Genetic programming of the visual forebrain in the absence of retinal input

Submitted by
Shachar Sherman

Graduate School of
Systemic Neurosciences

LMU Munich

Dissertation der
Graduate School of Systemic Neurosciences der
Ludwig-Maximilians-Universität München
June 29th 2022

First reviewer and Supervisor
Prof. Dr. Herwig Baier
Department of Genes-Circuits-Behavior
Max Planck Institute of Neuroscience

Second Reviewer: Prof. Dr. Wolfgang Enard
Third Reviewer: Dr. Christian Mayer
External Reviewer Prof. Filippo Del Bene

Date of Submission: June 29th 2022
Date of Defense : February 8th 2023

ABSTRACT

The brain is assembled during development by both innate and experience-dependent mechanisms[1–7], but the relative contribution of these factors is poorly understood. Axons of retinal ganglion cells (RGCs) connect the eye to the brain, forming a bottleneck for the transmission of visual information to central visual areas. RGCs secrete molecules from their axons that control proliferation, differentiation and migration of downstream components[7–9]. Spontaneously generated waves of retinal activity, but also intense visual stimulation, can entrain responses of RGCs[10] and central neurons[11–16]. Here I asked how the cellular composition of central targets is altered in a vertebrate brain that is depleted of retinal input throughout development. For this, I first established a molecular catalog[17] and gene expression atlas[18] of neuronal subpopulations in the retinorecipient areas of larval zebrafish. I then searched for changes in *lakritz* (*atoh7*-) mutants, in which RGCs do not form[19]. Although individual forebrain-expressed genes are dysregulated in *lakritz* mutants, the complete set of 77 putative neuronal cell types in thalamus, pretectum and tectum are present. While neurogenesis and differentiation trajectories are overall unaltered, a greater proportion of cells remain in an uncommitted progenitor stage in the mutant. Optogenetic stimulation of a pretectal area[20,21] evokes a visual behavior in blind mutants indistinguishable from wildtype. My analysis shows that, in this vertebrate visual system, neurons are produced more slowly, but specified and wired up in a proper configuration in the absence of any retinal signals.

ACKNOWLEDGMENTS

To friends; to colleagues; to mentors.

For my family

LIST OF FIGURES

ABBREVIATIONS

AC	amacrine cell
AF	arborization field
ANTs	advanced normalization tools
BC	bipolar cells
ChR	channelrhodopsin
CMZ	ciliary marginal zone
CRISPR	clustered regularly interspaced short palindromic repeats
DGE	digital gene expression
dLGN	dorsolateral geniculate nucleus
EM	electron microscopy
FACS	fluorescent activated cell sorting
FISH	fluorescent in-situ hybridization
GABA	γ-aminobutyric acid
GCL	ganglion cell layer
GFP	green fluorescent protein
GH	growth hormone
HC	horizontal cells
HCR	hybridization chain reaction
HVG	highly variable gene
IGF-I	insulin-like growth factor-I
INL	inner nuclear layer
mRNA	messenger RNA
NpHR	halorhodopsin
OMR	optomotor reflex
ONL	outer nuclear layer
OKR	optokinetic reflex

PCA	principal component analysis
PPd	dorsal periventricular pretectal
PPv	ventral periventricular pretectal
PTU	phenylthiourea
RFP	red fluorescent protein
RGC	retinal ganglion cell
Tg	transgene
UAS	upstream activating sequence
UMAP	Uniform Manifold Approximation and Projection
UMI	unique molecular identifier
VBA	visual background adaptation
WT	wild type

CONTENTS

INTRODUCTION

Uncovering neuronal cell types to understand brain function and development

While our brain is only one of many organs all responsible for our body's proper function, it is by far the most complex one. Looking into the structure and action of different organ systems, it is easy to understand how their function enables an animal's survival: the digestive tract allows the processing of food and extraction of nutrients. The lungs exchange gasses with the environment, and together with the circulatory system allows the flow of dissolved gasses into the body. The heart pumps blood. The skeletal system gives the body support, and muscles allow it to move. Nothing in the brain is that simple. The structure of the mammalian cortex does not inform of its function. Dissecting the cerebellum does not in any way tell us that it is involved in motor control. Or, that the hippocampus is critical for memory. Thus, a central goal of neuroscience remains to understand how the brain functions: how does the brain perceive the world, encodes experience, and transforms both into action (behavior).

