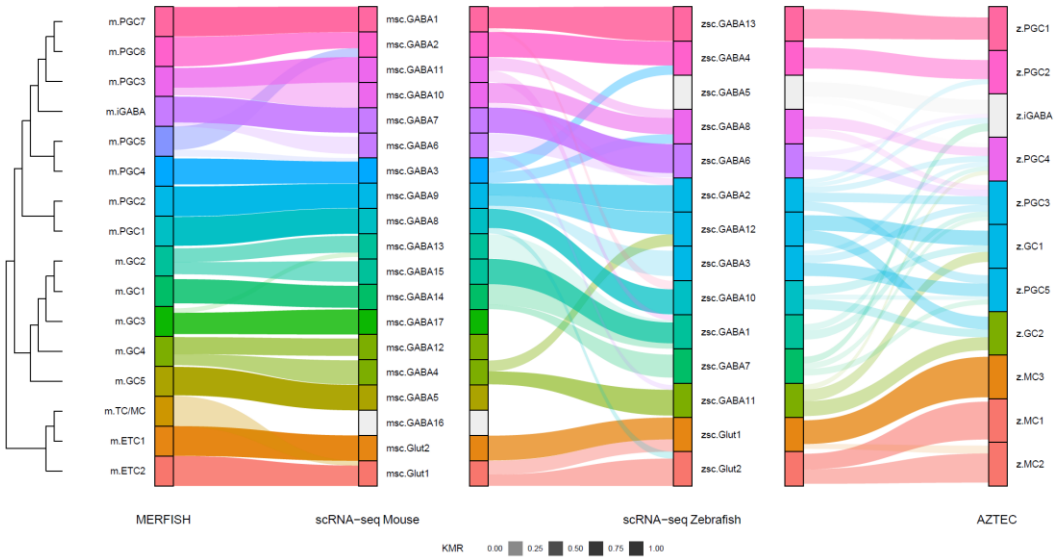

Figure 58. Alluvial plot of MERFISH, mouse scRNA-seq, scRNA-seq zebrafish, and AZTEC datasets. Dendrogram was created based on genetic similarity of MERFISH clusters.

4.2 Overcluster Analysis of Spatial Transcriptomics Datasets (AZTEC MERFISH) Revealed Topographical Organization in Olfactory Bulb

4.2.1 Overview of AZTEC and MERFISH spatial topography

Before diving into glutamatergic, and GABAergic overclustering analysis, all cell types were assigned to a color code, and visualized using their spatial coordinates, and UMAP coordinates (Figure 59,60). According to these 3D spatial plots, it is clear that all cell types show topographical organization which will be detailed in later sections.

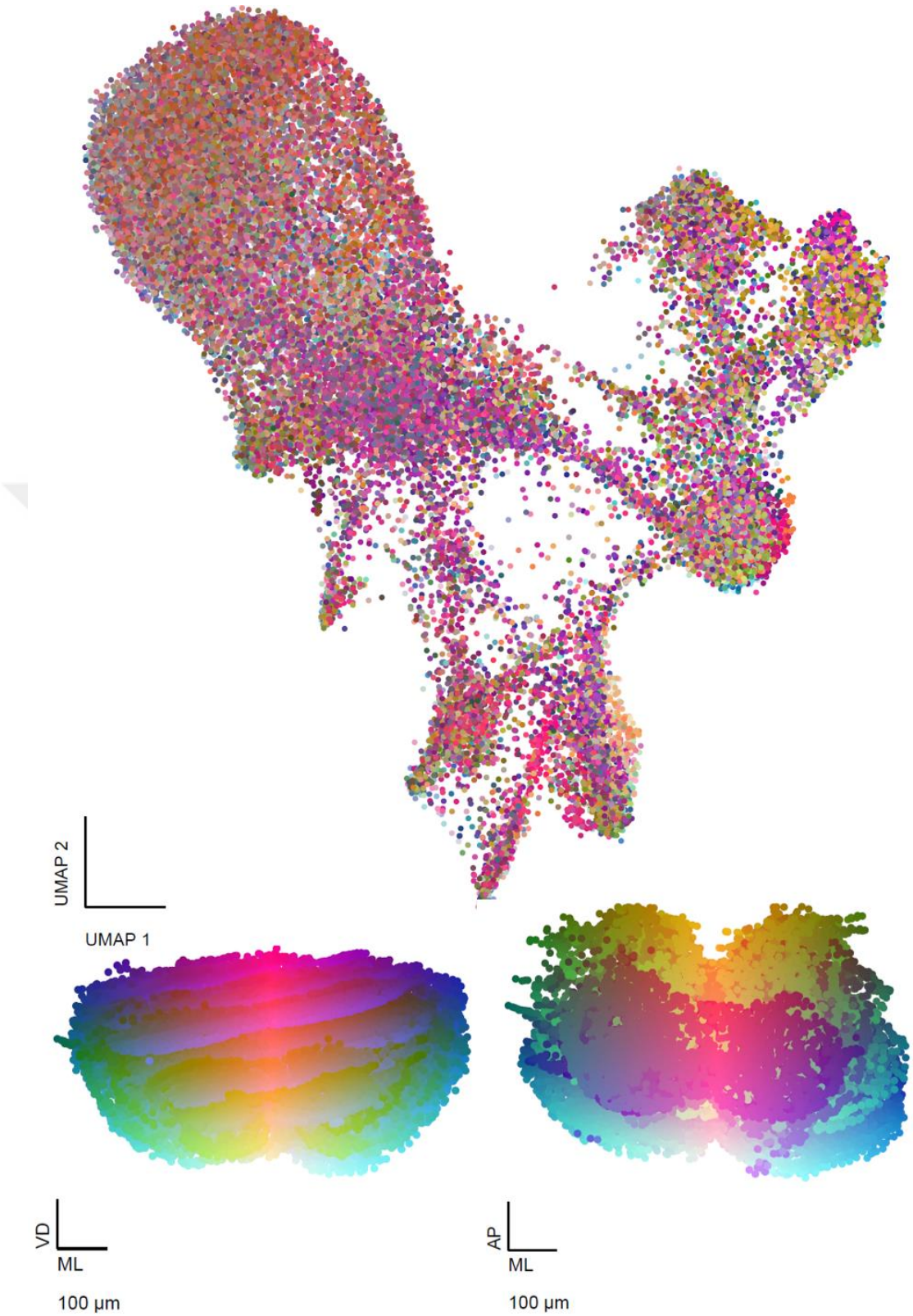

Figure 59. 3D spatial locations of AZTEC cells in both UMAP, and 2D spatial maps. VD: Ventral Dorsal, ML: Medial Lateral, AP: Anterior Posterior.

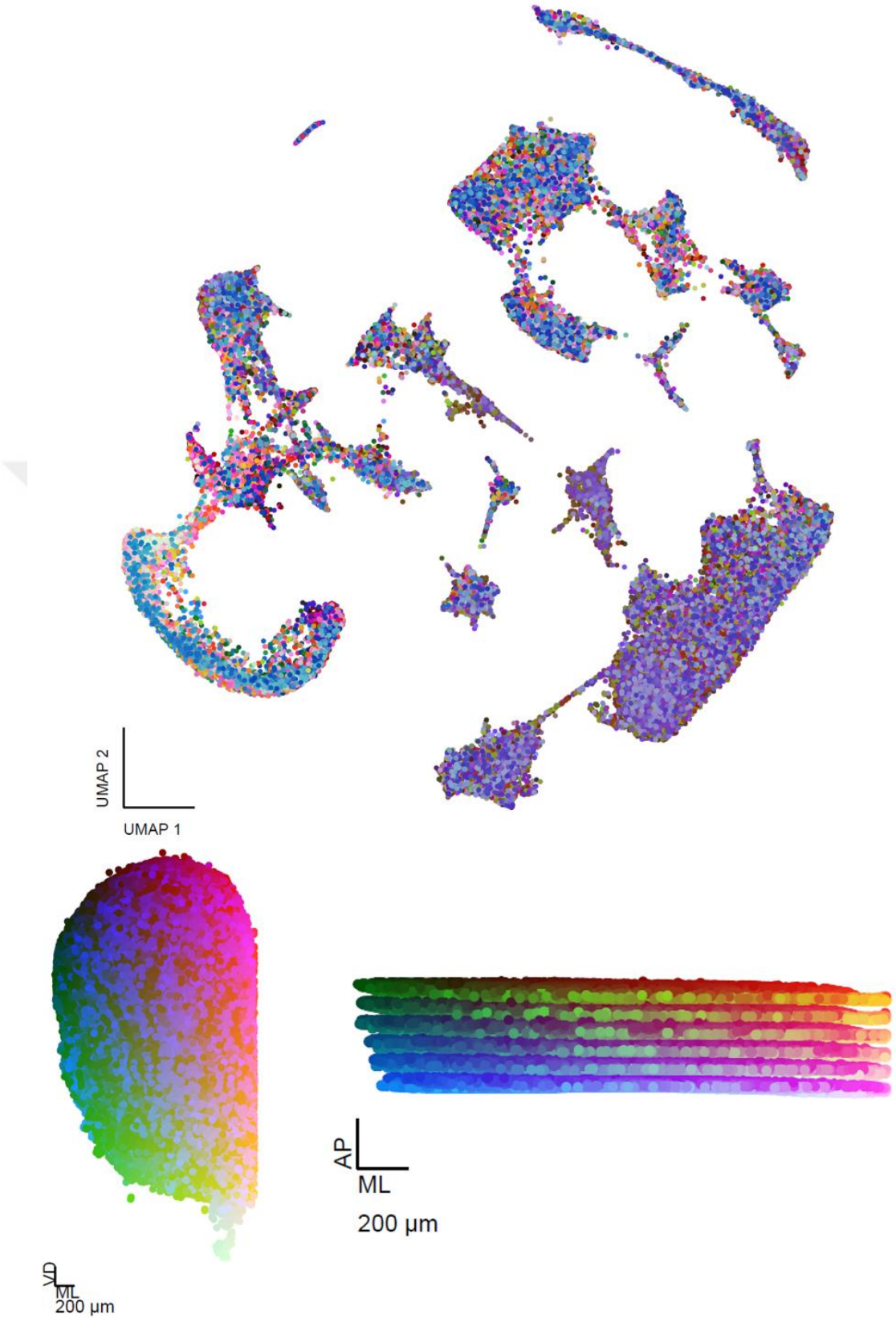

Figure 60. 3D spatial locations of MERFISH cells in both UMAP, and 2D spatial maps. VD: Ventral Dorsal, ML: Medial Lateral.

4.2.2 Overcluster of glutamatergic cells

4.2.2.1 Zebrafish spatial transcriptomics (AZTEC) dataset

Before overclustering AZTEC glutamatergic (GLUT) clusters, pairwise correlations of cells vs distance were plotted for all glutamatergic clusters. In this graph the spatial coordinates of the cells were also shuffled and represented in the graph as a control group (Figure 61).

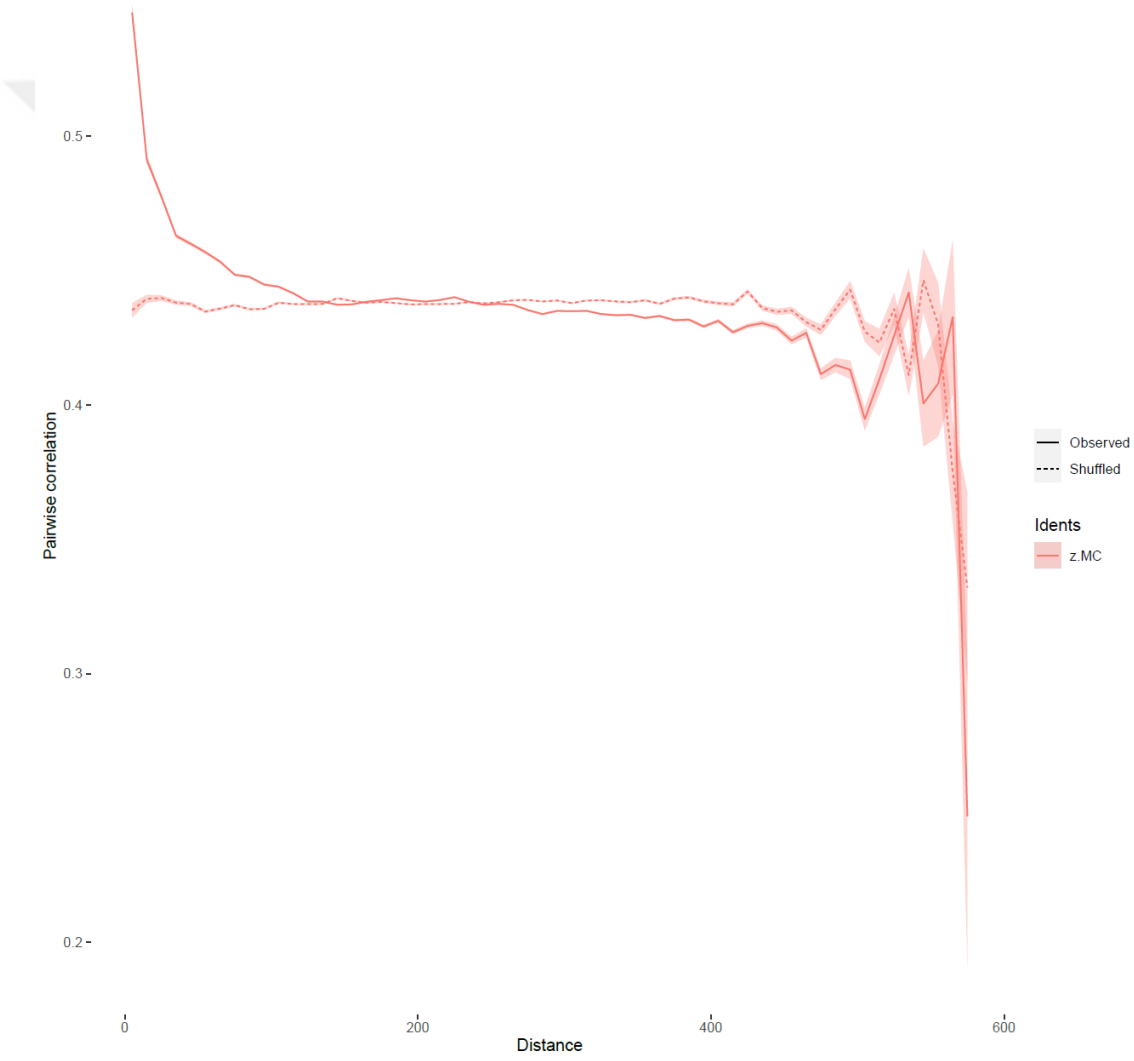

Figure 61. AZTEC GLUT cells correlation vs distance plot. MC: Mitral Cells.

Before diving into zebrafish glutamatergic overclustering analysis, all cell types were assigned to a color code, and visualized using their spatial coordinates, and UMAP coordinates (Figure 62).

Cells were analyzed using old normalization values but using 3.0 resolution in Louvain community detection method. In total of 20 clusters were identified, clusters were ordered based on genetic similarity, and UMAP was constructed (Figure 63). Spatially organized clusters were noted and plotted using spatial locations (Figure 64). It observed that some clusters have spatial patterns in OB. For example, cluster 8 and 19 are present at dorso-medial OB, cluster 18 is present at ventro-medial OB, cluster 17 is present at dorso-lateral OB, cluster 9, 12, and 16 are present at dorso-ventral whereas cluster 15 and 13 are present mostly at lateral OB. These cluster are mostly present at outer layer of OB, however there are clusters which are present at inner layer of OB as well which is cluster 10 (Figure 64).

Afterwards, differential expression analysis was applied to AZTEC GLUT clusters. Wilcoxon rank-sum test was used in this analysis, and top 10 most expressed genes for each cluster was represented in gene expression heatmap (Figure 64). As a result of differential expression analysis, 3 important marker genes abundantly expressed in specific regions of OB were identified. These genes are as follows: *fezf2* for ventro-medial, *pax6a* for dorso-medial, *tacr3l* for dorso-lateral OB (Figure 65).

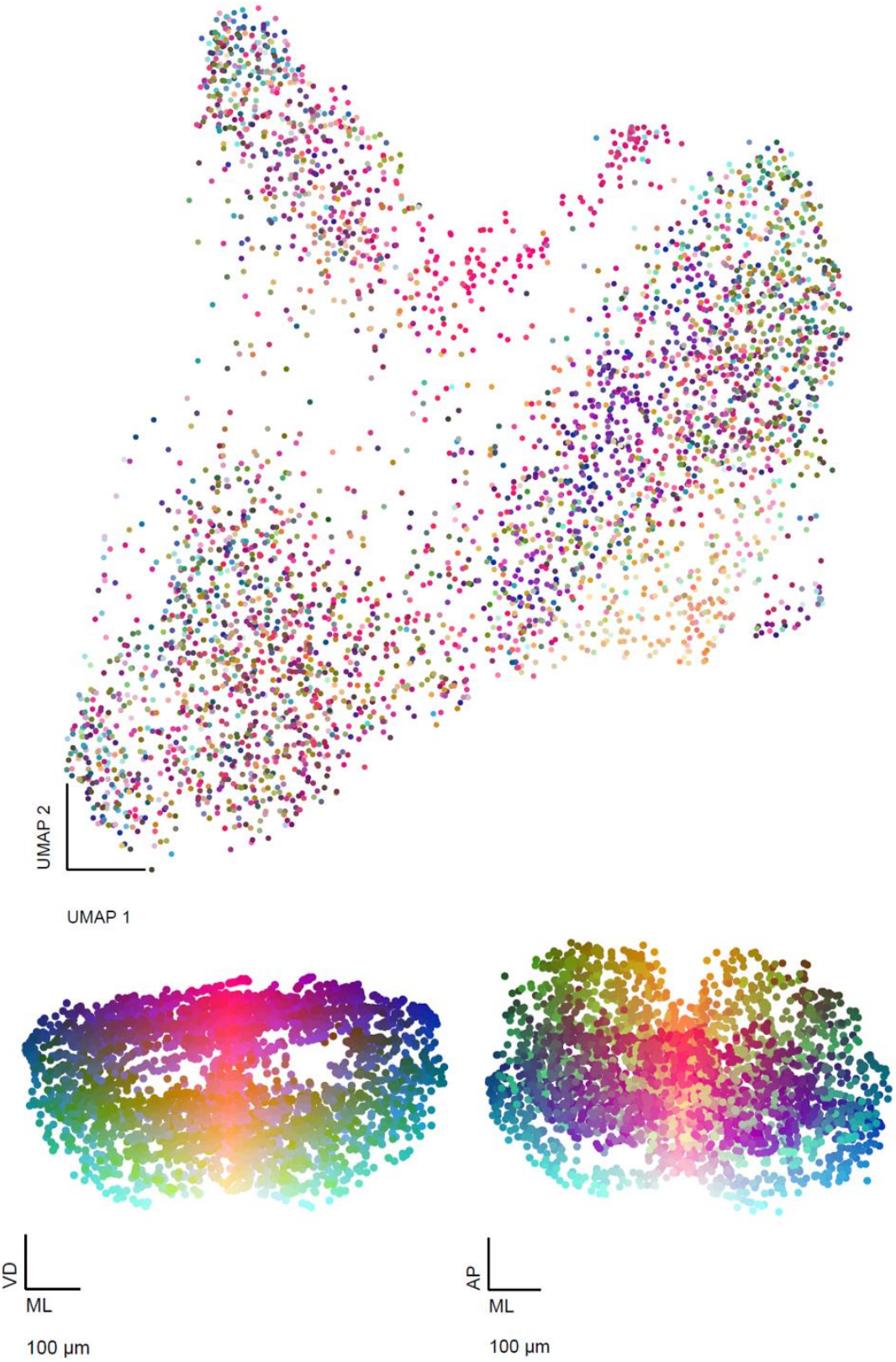

Figure 62. 3D spatial locations of AZTEC GLUT cells in both UMAP, and 2D spatial maps. VD: Ventral Dorsal, ML: Medial Lateral, AP: Anterior Posterior.

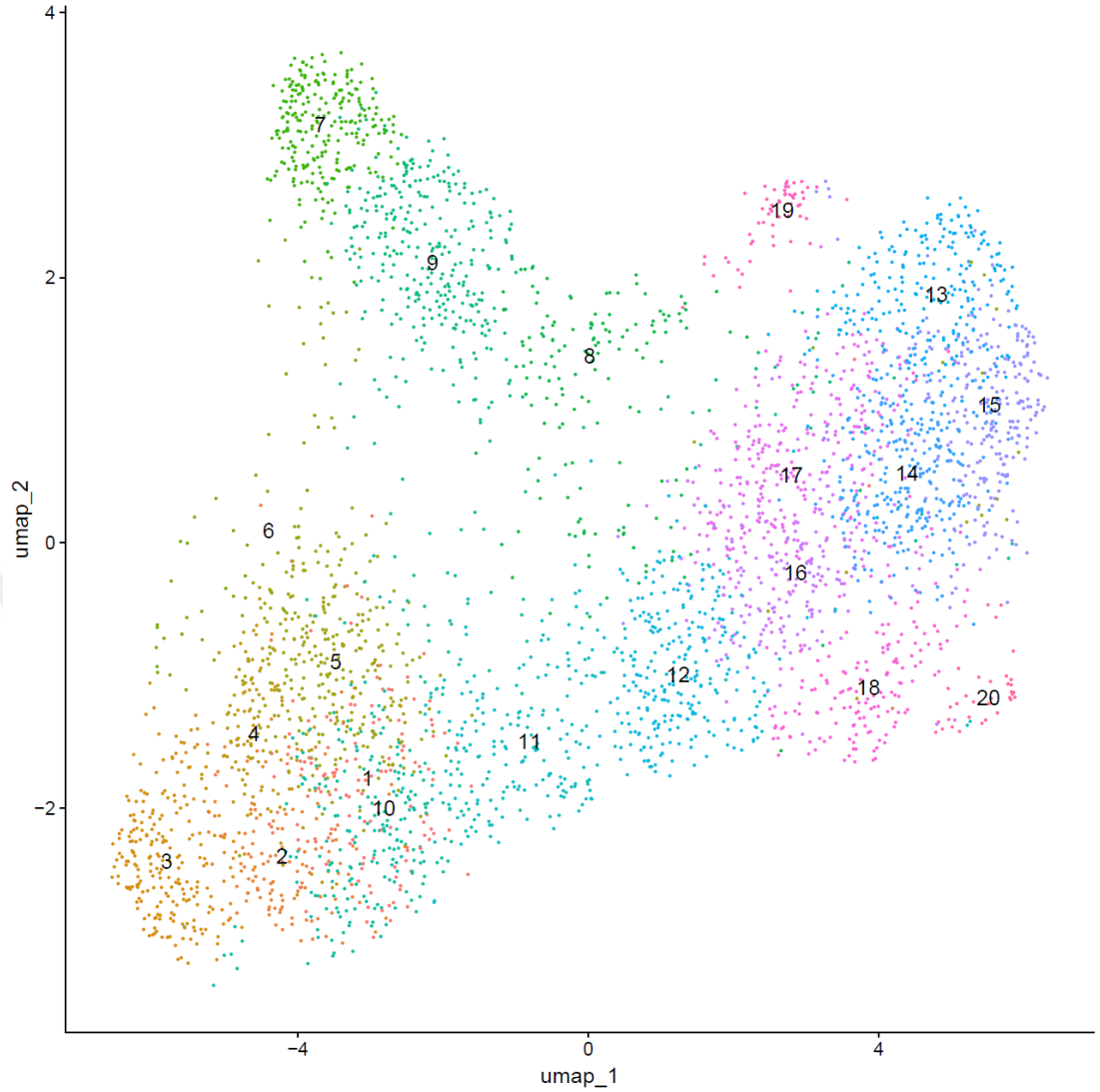

Figure 63. Overclustered AZTEC GLUT clusters in UMAP.

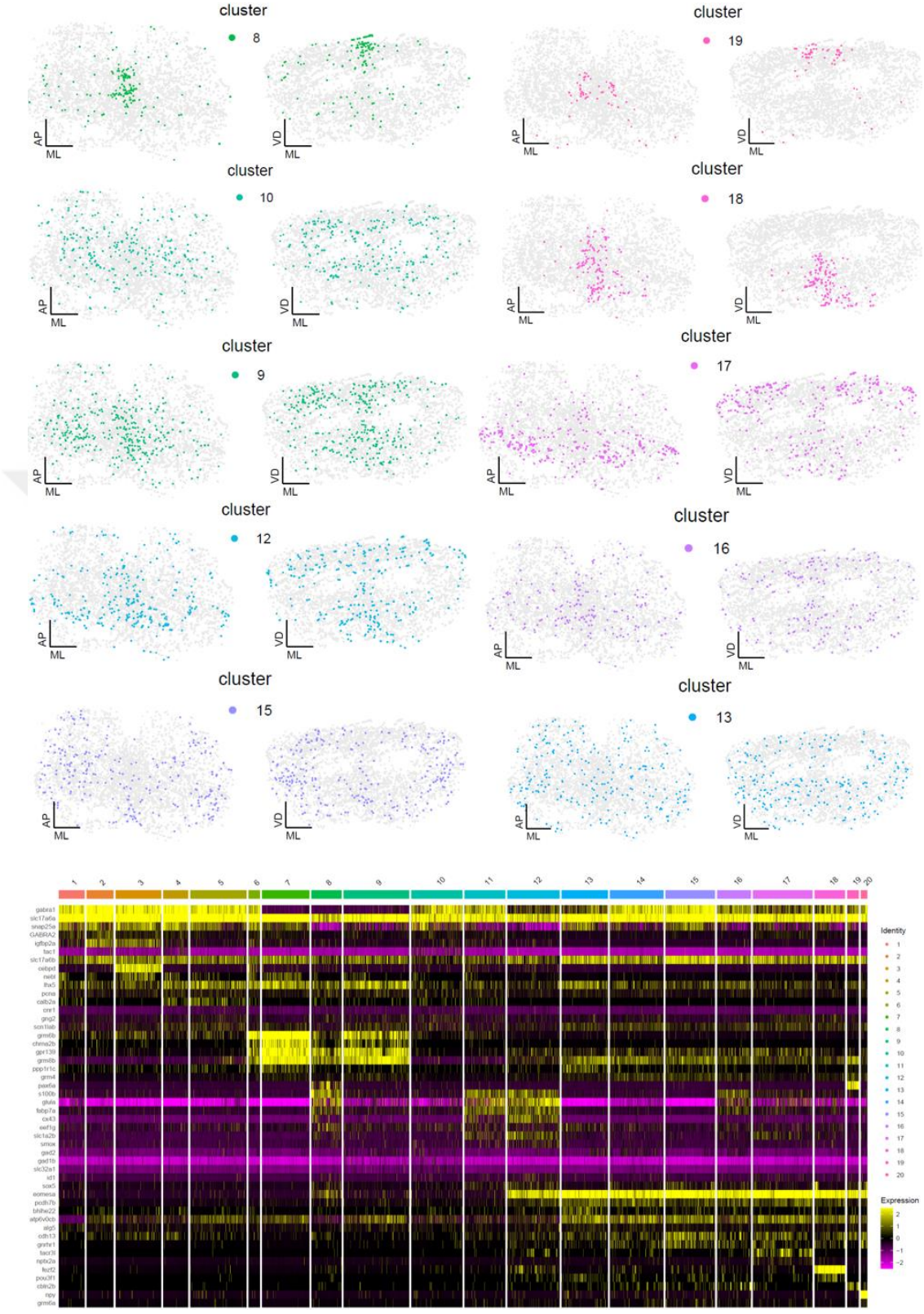

Figure 64. 2D spatial locations of spatially organized AZTEC GLUT clusters and all clusters' gene expression heatmap. VD: Ventral Dorsal, ML: Medial Lateral, AP: Anterior Posterior.

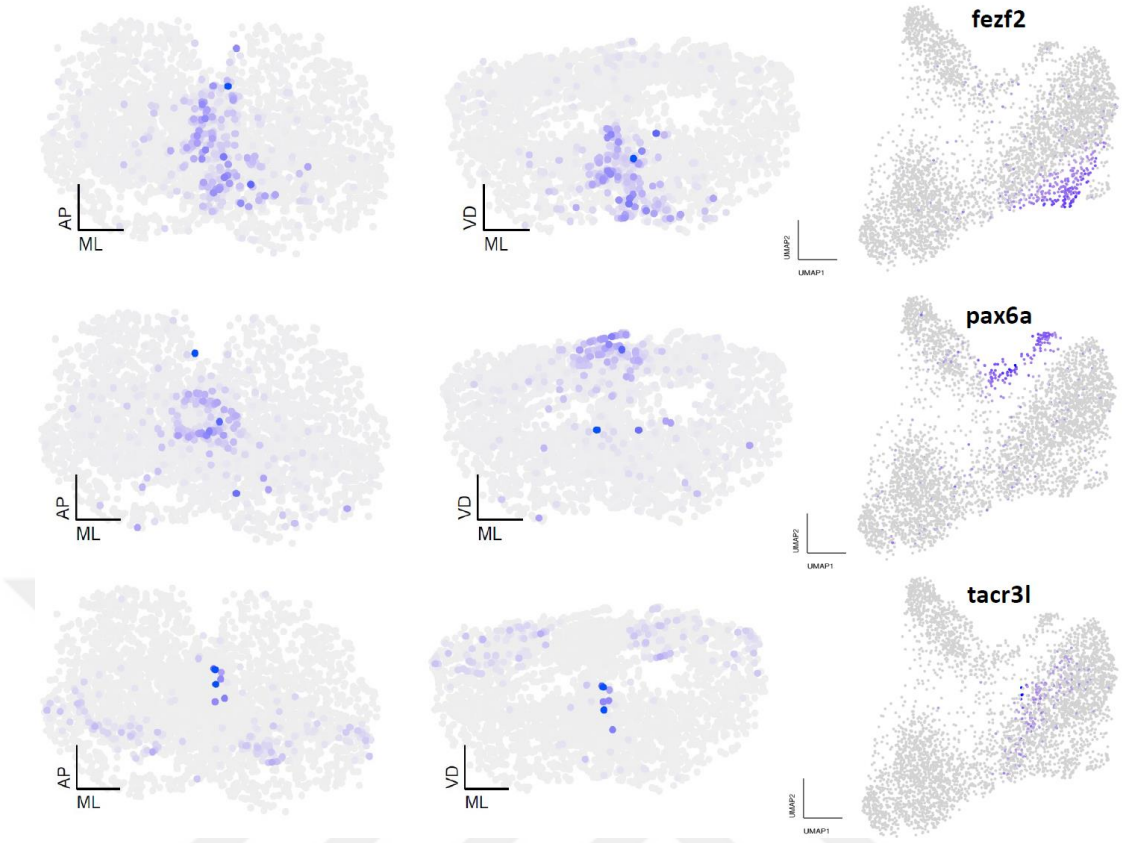

Figure 65. AZTEC GLUT spatial marker genes expression: *fezf2* for ventro-medial, *pax6a* for dorso-medial, *tacr3l* for dorso-lateral. VD: Ventral Dorsal, ML: Medial Lateral, AP: Anterior Posterior.

4.2.2.2 Mouse spatial transcriptomics (MERFISH) dataset

Before overclustering MERFISH glutamatergic (GLUT) clusters, pairwise correlations of cells vs distance were plotted for m.ETC1, m.ETC2, m.TC/MC clusters separately. In this graph the spatial coordinates of the cells were also shuffled and represented in the graph as a control group (Figure 66).