The brain's complexity stems not so much from its large number of cells, but rather from its incredible diversity of cells and how they are connected. And, while there are indeed many types of brain cells, none is more important than the neuron, which both receives and transmits electrochemical signals in the brain - transforming perception into action (behavior). It was the work by Santiago Ramón y Cajal that initially cataloged the incredible diversity of neuronal cell types[22–24]. By sparsely labeling only a handful of neurons using golgi stains in individual brain sections, he was able to correctly conclude that neurons exist as discrete units while still being able to communicate with other neurons (Figure 1). Using brain slices across animals and across brain areas he was able to show that neurons appear in different shapes and sizes. Critically, he also observed that neurons in different animals, in the same brain areas, have similar morphologies. Essentially, he showed for the first time that different neuronal types exist, but also that similar areas harbor the same types.

Later advances in electrophysiology allowed to characterize the function of individual neurons[25,26]. Advances in immunohistochemistry and molecular biology allowed to characterize neurons according to the proteins and mRNA they express[17,27,28]. More recently, advances in electron microscopy (EM) and machine learning have allowed to comprehensively characterize the connectivity of neuronal cell types[29]. Different areas in the brain house neurons showing different combinations of morphology, function, molecules, and synaptic partners. While there is still much debate on which of these properties is most important for describing a neuronal type, there is a general agreement that in order to understand brain function, we must describe and catalog its neuronal units.

The brain also changes dramatically during early development and during the life of an organism[30–32]. Non-neuronal progenitor cells divide to give rise to new neurons that can migrate to various areas in the brain. It is then that they often mature into more specific neuronal types, synapsing with cellular partners and developing specific functions within

the larger neuronal network. Even after embryonic development, cells in the brain continue to undergo changes. The experiences the organism undergoes can dramatically alter the cells themselves, their morphology, connections, and function. Over the lifetime of the organism, young animals mature into adults and develop new behaviors. Whether these are new neurons that accommodate these changes or existing neurons that have changed, the types of neurons composing the brain surely changes throughout an animal's life. Understanding how changes in the neurons themselves over the lifetime of an organism contribute to a change in perception, brain state, and behavior poses an enduring challenge for developmental- and systems-neuroscientists alike.