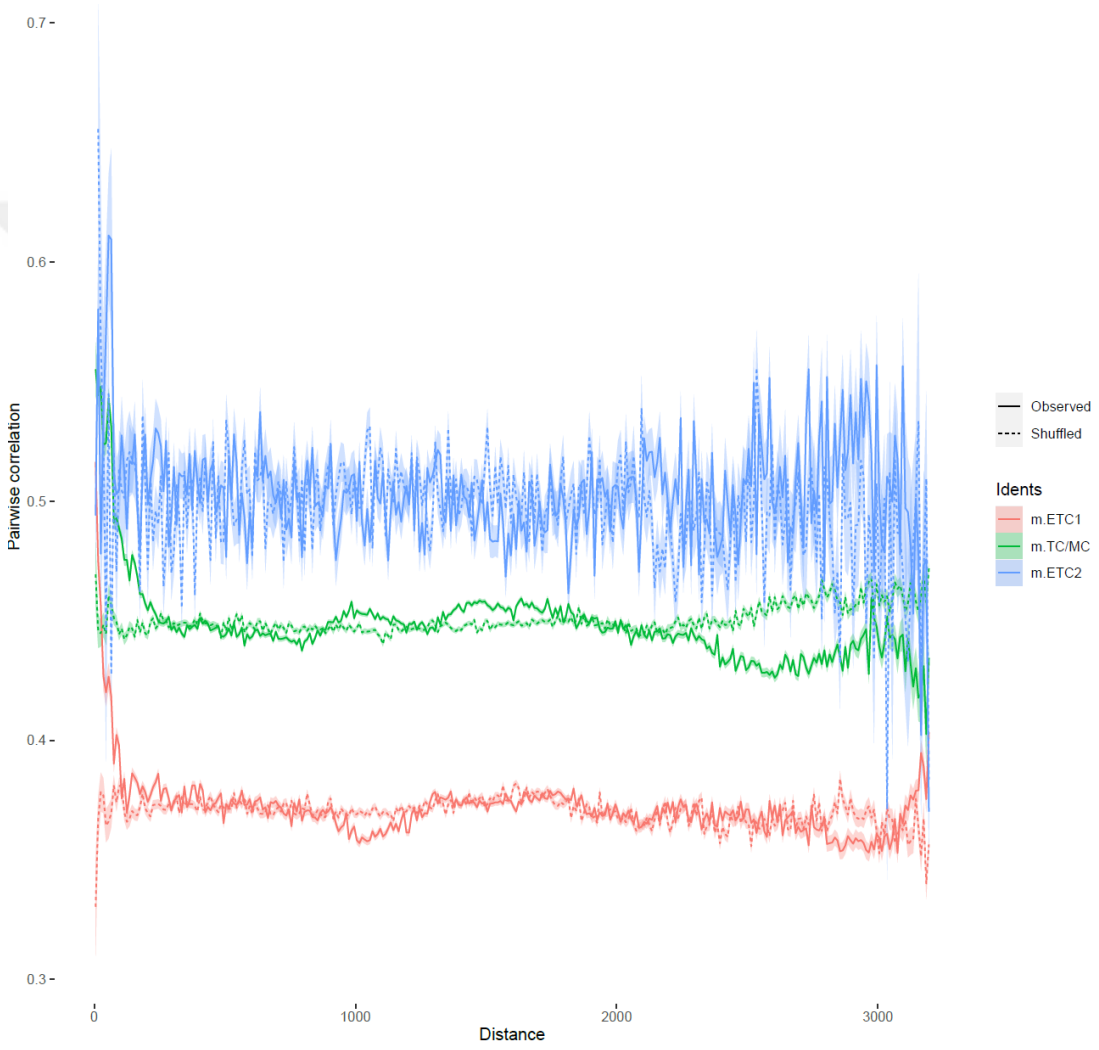

Figure 66. MERFISH GLUT cells correlation vs distance plot. ETC: External Tufted Cells, TC/MC: Tufted Cells, and Mitral Cells mix cluster.

Before diving into mouse glutamatergic overclustering analysis, all cell types were assigned to a color code, and visualized using their spatial coordinates, and UMAP coordinates (Figure 67).

Cells were analyzed using old normalization values but using 3.0 resolution in Louvain community detection method. In total of 21 clusters were identified, clusters were ordered based on genetic similarity, and UMAP was constructed (Figure 68). Spatially organized clusters were noted and plotted using spatial locations (Figure 69). It was observed that mouse glutamatergic cells are organized in dorso-ventral and medio-lateral axis. Examples are given in Figure 69. Also, it was noted that, not only external glutamatergic cell shows this topographical organization in dorso-ventral and medio-lateral axis, but also inner glutamatergic cells (Figure 69).

Afterwards, differential expression analysis was applied to MERFISH GLUT clusters. Wilcoxon rank-sum test was used in this analysis, and top 10 most expressed genes for each cluster was represented in gene expression heatmap (Figure 69). However, no marker genes were found for dorso-ventral and medio-lateral regions of OB therefore this organization was found to be caused by combinatorial gene expressions.

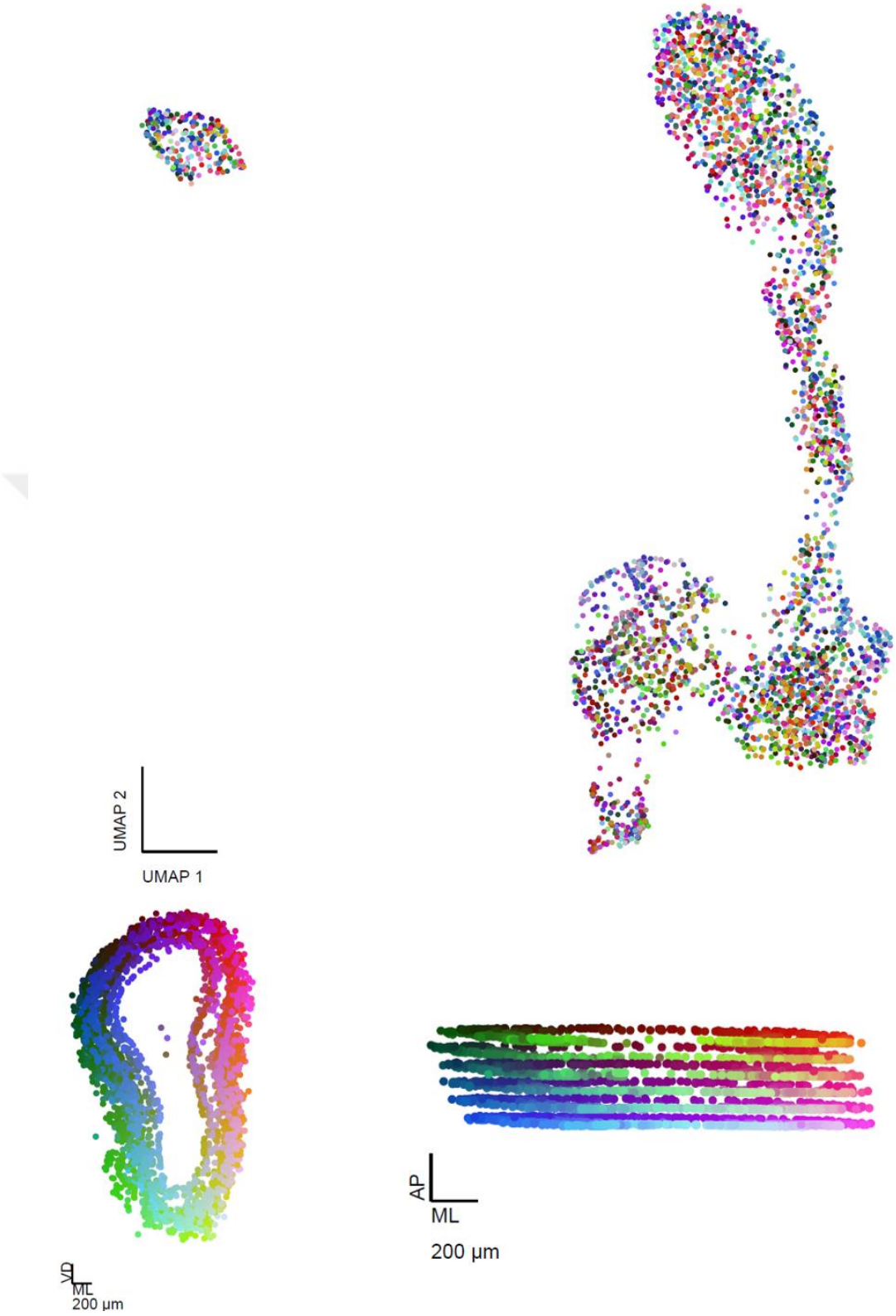

Figure 67. 3D spatial locations of MERFISH GLUT cells in both UMAP, and 2D spatial maps. VD: Ventral Dorsal, ML: Medial Lateral.

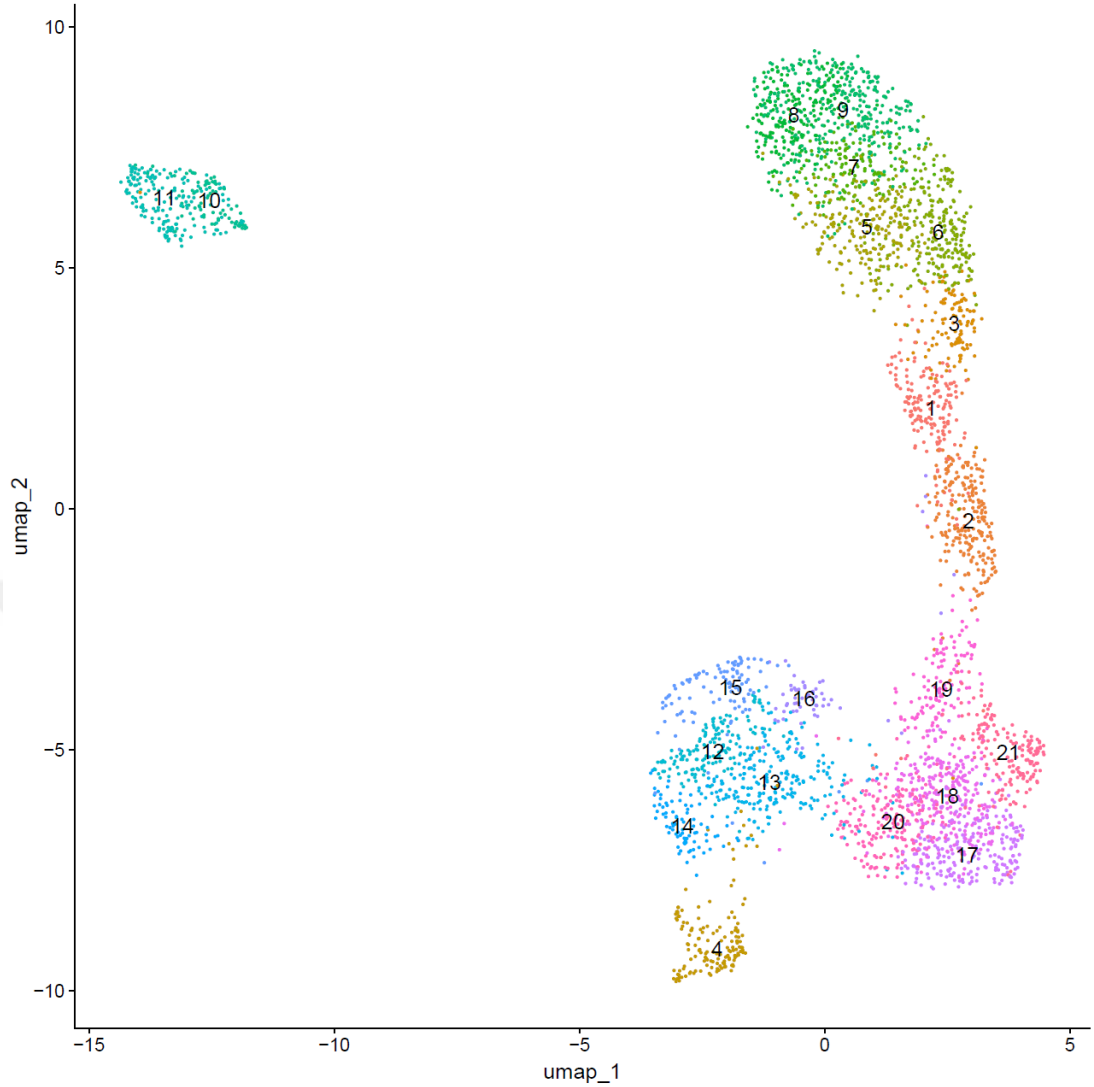

Figure 68. Overclustered MERFISH GLUT clusters in UMAP.

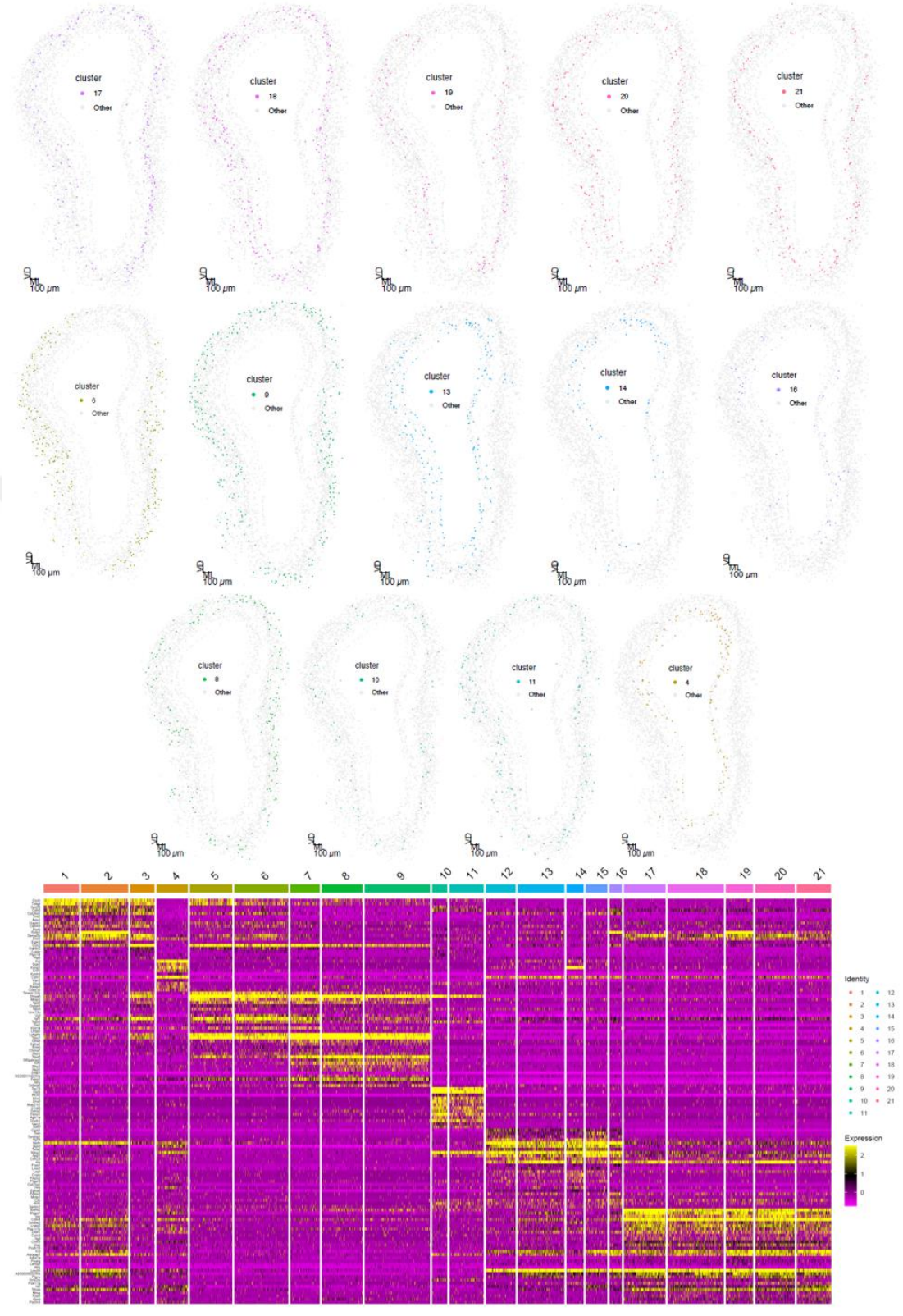

Figure 69. 2D spatial locations of spatially organized MERFISH GLUT clusters and all clusters' gene expression heatmap. VD: Ventral Dorsal, ML: Medial Lateral.

4.2.3 Overcluster of GABAergic cells

4.2.3.1 Zebrafish spatial transcriptomics (AZTEC) dataset

Before overclustering AZTEC GABAergic clusters, pairwise correlations of cells vs distance were plotted for all GABAergic clusters. In this graph the spatial coordinates of the cells were also shuffled and represented in the graph as a control group (Figure 70).

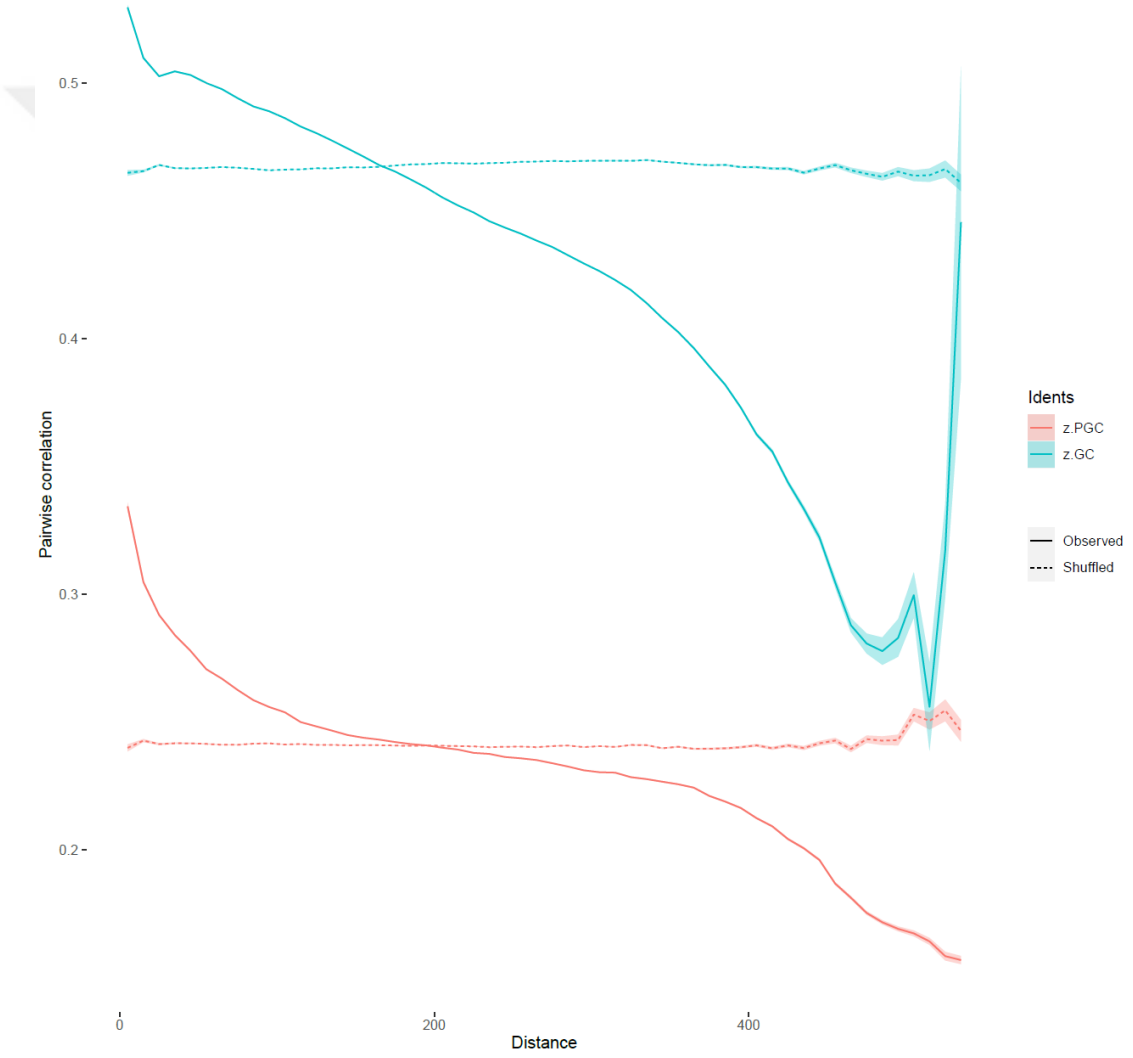

Figure 70. AZTEC GABAergic cells correlation vs distance plot. PGC: Periglomerular Cells, GC: Granule Cells.

Before diving into zebrafish GABAergic overclustering analysis, all cell types were assigned to a color code, and visualized using their spatial coordinates, and UMAP coordinates (Figure 71).

Cells were analyzed using old normalization values but using 3.0 resolution in Louvain community detection method. In total of 27 clusters were identified, UMAP was constructed (Figure 72). Spatially organized clusters were noted and plotted using spatial locations (Figure 73). The clusters were organized based on the “depth” of OB in GABAergic neurons. Though depth is the main axis, there are also dorso-medial, ventro-lateral, and dorso-lateral clusters present.

Afterwards, differential expression analysis was applied to AZTEC GABAergic clusters. Wilcoxon rank-sum test was used in this analysis, clusters were ordered based on genetic similarity top 10 most expressed genes for each cluster was represented in gene expression heatmap (Figure 73). As mentioned before, the organization based on depth was found to be strongly related with *tac1* granule cell type marker gene (Figure 19). However, no marker genes were found for dorso-medial, ventro-lateral, and dorso-lateral regions of OB therefore this organization was found to be caused by combinatorial gene expressions.

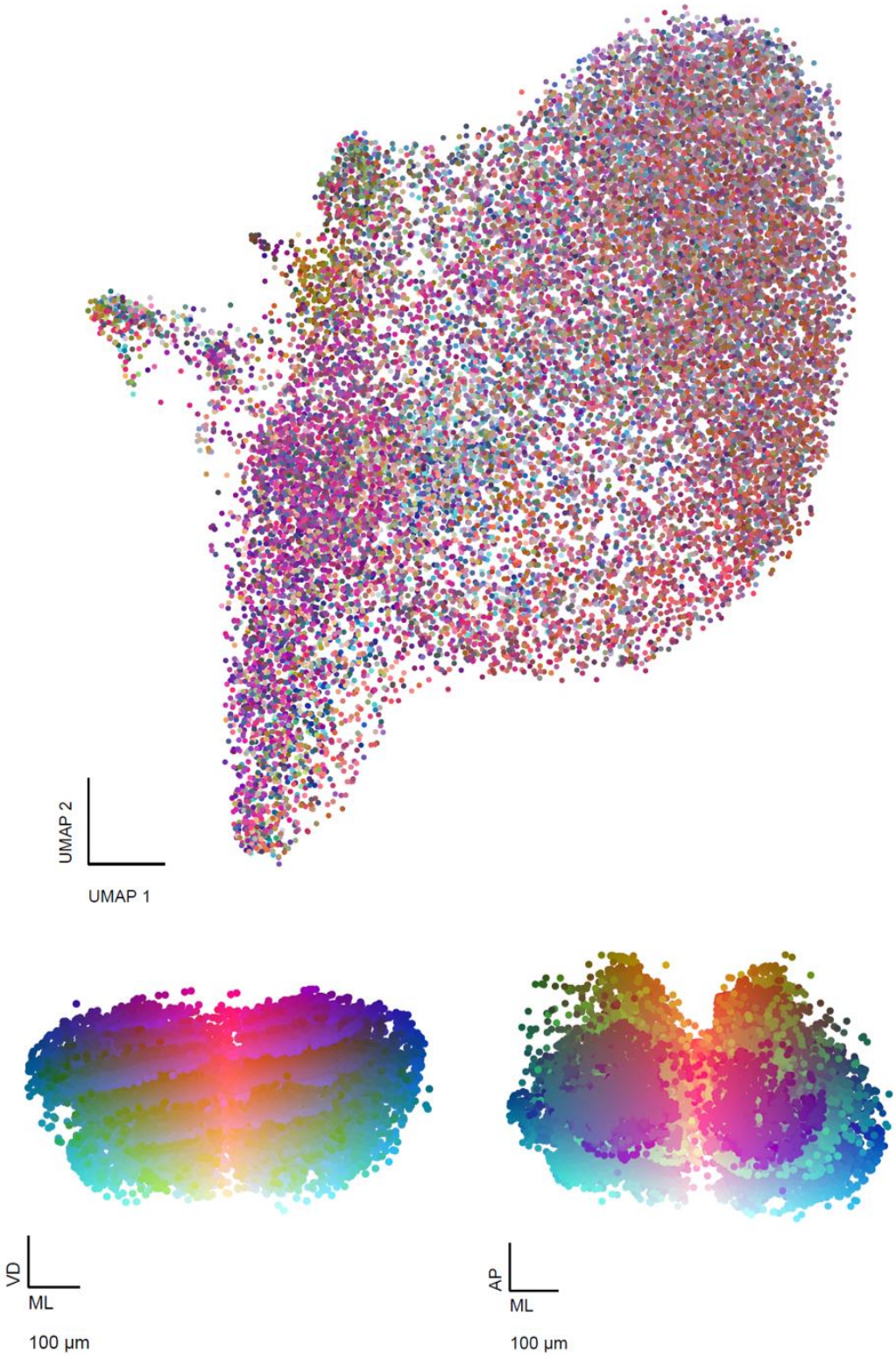

Figure 71. 3D spatial locations of AZTEC GABAergic cells in both UMAP, and 2D spatial maps. VD: Ventral Dorsal, ML: Medial Lateral, AP: Anterior Posterior.

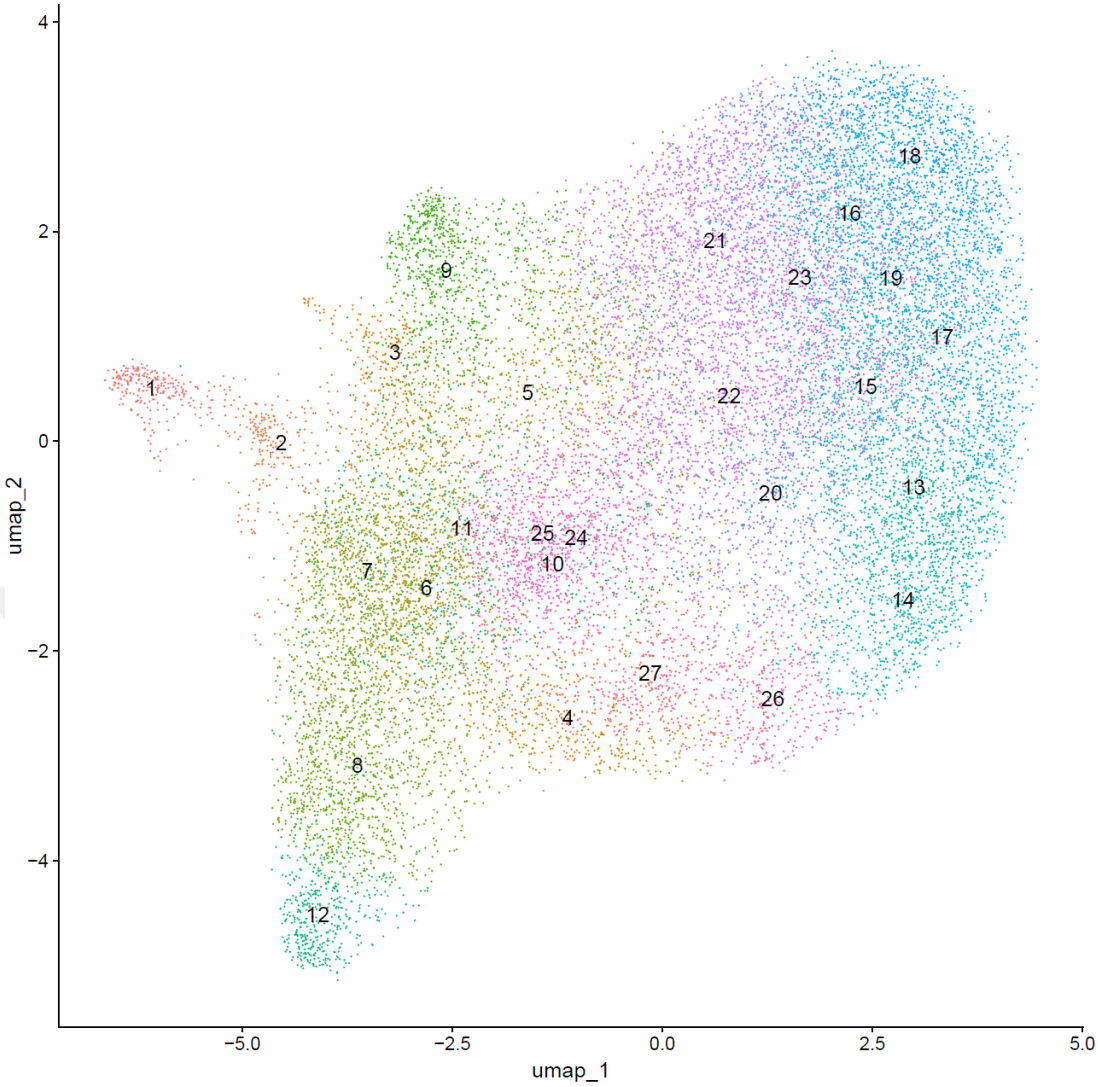

Figure 72. Overclustered AZTEC GABAergic clusters in UMAP.

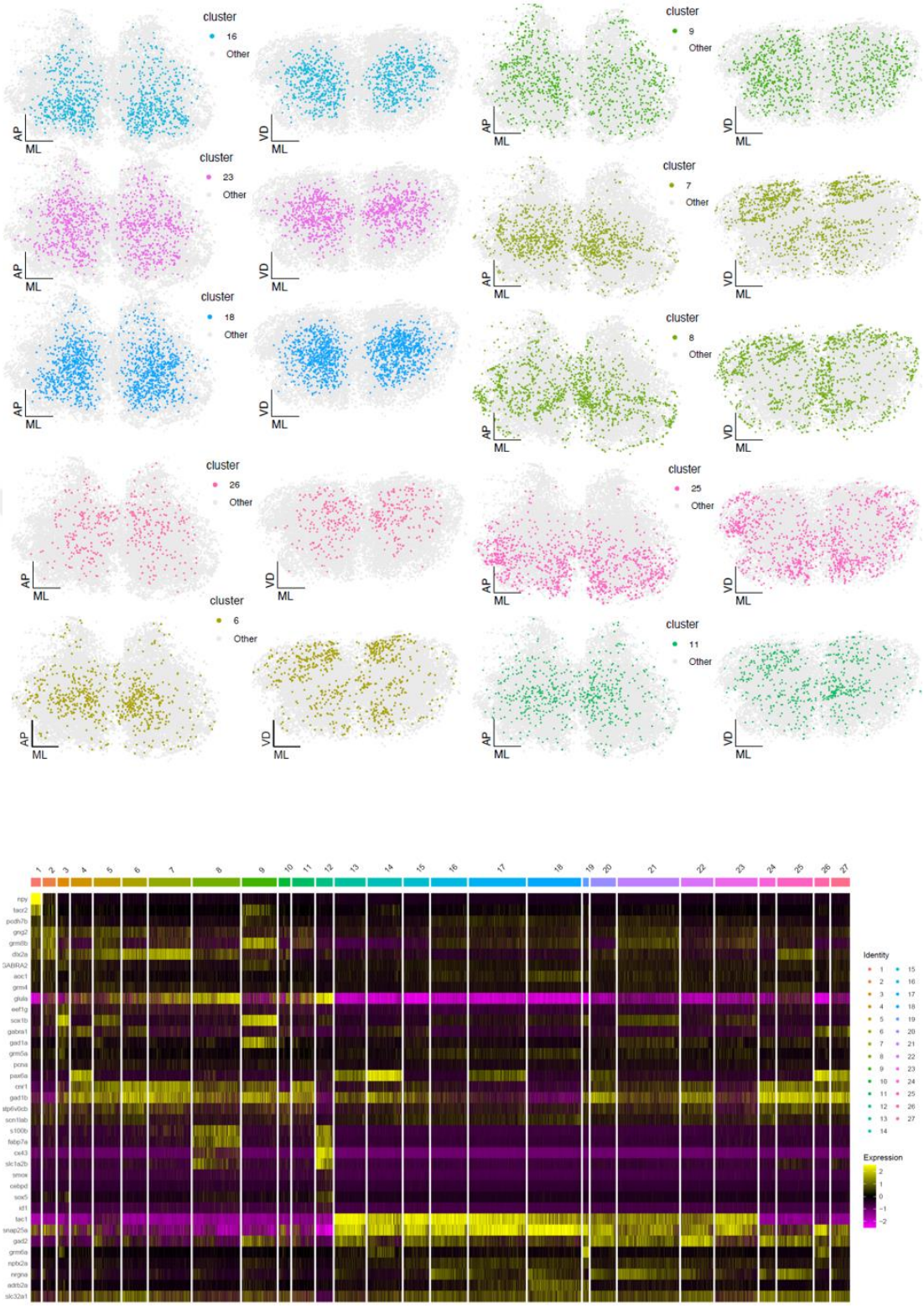

Figure 73. 2D spatial locations of spatially organized AZTEC GABAergic clusters and all clusters' gene expression heatmap. VD: Ventral Dorsal, ML: Medial Lateral, AP: Anterior Posterior.

4.2.3.2 Mouse spatial transcriptomics (MERFISH) dataset

Before overclustering MERFISH GABAergic clusters, pairwise correlations of cells vs distance were plotted for all GABAergic clusters. In this graph the spatial coordinates of the cells were also shuffled and represented in the graph as a control group (Figure 74).