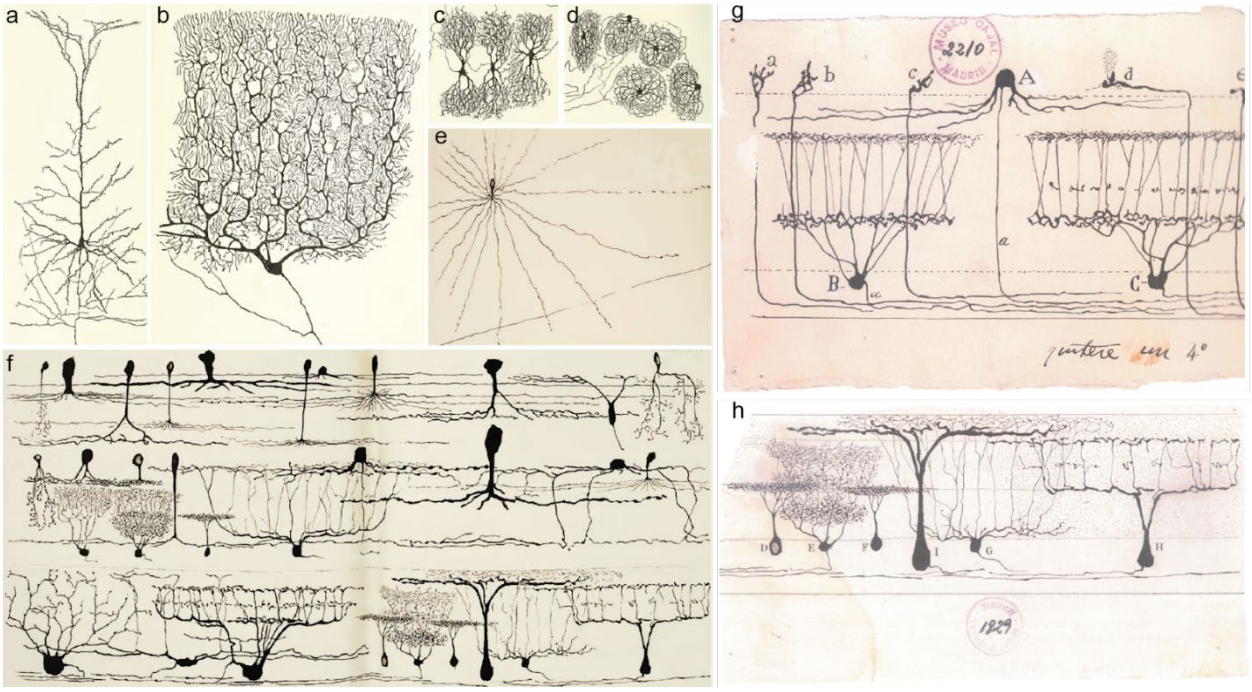


Figure 1: Diversity of neuronal types and their conservation across species

Drawing from Ramón y Cajal depicting neuronal cell type diversity and conservation across species. (a–e) Gallery of drawings by Ramón y Cajal, depicting the characteristic dendritic arbors of different cell types throughout the brain. Arbors vary in their geometry and branching patterns (adapted with permission from Lefebvre JL et al. Annu Rev Cell Dev Biol. 2015). (f,g) Drawings by Ramón y Cajal showing conservation of cell types across retinas of different species (adapted with permission from The Beautiful Brain). (a) Cortical pyramidal cell (rabbit) with apical and basal dendrites and axon. (b) Cerebellar Purkinje cell (human). (c) Neurons of the dorsal column nuclei (human). (d) Neurons of the inferior olive (human). (e) Retinal amacrine cell (green lizard). (f) Ramón y Cajal illustration of cell types found in lizard retina. Drawings depict two cell classes: the amacrine cells (top row) and retinal ganglion cells (bottom row; middle row contains some of each), as seen in transverse sections. Even within a single cell class, the variation in dendrite size and shape is remarkable. Ramón y Cajal used drawings such as these to enumerate and classify the cell types that populate various nervous tissues. Panels a–d from Ramón y Cajal (1909); panels e and f from Ramón y Cajal (1893). (g) Section from a lizard retina. (h) Section from a sparrow retina.

What is a neuronal cell type?

How can we define a cell type in a way that distinguishes it from others but places it together with similar cells? There is general agreement on what distinguishes large cell classes such as neurons, immune cells, hepatocytes, myocytes etc. The consensus is that cells of the same types should occupy the same niche and be more similar to each other than other types. However, as we increase in resolution, attempting to distinguish more closely related cell types, the consensus disintegrates[25,26].

Defining neuronal cell types should be multimodal, spanning multiple categories such as: anatomical location, developmental origin, molecular signature, morphology, connectivity, physiology, and functionality. These different modalities are not always easy to distinguish, some overlap, and they can influence one another. The developmental origin can influence the molecular signature[30,33]. The molecules in a neuron influence its morphology and connectivity[34,35]. Functional properties are determined by synaptic partners, morphology, and molecular components, but can in turn reciprocally affect a neuron's molecular profile[36]. It is the synergy of properties that define the identity of individual neurons.

Discussion on how to prioritize neuronal properties are ongoing. As the activity of the brain is its main signature - contributing directly to behavioral output and learning, some argue that the functional properties of a neuron are the most important for distinguishing neuronal types. On the other hand, some argue that the genetic composition of the cell (expressed mRNA and translated proteins) is most important for defining types. It is easy to see how the latter offers a stronger approach: molecules can be easily classified and distinguished. Behavior and function, on the other hand, are fuzzy. Ideally, all of the different properties can be incorporated into a complete catalog of neuronal cell types, giving a clearer view of what distinguishes one neuronal type from another.

There are currently multiple attempts at generating cell type atlases for entire organisms[37–40]. At the forefront lies the human cell atlas. However, even more ambitious projects have been suggested. Like the old taxonomist cataloging bones of different animals, there is an ongoing endeavor to build a cell type tree of life – one that will eventually grow to contain all cell types in a wide range of animals. It is clear how comparing the genetic composition of cell types is more constructive to building a cell type tree of life over other cellular properties. While it is not always possible to compare either the morphology or function of different cells, there is no apparent limit to comparing nucleotide sequences across different cells (Figure 2). Hopefully, understanding the evolutionary origin of different cell types will prove useful in generating an unbiased catalog of cell types[41].