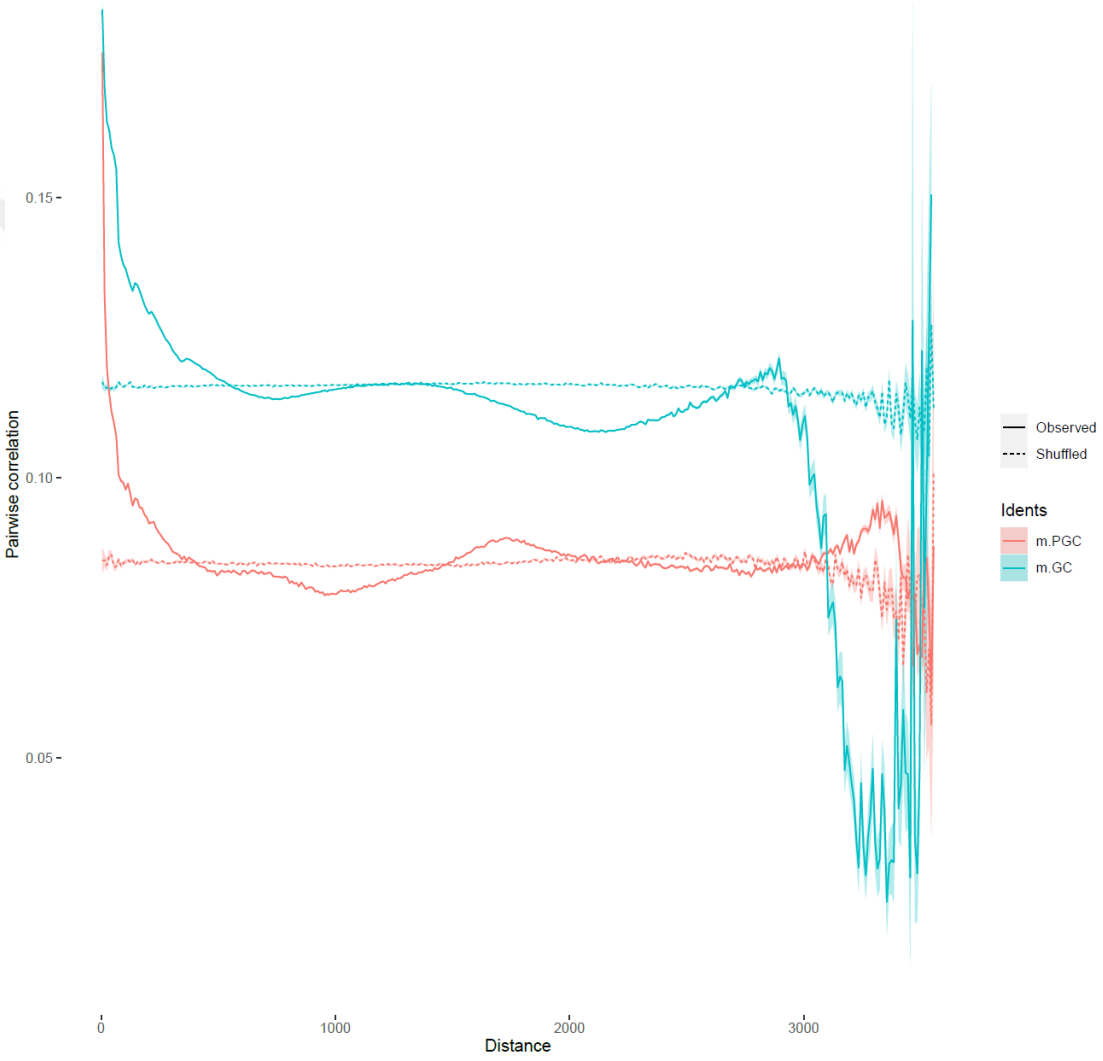

Figure 74. MERFISH GABAergic cells correlation vs distance plot. PGC: Periglomerular Cells, GC: Granule Cells.

Before diving into zebrafish GABAergic overclustering analysis, all cell types were assigned to a color code, and visualized using their spatial coordinates, and UMAP coordinates (Figure 75).

Cells were analyzed using old normalization values but using 3.0 resolution in Louvain community detection method. In total of 44 clusters were identified, clusters were ordered based on genetic similarity, and UMAP was constructed (Figure 76). Spatially organized clusters were noted and plotted using spatial locations (Figure 77). The clusters were organized based on the "depth" of OB in GABAergic neurons. Though depth is the main axis, there are also clusters which are organized at medio-lateral axis (Figure 77).

Afterwards, differential expression analysis was applied to MERFISH GABAergic clusters. Wilcoxon rank-sum test was used in this analysis, and top 10 most expressed genes for each cluster was represented in gene expression heatmap (Figure 77). Granule, and periglomerular cell type marker genes can be counted as relevant marker genes for outer and inner layer topography of GABAergic cells in olfactory bulb. However, *Cbln1* gene was abundantly expressed in GABAergic neurons at medio-lateral regions of OB (Figure 78).

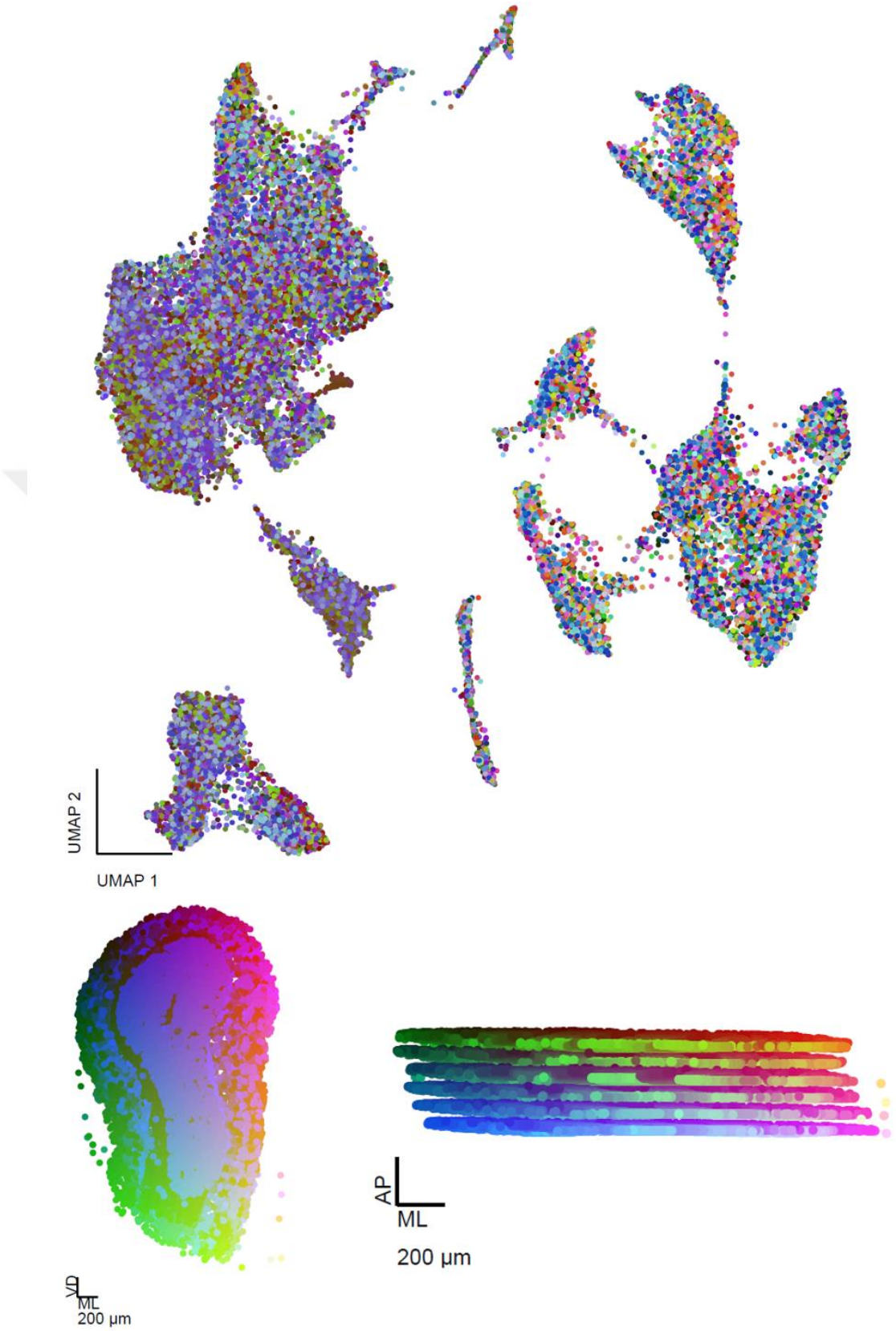

Figure 75. 3D spatial locations of MERFISH GABAergic cells in both UMAP, and 2D spatial maps. VD: Ventral Dorsal, ML: Medial Lateral.

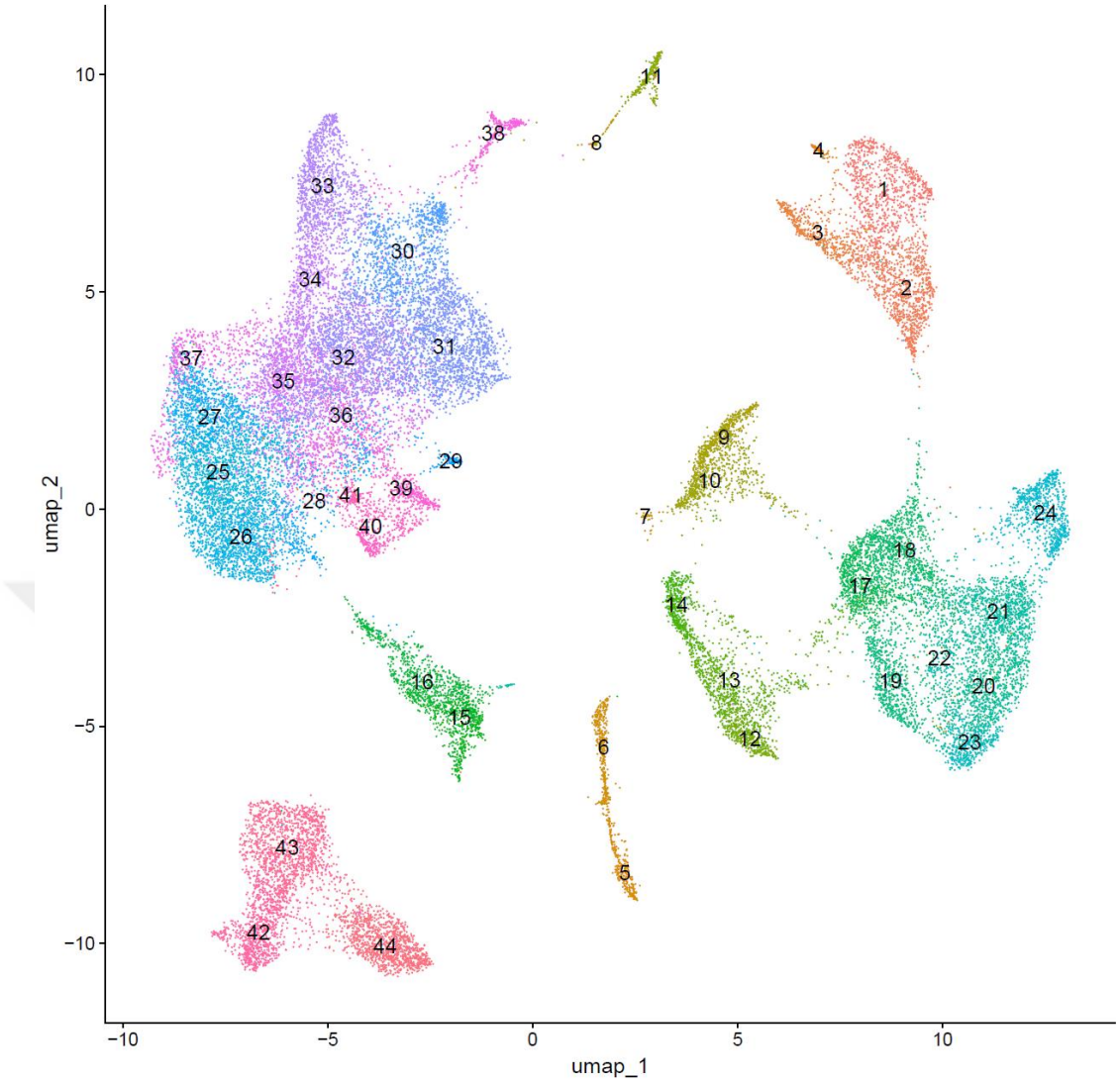

Figure 76. Overclustered MERFISH GABAergic clusters in UMAP.

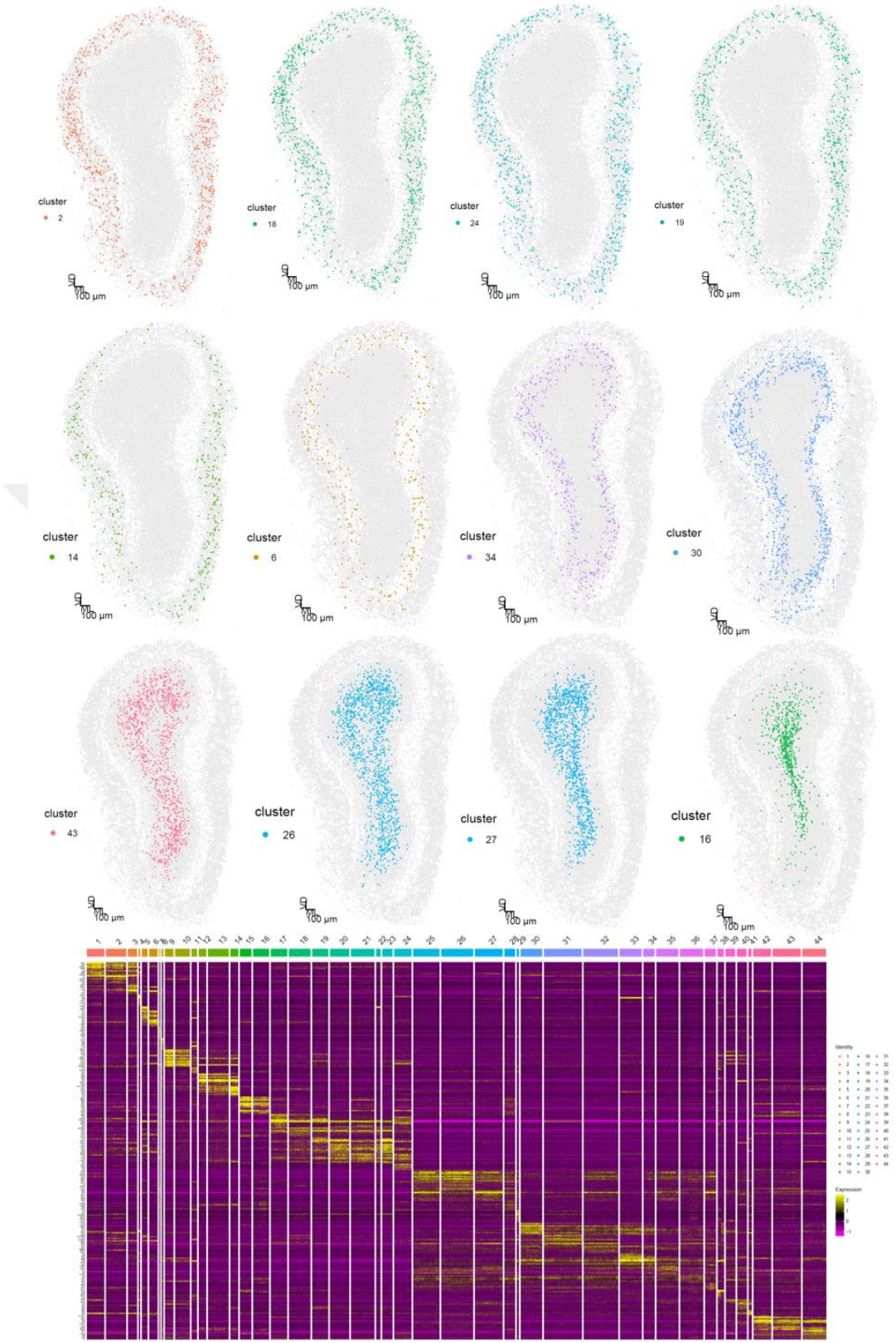

Figure 77. 2D spatial locations of spatially organized MERFISH GABAergic clusters and all clusters' gene expression heatmap. VD: Ventral Dorsal, ML: Medial Lateral.

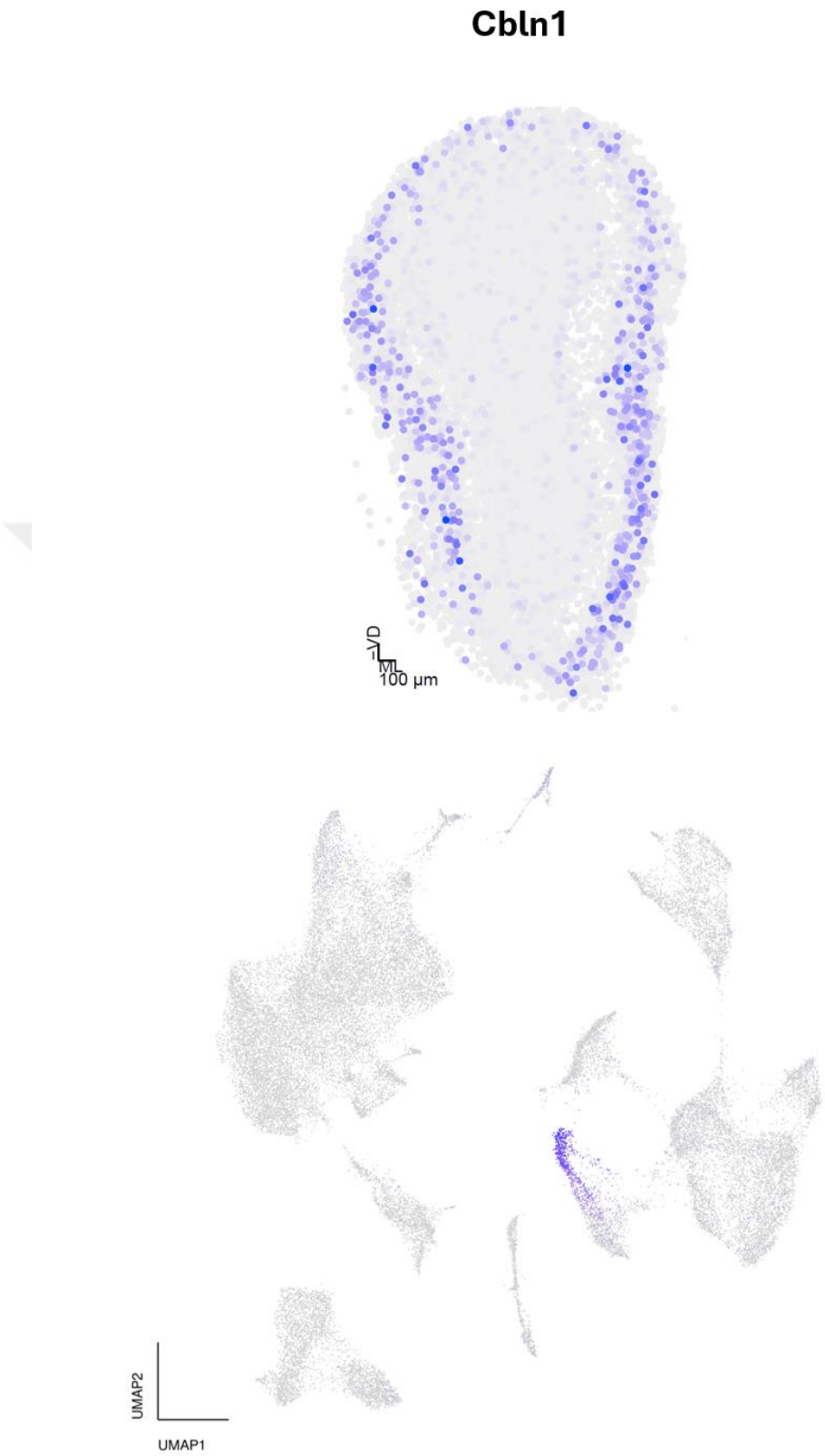

Figure 78. Potential medio-lateral marker gene: *Cbln1* in 2D spatial plot, and UMAP. VD: Ventral Dorsal, ML: Medial Lateral.

5 DISCUSSION

5.1 Comparison of Single Cell and Spatial Transciptomics

In this study, two different types of datasets were used which are single cell transcriptomics, and spatial transcriptomics. Spatial transcriptomics and single-cell transcriptomics are both advanced techniques for studying gene expression, but they differ significantly in methodology and focus. Single-cell transcriptomics dissociates tissues into individual cells for sequencing, providing high-resolution gene expression profiles that highlight cellular heterogeneity but lose spatial context. In contrast, spatial transcriptomics retains the spatial arrangement of tissues, allowing for the analysis of gene expression patterns within their original tissue architecture, although often with lower resolution. Single-cell transcriptomics is ideal for identifying and characterizing different cell types and understanding cell-to-cell variability, while spatial transcriptomics excels in studying the spatial organization of tissues and how gene expression varies across different regions. Each technique has its own challenges and applications, with single-cell methods facing technical difficulties in isolating cells and spatial methods dealing with the complexity of analyzing spatial data. These two methods also have strengths and weaknesses. For example, AZTEC, and MERFISH data had lover gene counts compared to single cell transcriptomics. However, in both datasets, more glutamatergic cells were captured and annotated compared to single cell transcriptomics. Another advantages of AZTEC, and MERFISH datasets is also they have higher cell counts which enabled us to capture more cells for each cell types. On the other hand, scRNA-seq datasets have greater number of genes compared to spatial transcriptomics datasets which makes them as mentioned more reliable for capturing different cell types. In this study, the spatial information coming from AZTEC, and MERFISH datasets, and genetic information coming from scRNA-seq datasets were combined by using KMR matrixes, and alluvial plots (Figure 58). So that the validations of our spatial clusters were successfully made. The framework structured for this study worked efficiently.

5.2 Identifying Optimum Resolution of Determining Cell Types

In order to identify different cell types in both species, cell resolution used in Louvain Community Method plays a crucial role. This resolution number may vary depending on what kind of datasets is used. For example, if the analyzed part of dataset is known to have lots of different cell types, this resolution number might be increased. In this case, since it does not know how many cell types, and subtypes there are in olfactory bulb exactly in both species, the resolution parameter was selected based on the checking the stability of the clusters using clustree. Although clustree tool does not exactly account for optimum resolution, it facilitates the process of selecting it by showing the stability of all clusters. Selecting the resolution number method was using a trial-and-error method. It is also important the different cell types may dominant the datasets in resolution perspective. For example, in our MERFISH dataset, there is a cluster named TC/MC. This clusters are essentially a mixed cluster of tufted and mitral cells however the resolution number is now high enough to divide them therefore they are taken as whole. The reason behind that is olfactory bulb is mostly composed of GABAergic neurons which dominates the dataset. If resolution number had been increased more to subdivide those glutamatergic clusters, GABAergic clusters in this case might have overclustered. Therefore, in order to analyze each cell type with minimum batch effects, GABAergic neurons and glutamatergic neurons were analyzed separately. With that way, the difference between genetically close cell types were identified better.

Another important point is that overclustering resolution was taken as 3.0 for all datasets. This might have been an issue for potential batch affects. This number is usually 5 times of what is generally used as a resolution number. However, since it is possible to generate spatial positions of the clusters, the effect of batch effects was minimized. In other words, the clusters that were not spatially organized or has very few cell numbers been not really considered as significant because their presence might be due to potential batch effects caused by high resolution.

5.3 Diversity of Cell Types in Vertebrate Olfactory Bulb

In this study, molecular features of cell types, and their spatial distribution were investigated in both mouse and zebrafish. To begin with, it was confirmed as it was shown before that glutamatergic cells in mouse olfactory are composed of at least three different cell types they are located in different depths which are mitral cells, tufted cells, and external tufted cells with distinct marker genes (Figure 38, 50-52) (144). In line with this observation, it was also observed that there are at least three different types of glutamatergic cells in zebrafish olfactory bulb which earlier was not shown (Figure 16, 20-21). Except z.MC2, distinct marker genes were found for z.MC1, and z.MC3. Although no marker gene is found for z.MC2, its presence can be identified combinatorically. Since AZTEC data has only 99 genes, marker genes for z.MC2 might not be present in AZTEC gene set.

Concordantly, GABAergic cells also show distinct subtypes. There are 7 genetically different types of periglomerular cell types with distinct marker genes in mouse olfactory bulb (Figure 37, 41-46). Such molecular diversity has not been shown before. In addition, there are 5 different subtypes of granule cell types in mouse olfactory bulb. Granule cell types in mouse olfactory bulb, although genetically different, show a continuum of clustering genetically with the exception of m.GC5. This pattern can also be observed in their spatial organization, and UMAP (Figure 36, 39, 47-49). Moreover, 1 immature GABAergic cell types were identified which is located at the deepest part of mouse olfactory bulb (Figure 36, 48). While the GABAergic cell types are not very well defined in zebrafish based on the limitation in the geneset in spatial transcriptome. Single cell transcriptome dataset of zebrafish also confirmed that there are multiple subtypes of GABAergic cells (Figure 8-9).

5.4 Evolutionary Comparison of Mouse and Zebrafish

In this study, it was also shown that the evolutionary relationship between mouse, and zebrafish using KMR matrixes, and alluvial plots. It was observed that both in KMR matrixes, and alluvial plots, mouse and zebrafish cluster are able to find their corresponding clusters (Figure 56, 57). Although the strength of this correlation differs depending on which cluster is considered, overall, most of these clusters were found to be relatable to each other. After validation the relationship between two species, it was also observed the relationship between spatial transcriptome datasets clusters using single cell transcriptomics as a bridge in between (Figure 58). Since AZTEC dataset is limited by its number of genes, the transition in between mouse and zebrafish was not intensely smooth, but still there are correspondences observed even with 99 genes.

Interestingly it was observed that glutamatergic cell types that are more on the external part of the zebrafish olfactory bulb (z.MC3) corresponds to the external tufted cells (m. ETC1) suggesting some level of evolutionary conservation in the way these cells are distributed (Figure 58). The relationship GABAergic clusters of spatial transcriptomics dataset were not evident due to being limited by AZTEC genes as mentioned previously. Therefore, although evolutionary conservation was observed in between GABAergic neurons in single cell transcriptomics, such information was limited for spatial transcriptomics clusters (Figure 58).

Overall, evolutionary conservation was shown between mouse and zebrafish which was not shown before. Although considerable relationship between external glutamatergic cells of both species was observed, same approach for GABAergic cells were limited due to limited number of genes in zebrafish spatial transcriptomics.

5.5 Spatial Features of Glutamatergic Cells

There are few studies showing that dorsal olfactory bulb is special, and topographically organized. For example, it was previously shown that only dorsomedial region of olfactory bulb is responsive to muscone (145). In another study, the absence of *Ocam* gene, which is crucial for olfaction as it encodes proteins involved in the proper targeting and connectivity of olfactory sensory neurons, ensuring accurate odor detection and processing, is only present at axons that project to dorsal zones of olfactory bulb (18). All these studies suggest that mouse dorsal olfactory bulb is special. Concordantly, the findings of this study support these studies. It was found that some clusters. It can be seen that some mouse glutamatergic cells are organized at medio-lateral axis of olfactory bulb avoiding dorsal axis (Figure 69). Although no marker gene that account for this spatial topography was identified, the combinatorial expression of 1122 gene selected for this data yields such results.

In addition, zebrafish glutamatergic cells show more organized topographical patterns compared to mouse olfactory bulb. Overclustering of the dataset revealed that there are different clusters (potential glomeruli) that are organized at ventro-medial, dorso-medial, and dorso-lateral axes of olfactory bulb (Figure 65). Such clusters also revealed distinct marker genes which are *fezf2, pax6a,* and *tacr3l* for each one of these axes respectively. Finding such genes in limited 99 geneset might be seen misleading due to potential batch effects. However, since the results were supported by spatial information of AZTEC data, it is highly unlikely that such genes, and clusters are organized randomly.

These spatio-molecular topography of glutamatergic neurons might also be related to functional organization of olfactory bulb. For example, in the zebrafish it is shown that appetitive food-derived odors and attractive sex pheromones activate the ventrolateral and ventromedial OB, respectively (146–148) . It is known that mouse of both sexes is strongly attracted to the odor of opposite-sex urine, which activates mitral cells located in the ventral OB (149,150). Also, it is shown that in rats, that floral, woody, fruity and herbaceous odorants, which are rated as pleasant by humans, preferentially activate the rat ventral OB (151). In the light of this, ventro-medial marker gene *fezf2* might be responsible for food odors, and social behaviours. In addition to which, it is known that the dorsal OB plays a specific role in the processing of odorants signaling danger. The OSNs expressing trace amine-associated receptors, which mediate detection of spoiled flesh and/or predator odorants, project onto the dorsal OB in both zebrafish and mice (152–154). In the zebrafish, aversive odorants signaling decaying flesh and fear-inducing alarm odorants mostly activate the dorsal and dorsolateral OB domains, respectively (148,154–156). Therefore, *pax6a*, and *tacr3l* genes are potential genes that might be related to danger signaling.

Overall, both zebrafish and mouse glutamatergic cells show topographical organization. Zebrafish glutamatergic cells show some evident topographical organization whereas mouse glutamatergic cells indicate that dorsal mouse OB is special, and there are cell types that avoid dorsal mouse olfactory bulbs.

5.6 Spatial Features of GABAergic Cells

As mentioned, 5 different types of granule cell types, 7 different types of periglomerular cell types were identified for mouse olfactory bulb (Figure 39). It was found depth is the main axis for GABAergic cell types for spatial topology. Although the main axis is depth, it can be clearly seen that periglomerular cells are organized in two dimensions which are dorso-ventral, and medio-lateral (Figure 37). After overclustering GABAergic cell types, this phenomenon was better understood (Figure 77). In contrast to glutamatergic cell types, a potential marker gene that is expressed at medio-lateral axis has been discovered which is *Cbln1*. Although this gene is mainly expressed in one type of periglomerular cell type, it is an important indicator of periglomerular cells are spatially organized. On the other hand, the depth axis is much more critical for granule cells than it is for periglomerular cells. That is because as mentioned, different granule cells represent a continuum of one another genetically, and spatially also (Figure 77). That is particularly interesting given the fact that the deepest part of the olfactory bulb is mainly composed of immature GABAergic cells (Figure 36). These findings are significant because periglomerular interneurons mediate interglomerular interactions (157), and it is known that olfactory fear learning enhances not only the representation of the learned odorant but also that of the unconditioned odorants in periglomerular interneurons (158) , and output neurons (159). Since danger signaling is related to dorsal olfactory bulb glomeruli, the dorsal organization that is shown in this study might be correlated to such functional topography. In parallel to this, a lot of periglomerular clusters that were observed in overclustering specifically avoided dorsal mouse olfactory bulb indicating that dorsal olfactory bulb is special for GABAergic neurons as well. Also, it is known that the odor pattern is regulated by feedback loops between output neurons and granule cells that are involved in olfactory discrimination and memory (15,19,160,161). Therefore, marker genes for granule cells might have important roles in odor discrimination and memory.

In zebrafish, similar to mouse OB, the main axis for spatial topography is depth. This can be seen in both basic GABAergic cell types (Figure 15), and in overclustered version of GABAergic cell types (Figure 73). Compared to mouse periglomerular cells, zebrafish periglomerular cell types show distinct organizations for example there are clusters that are located at dorsal olfactory bulb, and there are also different clusters which are present at anterio-ventral, and medio-lateral regions of zebrafish olfactory bulb (Figure 73). However, due to being limited by 99 genes in AZTEC data, potential marker genes for these clusters could not be identified. This topographical organization seems to be due to combinatorial gene expression of all 99 genes. Although AZTEC dataset is limited with gene count, it is clear that such topography is not random because it is supported by spatial information. In granule cells however, cells are organized based on depth. These findings are interesting because in glutamatergic cells, we have observed specific clusters of dorsal, dorso-lateral, and ventro-medial regions. Since GABAergic neurons play an important role in regulation of glutamatergic neurons, these spatial patterns might be related to danger signaling, social interactions, and food odors as mentioned previously. Also, *tac1* gene is an important marker gene for granule cell types therefore this gene might be related to fish memory. Since it is previously shown that GABAergic neurons have their own topology in zebrafish (65), the findings of this study support previous works.

Overall, GABAergic neurons are spatially organized in both mouse, and zebrafish. The main axis in both species is depth, and cells are structured based on this axis. In addition to depth, mouse periglomerular cells are organized at dorso-ventral, and medio-lateral axes whereas in zebrafish multiple dorsal, anterio-ventral clusters were identified. Cells that are at different depths should have different functions. However, these functions have not been known yet. One possibility is that cells that are at different depths might reach to higher brain regions whereas cells that are located at outer layer of olfactory bulb might regulate local organizations.

6 CONCLUSION

In this study, we generated and utilized single-cell and spatial transcriptomic atlases of both zebrafish and mouse to compare molecularly defined cell types of the vertebrate olfactory bulb and investigated the molecular correlates of olfactory bulb topography. We first showed that basic cell types of zebrafish and mouse olfactory bulbs can be identified and matched across species. Main glutamatergic and GABAergic neuron types were identified with distinct marker genes. Olfactory bulb neuron types in zebrafish and mice are largely homologous, except for some mouse GABAergic populations. A distinct peripheral population of glutamatergic neurons in the zebrafish olfactory bulb is homologous to external tufted cells in mice.

Next, we demonstrated that in each species, both glutamatergic projection neurons and GABAergic interneurons exhibit a range of molecular diversity represented by a continuum of gene expression profiles rather than distinct molecular clusters. This continuum of molecular diversity in GABAergic and glutamatergic neurons correlates with distinct topographical organization principles, where nearby neurons exhibit a higher degree of molecular similarity. Olfactory bulb neurons exhibit a continuum of gene expression diversity, correlating with olfactory bulb topography.

Both zebrafish and mouse projection neurons in spatially distinct olfactory bulb zones across external plexiform layers exhibit high molecular similarity, clearly different from spatially distant olfactory bulb zones. In both species, these molecularly identifiable groups of projection neurons are organized into non-overlapping regions across the dorsoventral and mediolateral axes of the olfactory bulbs. In zebrafish, we identified that these molecularly distinct groups of projections are organized into distinct olfactory bulb zones encoding alarm odors, social odors, and food odors. Conversely, interneurons exhibit a different logic of molecular topography.

Interneurons display different molecular topography, with diversity primarily along the depth axis in both species. In both species, the major axis of molecular diversity is the depth from the surface of the olfactory bulb, where interneurons at different depths exhibit prominent transcriptomic differences. While there is less zonal organization of interneuron molecular topography along the dorsoventral and mediolateral axes of the mouse olfactory bulb, the zebrafish olfactory bulb shows a prominent separation of molecularly distinct interneuron types along these dimensions. Mouse olfactory bulb interneurons show less zonal organization along the dorsoventral, and mediolateral axes compared to zebrafish.

Nearby neurons show higher transcriptomic similarity, indicating a molecular correlation with spatial arrangement. Glutamatergic neurons in spatially distinct zones of the zebrafish olfactory bulb exhibit high molecular similarity, which is less prominent in mice except in the dorsal olfactory bulb. This study provides a comprehensive view of the spatio-molecular distribution of olfactory bulb neurons in vertebrates and highlights the conserved and divergent molecular features underlying the organization of the olfactory bulb across species.

7 REFERENCES

1. Mountcastle V. The columnar organization of the neocortex. Brain. 1997 Apr 1;120(4):701–22.
2. Ferrier lecture - Functional architecture of macaque monkey visual cortex. Proc R Soc Lond B Biol Sci. 1977 Jul 28;198(1130):1–59.
3. Knudsen EI, Konishi M. A Neural Map of Auditory Space in the Owl. Science (1979). 1978 May 19;200(4343):795–7.
4. PENFIELD W, BOLDREY E. SOMATIC MOTOR AND SENSORY REPRESENTATION IN THE CEREBRAL CORTEX OF MAN AS STUDIED BY ELECTRICAL STIMULATION. Brain. 1937;60(4):389–443.
5. TRAVERS JB, DINARDO LA, KARIMNAMAZI H. Motor and Premotor Mechanisms of Licking. Neurosci Biobehav Rev. 1997;21(5):631–47.
6. Fanselow MS, Dong HW. Are the Dorsal and Ventral Hippocampus Functionally Distinct Structures? Neuron. 2010 Jan;65(1):7–19.
7. Strange BA, Witter MP, Lein ES, Moser EI. Functional organization of the hippocampal longitudinal axis. Nat Rev Neurosci. 2014 Oct 19;15(10):655–69.
8. Rakic P, Ayoub AE, Breunig JJ, Dominguez MH. Decision by division: making cortical maps. Trends Neurosci. 2009 May;32(5):291–301.
9. Marín O, Rubenstein JLR. Cell migration in the forebrain. Annu Rev Neurosci. 2003 Mar;26(1):441–83.
10. Tessier-Lavigne M, Goodman CS. The Molecular Biology of Axon Guidance. Science (1979). 1996 Nov 15;274(5290):1123–33.
11. O'Leary DDM, Koester SE. Development of projection neuron types, axon pathways, and patterned connections of the mammalian cortex. Neuron. 1993 Jun;10(6):991–1006.
12. Mombaerts P, Wang F, Dulac C, Chao SK, Nemes A, Mendelsohn M, et al. Visualizing an Olfactory Sensory Map. Cell. 1996 Nov;87(4):675–86.
13. Ressler KJ, Sullivan SL, Buck LB. Information coding in the olfactory system: Evidence for a stereotyped and highly organized epitope map in the olfactory bulb. Cell. 1994 Dec;79(7):1245 55.
14. Vassar R, Ngai J, Axel R. Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. Cell. 1993 Jul;74(2):309–18.
15. Mori K, Yoshihara Y. Molecular recognition and olfactory processing in the mammalian olfactory system. Prog Neurobiol. 1995 Apr;45(6):585–619.
16. Sosulski DL, Bloom ML, Cutforth T, Axel R, Datta SR. Distinct representations of olfactory information in different cortical centres. Nature. 2011 Apr 30;472(7342):213–6.
17. Miyasaka N, Arganda-Carreras I, Wakisaka N, Masuda M, Sümbül U, Seung HS, et al. Olfactory projectome in the zebrafish forebrain revealed by genetic single-neuron labelling. Nat Commun. 2014 Apr 9;5(1):3639.
18. Mori K, Nagao H, Yoshihara Y. The Olfactory Bulb: Coding and Processing of Odor Molecule

Information. Science (1979). 1999 Oct 22;286(5440):711–5.
19. Mori K, Sakano H. How Is the Olfactory Map Formed and Interpreted in the Mammalian Brain? Annu Rev Neurosci. 2011 Jul 21;34(1):467–99.
20. Mombaerts P. Axonal Wiring in the Mouse Olfactory System. Annu Rev Cell Dev Biol. 2006 Nov 1;22(1):713–37.
21. Wachowiak M, Shipley MT. Coding and synaptic processing of sensory information in the glomerular layer of the olfactory bulb. Semin Cell Dev Biol. 2006 Aug;17(4):411–23.
22. Galizia CG, Menzel R. The role of glomeruli in the neural representation of odours: results from optical recording studies. J Insect Physiol. 2001 Feb;47(2):115–30.
23. Mori K, Mataga N, Imamura K. Differential specificities of single mitral cells in rabbit olfactory bulb for a homologous series of fatty acid odor molecules. J Neurophysiol. 1992 Mar 1;67(3):786–9.
24. Katoh K, Koshimoto H, Tani A, Mori K. Coding of odor molecules by mitral/tufted cells in rabbit olfactory bulb. II. Aromatic compounds. J Neurophysiol. 1993 Nov 1;70(5):2161–75.
25. Imamura K, Mataga N, Mori K. Coding of odor molecules by mitral/tufted cells in rabbit olfactory bulb. I. Aliphatic compounds. J Neurophysiol. 1992 Dec 1;68(6):1986–2002.
26. Salcedo E, Zhang C, Kronberg E, Restrepo D. Analysis of Training-Induced Changes in Ethyl Acetate Odor Maps Using a New Computational Tool to Map the Glomerular Layer of the Olfactory Bulb. Chem Senses. 2005 Sep 1;30(7):615–26.
27. Montag-Sallaz M, Buonviso N. Altered odor-induced expression of c- *fos* and arg 3.1 immediate early genes in the olfactory system after familiarization with an odor. J Neurobiol. 2002 Jul 17;52(1):61–72.
28. Inaki K, Takahashi YK, Nagayama S, Mori K. Molecular-feature domains with posterodorsal–anteroventral polarity in the symmetrical sensory maps of the mouse olfactory bulb: mapping of odourant-induced Zif268 expression. European Journal of Neuroscience. 2002 May 10;15(10):1563–74.
29. Schaefer ML, Yamazaki K, Osada K, Restrepo D, Beauchamp GK. Olfactory Fingerprints for Major Histocompatibility Complex-Determined Body Odors II: Relationship among Odor Maps, Genetics, Odor Composition, and Behavior. The Journal of Neuroscience. 2002 Nov 1;22(21):9513–21.
30. Schaefer ML, Young DA, Restrepo D. Olfactory Fingerprints for Major Histocompatibility Complex-Determined Body Odors. The Journal of Neuroscience. 2001 Apr 1;21(7):2481–7.
31. Sallaz M, Jourdan F. C-fos expression and 2-deoxyglucose uptake in the olfactory bulb of odour-stimulated awake rats. Neuroreport. 1993 Jan;4(1):55–8.
32. Guthrie K, Rayhanabad J, Kuhl D, Gall C. Odors regulate Arc expression in neuronal ensembles engaged in odor processing. Neuroreport. 2000 Jun;11(9):1809–13.
33. Guthrie KM, Anderson AJ, Leon M, Gall C. Odor-induced increases in c-fos mRNA expression reveal an anatomical “unit” for odor processing in olfactory bulb. Proceedings of the National Academy of Sciences. 1993 Apr 15;90(8):3329–33.

34. Onoda N. Odor-induced fos-like immunoreactivity in the rat olfactory bulb. Neurosci Lett. 1992 Mar;137(2):157–60.
35. Igarashi KM, Mori K. Spatial Representation of Hydrocarbon Odorants in the Ventrolateral Zones of the Rat Olfactory Bulb. J Neurophysiol. 2005 Feb;93(2):1007–19.
36. Takahashi YK, Kurosaki M, Hirono S, Mori K. Topographic Representation of Odorant Molecular Features in the Rat Olfactory Bulb. J Neurophysiol. 2004 Oct;92(4):2413–27.
37. Meister M, Bonhoeffer T. Tuning and Topography in an Odor Map on the Rat Olfactory Bulb. The Journal of Neuroscience. 2001 Feb 15;21(4):1351–60.
38. Uchida N, Takahashi YK, Tanifuji M, Mori K. Odor maps in the mammalian olfactory bulb: domain organization and odorant structural features. Nat Neurosci. 2000 Oct;3(10):1035–43.
39. Rubin BD, Katz LC. Optical Imaging of Odorant Representations in the Mammalian Olfactory Bulb. Neuron. 1999 Jul;23(3):499–511.
40. Spors H, Wachowiak M, Cohen LB, Friedrich RW. Temporal Dynamics and Latency Patterns of Receptor Neuron Input to the Olfactory Bulb. The Journal of Neuroscience. 2006 Jan 25;26(4):1247–59.
41. Spors H, Grinvald A. Spatio-Temporal Dynamics of Odor Representations in the Mammalian Olfactory Bulb. Neuron. 2002 Apr;34(2):301–15.
42. Fried HU, Fuss SH, Korsching SI. Selective imaging of presynaptic activity in the mouse olfactory bulb shows concentration and structure dependence of odor responses in identified glomeruli. Proceedings of the National Academy of Sciences. 2002 Mar 5;99(5):3222–7.
43. Wachowiak M, Cohen LB. Correspondence Between Odorant-Evoked Patterns of Receptor Neuron Input and Intrinsic Optical Signals in the Mouse Olfactory Bulb. J Neurophysiol. 2003 Mar 1;89(3):1623–39.
44. Wachowiak M, Cohen LB. Representation of Odorants by Receptor Neuron Input to the Mouse Olfactory Bulb. Neuron. 2001 Nov;32(4):723–35.
45. Fuss SH, Korsching SI. Odorant Feature Detection: Activity Mapping of Structure Response Relationships in the Zebrafish Olfactory Bulb. The Journal of Neuroscience. 2001 Nov 1;21(21):8396–407.
46. Friedrich RW, Korsching SI. Chemotopic, Combinatorial, and Noncombinatorial Odorant Representations in the Olfactory Bulb Revealed Using a Voltage-Sensitive Axon Tracer. The Journal of Neuroscience. 1998 Dec 1;18(23):9977–88.
47. Friedrich RW, Korsching SI. Combinatorial and Chemotopic Odorant Coding in the Zebrafish Olfactory Bulb Visualized by Optical Imaging. Neuron. 1997 May;18(5):737–52.
48. Sachse S, Rappert A, Galizia CG. The spatial representation of chemical structures in the antennal lobe of honeybees: steps towards the olfactory code. European Journal of Neuroscience. 1999 Nov 27;11(11):3970–82.
49. Galizia CG, Sachse S, Rappert A, Menzel R. The glomerular code for odor representation is species specific in the honeybee Apis mellifera. Nat Neurosci. 1999 May;2(5):473–8.
50. Joerges J, Küttner A, Galizia CG, Menzel R. Representations of odours and odour mixtures

visualized in the honeybee brain. Nature. 1997 May;387(6630):285–8.
51. Lei H, Christensen TA, Hildebrand JG. Spatial and Temporal Organization of Ensemble Representations for Different Odor Classes in the Moth Antennal Lobe. The Journal of Neuroscience. 2004 Dec 8;24(49):11108–19.
52. Skiri HT. Representation of Primary Plant Odorants in the Antennal Lobe of the Moth Heliothis virescens Using Calcium Imaging. Chem Senses. 2004 Mar 1;29(3):253–67.
53. Collmann C, Carlsson MA, Hansson BS, Nighorn A. Odorant-Evoked Nitric Oxide Signals in the Antennal Lobe of *Manduca sexta*. The Journal of Neuroscience. 2004 Jul 7;24(27):6070–7.
54. Carlsson MA. Spatial Representation of Odours in the Antennal Lobe of the Moth Spodoptera littoralis (Lepidoptera: Noctuidae). Chem Senses. 2002 Mar 1;27(3):231–44.
55. Giovanni Galizia C, Sachse S, Mustaparta H. Calcium responses to pheromones and plant odours in the antennal lobe of the male and female moth Heliothis virescens. J Comp Physiol A. 2000 Dec 12;186(11):1049–63.
56. Galizia CG, Menzel R, Hölldobler B. Optical Imaging of Odor-Evoked Glomerular Activity Patterns in the Antennal Lobes of the Ant Camponotus rufipes. Naturwissenschaften. 1999 Nov 3;86(11):533–7.
57. Kreher SA, Kwon JY, Carlson JR. The Molecular Basis of Odor Coding in the Drosophila Larva. Neuron. 2005 May;46(3):445–56.
58. Wang JW, Wong AM, Flores J, Vosshall LB, Axel R. Two-Photon Calcium Imaging Reveals an Odor-Evoked Map of Activity in the Fly Brain. Cell. 2003 Jan;112(2):271–82.
59. Rodrigues V. Spatial coding of olfactory information in the antennal lobe of Drosophila melanogaster. Brain Res. 1988 Jun;453(1–2):299–307.
60. Vosshall LB, Stocker RF. Molecular Architecture of Smell and Taste in *Drosophila*. Annu Rev Neurosci. 2007 Jul 1;30(1):505–33.
61. Su CY, Menuz K, Carlson JR. Olfactory Perception: Receptors, Cells, and Circuits. Cell. 2009 Oct;139(1):45–59.
62. Li Q, Liberles SD. Aversion and Attraction through Olfaction. Current Biology. 2015 Feb;25(3):R120–9.
63. Galizia CG, Rössler W. Parallel Olfactory Systems in Insects: Anatomy and Function. Annu Rev Entomol. 2010 Jan 1;55(1):399–420.
64. Kay LM, Sherman SM. An argument for an olfactory thalamus. Trends Neurosci. 2007 Feb;30(2):47–53.
65. Yaksi E, Judkewitz B, Friedrich RW. Topological Reorganization of Odor Representations in the Olfactory Bulb. PLoS Biol. 2007 Jul 3;5(7):e178.
66. Firestein S. How the olfactory system makes sense of scents. Nature. 2001 Sep;413(6852):211–8.
67. Buck L, Axel R. A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. Cell. 1991 Apr;65(1):175–87.
68. Wyatt TD. Pheromones and Animal Behavior. Cambridge University Press; 2014.
69. SHEPHERD GM, CHEN WR, GREER CA. Olfactory Bulb. In: The Synaptic Organization of the

Brain. Oxford University Press; 2004. p. 165–216.
70. Wilson DA, Sullivan RM. Cortical Processing of Odor Objects. Neuron. 2011 Nov;72(4):506–19.
71. Nagayama S, Homma R, Imamura F. Neuronal organization of olfactory bulb circuits. Front Neural Circuits. 2014 Sep 3;8.
72. Kosaka T. Structural Organization of the Glomerulus in the Main Olfactory Bulb. Chem Senses. 2005 Jan 1;30(Supplement 1):i107–8.
73. Cleland TA, Linster C. Central olfactory structures. In 2019. p. 79–96.
74. Shipley MT, Ennis M. Functional organization of olfactory system. J Neurobiol. 1996 May;30(1):123–76.
75. Wilson RI, Mainen ZF. EARLY EVENTS IN OLFACTORY PROCESSING. Annu Rev Neurosci 2006 Jul 21;29(1):163–201.
76. Brennan PA, Zufall F. Pheromonal communication in vertebrates. Nature. 2006 Nov 15;444(7117):308–15.
77. Animals in a chemical world. In: Pheromones and Animal Behaviour. Cambridge University Press; 2003. p. 1–22.
78. Buck LB. The Molecular Architecture of Odor and Pheromone Sensing in Mammals. Cell. 2000 Mar;100(6):611–8.
79. Mombaerts P. Genes and ligands for odorant, vomeronasal and taste receptors. Nat Rev Neurosci. 2004 Apr;5(4):263–78.
80. Rodriguez I, Greer CA, Mok MY, Mombaerts P. A putative pheromone receptor gene expressed in human olfactory mucosa. Nat Genet. 2000 Sep;26(1):18–9.
81. Capecchi MR. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet. 2005 Jun;6(6):507–12.
82. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. Developmental Dynamics. 1995 Jul 3;203(3):253–310.
83. Baier H, Scott EK. Genetic and optical targeting of neural circuits and behavior—zebrafish in the spotlight. Curr Opin Neurobiol. 2009 Oct;19(5):553–60.
84. Stainier DYR, Fishman MC. The zebrafish as a model system to study cardiovascular development. Trends Cardiovasc Med. 1994 Sep;4(5):207–12.
85. Merrick C, Godwin CA, Geisler MW, Morsella E. The olfactory system as the gateway to the neural correlates of consciousness. Front Psychol. 2014;4.
86. Satou M. Synaptic organization, local neuronal circuitry, and functional segregation of the teleost olfactory bulb. Prog Neurobiol. 1990 Jan;34(2):115–42.
87. Kermen F, Franco LM, Wyatt C, Yaksi E. Neural circuits mediating olfactory-driven behavior in fish. Front Neural Circuits. 2013;7.
88. Dryer L, Graziadei PPC. Mitral cell dendrites: a comparative approach. Anat Embryol (Berl). 1994 Feb;189(2).
89. Shepherd GM, Chen WR, Willhite D, Migliore M, Greer CA. The olfactory granule cell: From classical enigma to central role in olfactory processing. Brain Res Rev. 2007 Oct;55(2):373–82.

90. Emre Yaksi. Analysis of odor-evoked activity pattern in the zebrafish brain by two-photon calcium imaging. Max Planck Institute, University of Heidelberg; 2003.
91. Shepherd GM. Discrimination of molecular signals by the olfactory receptor neuron. Neuron. 1994 Oct;13(4):771–90.
92. Johnson BA, Farahbod H, Leon M. Interactions between odorant functional group and hydrocarbon structure influence activity in glomerular response modules in the rat olfactory bulb. J Comp Neurol. 2005 Mar 7;483(2):205–16.
93. Joerges J, Küttner A, Galizia CG, Menzel R. Representations of odours and odour mixtures visualized in the honeybee brain. Nature. 1997 May;387(6630):285–8.
94. Duchamp-Viret P, Coronas V, Delaleu JC, Moyse E, Duchamp A. Dopaminergic modulation of mitral cell activity in the frog olfactory bulb: a combined radioligand binding–electrophysiological study. Neuroscience. 1997 Apr;79(1):203–16.
95. Cinelli AR, Hamilton KA, Kauer JS. Salamander olfactory bulb neuronal activity observed by video rate, voltage-sensitive dye imaging. III. Spatial and temporal properties of responses evoked by odorant stimulation. J Neurophysiol. 1995 May 1;73(5):2053–71.
96. Guthrie KM, Anderson AJ, Leon M, Gall C. Odor-induced increases in c-fos mRNA expression reveal an anatomical "unit" for odor processing in olfactory bulb. Proceedings of the National Academy of Sciences. 1993 Apr 15;90(8):3329–33.
97. Kauer JS. Contributions of topography and parallel processing to odor coding in the vertebrate olfactory pathway. Trends Neurosci. 1991 Feb;14(2):79–85.
98. Kauer JS, White J. Imaging and Coding in the Olfactory System. Annu Rev Neurosci. 2001 Mar;24(1):963–79.
99. Stewart WB, Kauer JS, Shepherd GM. Functional organization of rat olfactory bulb analysed by the 2-deoxyglucose method. Journal of Comparative Neurology. 1979 Jun 15;185(4):715–34.
100. Moulton DG. Spatial patterning of response to odors in the peripheral olfactory system. Physiol Rev. 1976 Jul 1;56(3):578–93.
101. Uchida N, Takahashi YK, Tanifuji M, Mori K. Odor maps in the mammalian olfactory bulb: domain organization and odorant structural features. Nat Neurosci. 2000 Oct;3(10):1035–43.
102. Friedrich RW, Laurent G. Dynamic Optimization of Odor Representations by Slow Temporal Patterning of Mitral Cell Activity. Science (1979). 2001 Feb 2;291(5505):889–94.
103. Laurent G, Davidowitz H. Encoding of Olfactory Information with Oscillating Neural Assemblies. Science (1979). 1994 Sep 23;265(5180):1872–5.
104. Toida K. Synaptic organization of the olfactory bulb based on chemical coding of neurons. Anat Sci Int. 2008 Dec;83(4):207–17.
105. Wellis DP, Scott JW, Harrison TA. Discrimination among odorants by single neurons of the rat olfactory bulb. J Neurophysiol. 1989 Jun 1;61(6):1161–77.
106. Harrison TA, Scott JW. Olfactory bulb responses to odor stimulation: analysis of response pattern and intensity relationships. J Neurophysiol. 1986 Dec 1;56(6):1571–89.
107. Macrides F, Chorover SL. Olfactory Bulb Units: Activity Correlated with Inhalation Cycles and

Odor Quality. Science (1979). 1972 Jan 7;175(4017):84–7.
108. Wehr M, Laurent G. Odour encoding by temporal sequences of firing in oscillating neural assemblies. Nature. 1996 Nov;384(6605):162–6.
109. Laurent G, Stopfer M, Friedrich RW, Rabinovich MI, Volkovskii A, Abarbanel H DI. Odor Encoding as an Active, Dynamical Process: Experiments, Computation, and Theory. Annu Rev Neurosci. 2001 Mar;24(1):263–97.
110. Murthy VN. Olfactory Maps in the Brain. Annu Rev Neurosci. 2011 Jul 21;34(1):233–58.
111. Ressler KJ, Sullivan SL, Buck LB. A zonal organization of odorant receptor gene expression in the olfactory epithelium. Cell. 1993 May;73(3):597–609.
112. Leon M, Johnson BA. Olfactory coding in the mammalian olfactory bulb. Brain Res Rev. 2003 Apr;42(1):23–32.
113. Rodriguez I. Singular Expression of Olfactory Receptor Genes. Cell. 2013 Oct;155(2):274–7.
114. Snyder SH, Sklar PB, Hwang PM, Pevsner J. Molecular mechanisms of olfaction. Trends Neurosci. 1989 Jan;12(1):35–8.
115. Hamdani EH, Døving KB. The functional organization of the fish olfactory system. Prog Neurobiol. 2007 Jun;82(2):80–6.
116. Shi Q, Chen X, Zhang Z. Decoding Human Biology and Disease Using Single-Cell Omics Technologies. Genomics Proteomics Bioinformatics. 2023 Oct 1;21(5):926–49.
117. Saliba AE, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res. 2014 Aug 18;42(14):8845–60.
118. Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y. Single-cell RNA sequencing technologies and applications: A brief overview. Clin Transl Med. 2022 Mar 29;12(3).
119. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, et al. Comparative Analysis of Single-Cell RNA Sequencing Methods. Mol Cell. 2017 Feb;65(4):631-643.e4.
120. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, et al. Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. Cell. 2015 May;161(5):1187–201.
121. Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA. The Technology and Biology of Single-Cell RNA Sequencing. Mol Cell. 2015 May;58(4):610–20.
122. Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD. Reverse Transcriptase Template Switching: A SMART™ Approach for Full-Length cDNA Library Construction. Biotechniques. 2001 Apr 5;30(4):892–7.
123. Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. Nat Methods. 2014 Feb 22;11(2):163–6.
124. Adil A, Kumar V, Jan AT, Asger M. Single-Cell Transcriptomics: Current Methods and Challenges in Data Acquisition and Analysis. Front Neurosci. 2021 Apr 22;15.
125. Trapnell C. Defining cell types and states with single-cell genomics. Genome Res. 2015 Oct 1;25(10):1491–8.
126. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods. 2009 May 6;6(5):377–82.

127. Jung N, Kim TK. Spatial transcriptomics in neuroscience. Exp Mol Med. 2023 Oct 2;55(10):2105–15.
128. Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science (1979). 2016 Jul;353(6294):78–82.
129. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2021.
130. Bredesen-Aa BA, HF, . . . , YE. Single-cell volumetric transcriptomic atlas of the adult zebrafish forebrain reveals teleost homologues of cortical and subcortical structures from terrestrial vertebrates. manuscript in preparation, unpublished. 2024;
131. Zappia L, Oshlack A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. Gigascience. 2018 Jul 1;7(7).
132. D'Gama PP, Qiu T, Cosacak MI, Rayamajhi D, Konac A, Hansen JN, et al. Diversity and function of motile ciliated cell types within ependymal lineages of the zebrafish brain. Cell Rep. 2021 Oct;37(1):109775.
133. Kenney JW, Steadman PE, Young O, Shi MT, Polanco M, Dubaishi S, et al. A 3D adult zebrafish brain atlas (AZBA) for the digital age. Elife. 2021 Nov 22;10.
134. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods. 2021 Jan 14;18(1):100–6.
135. Yao Z, van Velthoven CTJ, Kunst M, Zhang M, McMillen D, Lee C, et al. A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. Nature. 2023 Dec 14;624(7991):317–32.
136. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol. 2024 Feb 25;42(2):293–304.
137. Zhang M, Pan X, Jung W, Halpern AR, Eichhorn SW, Lei Z, et al. Molecularly defined and spatially resolved cell atlas of the whole mouse brain. Nature. 2023 Dec 14;624(7991):343–54.
138. Choudhary S, Satija R. Comparison and evaluation of statistical error models for scRNA-seq. Genome Biol. 2022 Jan 18;23(1):27.
139. Greenacre M, Groenen PJF, Hastie T, D'Enza AI, Markos A, Tuzhilina E. Publisher Correction: Principal component analysis. Nature Reviews Methods Primers. 2023 Mar 8;3(1):22.
140. Lenz M, Müller FJ, Zenke M, Schuppert A. Principal components analysis and the reported low intrinsic dimensionality of gene expression microarray data. Sci Rep. 2016 Jun 2;6(1):25696.
141. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol. 2019 Jan 3;37(1):38–44.
142. Li Y, Ge X, Peng F, Li W, Li JJ. Exaggerated false positives by popular differential expression methods when analyzing human population samples. Genome Biol. 2022 Dec 15;23(1):79.
143. Muller J, Szklarczyk D, Julien P, Letunic I, Roth A, Kuhn M, et al. eggNOG v2.0: extending the evolutionary genealogy of genes with enhanced non-supervised orthologous groups, species and

functional annotations. Nucleic Acids Res. 2010 Jan;38(suppl_1):D190–5.
144. Zeppilli S, Ackels T, Attey R, Klimpert N, Ritola KD, Boeing S, et al. Molecular characterization of projection neuron subtypes in the mouse olfactory bulb. Elife. 2021 Jul 22;10.
145. Shirasu M, Yoshikawa K, Takai Y, Nakashima A, Takeuchi H, Sakano H, et al. Olfactory Receptor and Neural Pathway Responsible for Highly Selective Sensing of Musk Odors. Neuron. 2014 Jan;81(1):165–78.
146. Yabuki Y, Koide T, Miyasaka N, Wakisaka N, Masuda M, Ohkura M, et al. Olfactory receptor for prostaglandin F2α mediates male fish courtship behavior. Nat Neurosci. 2016 Jul 30;19(7):897–904.
147. Kermen F, Lal P, Faturos NG, Yaksi E. Interhemispheric connections between olfactory bulbs improve odor detection. PLoS Biol. 2020 Apr 20;18(4):e3000701.
148. Calvo-Ochoa E, Byrd-Jacobs CA. The Olfactory System of Zebrafish as a Model for the Study of Neurotoxicity and Injury: Implications for Neuroplasticity and Disease. Int J Mol Sci. 2019 Apr 2;20(7):1639.
149. Xu F, Schaefer M, Kida I, Schafer J, Liu N, Rothman DL, et al. Simultaneous activation of mouse main and accessory olfactory bulbs by odors or pheromones. Journal of Comparative Neurology. 2005 Sep 5;489(4):491–500.
150. Kang N, Baum MJ, Cherry JA. A direct main olfactory bulb projection to the 'vomeronasal' amygdala in female mice selectively responds to volatile pheromones from males. European Journal of Neuroscience. 2009 Feb 2;29(3):624–34.
151. Auffarth B. Continuous spatial representations in the olfactory bulb may reflect perceptual categories. Front Syst Neurosci. 2011;5.
152. Pacifico R, Dewan A, Cawley D, Guo C, Bozza T. An Olfactory Subsystem that Mediates High-Sensitivity Detection of Volatile Amines. Cell Rep. 2012 Jul;2(1):76–88.
153. Hussain A, Saraiva LR, Ferrero DM, Ahuja G, Krishna VS, Liberles SD, et al. High-affinity olfactory receptor for the death-associated odor cadaverine. Proceedings of the National Academy of Sciences. 2013 Nov 26;110(48):19579–84.
154. Dieris M, Ahuja G, Krishna V, Korsching SI. A single identified glomerulus in the zebrafish olfactory bulb carries the high-affinity response to death-associated odor cadaverine. Sci Rep. 2017 Jan 19;7(1):40892.
155. Mathuru AS, Kibat C, Cheong WF, Shui G, Wenk MR, Friedrich RW, et al. Chondroitin Fragments Are Odorants that Trigger Fear Behavior in Fish. Current Biology. 2012 Mar;22(6):538–44.
156. Kermen F, Darnet L, Wiest C, Palumbo F, Bechert J, Uslu O, et al. Stimulus-specific behavioral responses of zebrafish to a large range of odors exhibit individual variability. BMC Biol. 2020 Dec 15;18(1):66.
157. Kermen F, Mandairon N, Chalençon L. Odor hedonics coding in the vertebrate olfactory bulb. Cell Tissue Res. 2021 Jan 30;383(1):485–93.
158. Kass MD, McGann JP. Persistent, generalized hypersensitivity of olfactory bulb interneurons after

olfactory fear generalization. Neurobiol Learn Mem. 2017 Dec;146:47–57.

159. Ross JM, Fletcher ML. Learning-Dependent and -Independent Enhancement of Mitral/Tufted Cell Glomerular Odor Responses Following Olfactory Fear Conditioning in Awake Mice. The Journal of Neuroscience. 2018 May 16;38(20):4623–40.

160. Tan J, Savigner A, Ma M, Luo M. Odor Information Processing by the Olfactory Bulb Analyzed in Gene-Targeted Mice. Neuron. 2010 Mar;65(6):912–26.

161. Grelat A, Benoit L, Wagner S, Moigneu C, Lledo PM, Alonso M. Adult-born neurons boost odor–reward association. Proceedings of the National Academy of Sciences. 2018 Mar 6;115(10):2514–9.

8 CURRICULUM VITAE