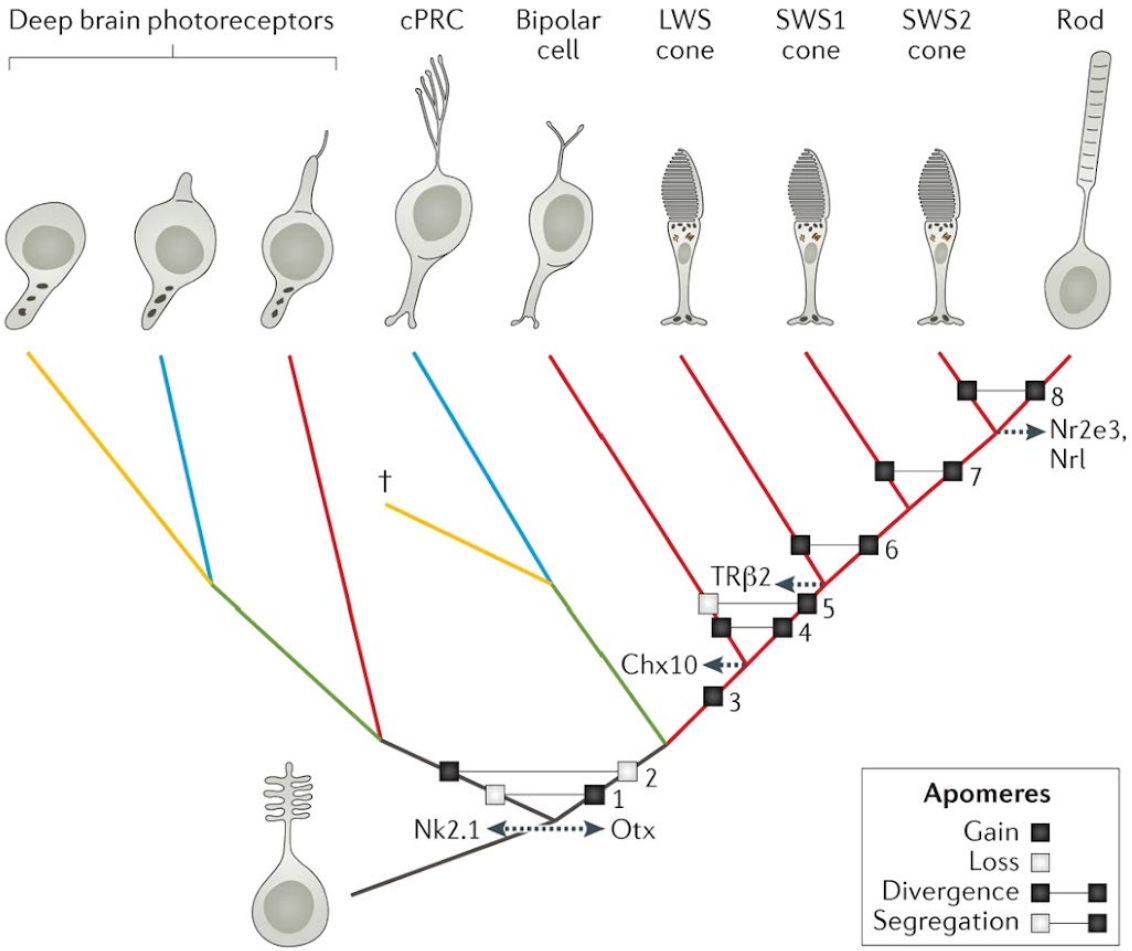


Figure 2: Understanding neuronal diversification by evolutionary comparison of cell types

Evolution of ciliary photoreceptors (cPRCs) in Bilateria (adapted with permission from Arendt, D., Musser, J., Baker, C. et al. Nat Rev Genet. 2016). Different colours indicate phylogenetic cell type splits, with red representing vertebrates, green representing protostomes, yellow representing ecdysozoans, and blue representing lophotrochozoans. Dashed arrows indicate known transcription factor changes 5,8. Different apomeres are indicated by numbers: 1, surface-extended cilium; 2, control of reproduction via neurosecretion; 3, visual function; 4, axonal projection or interneuron function; 5, light sensitivity via c-opsin; 6, deployment of the vertebrate long wavelength-sensitive (LWS) or of the vertebrate short wavelength-sensitive (SWS)1–2–rhodopsin duplicate; 7, deployment of the SWS1 or of the SWS27–rhodopsin duplicate; 8, deployment of the SWS2 or of the rhodopsin duplicate. Cross denotes cell type loss. Chx10, ceh10 homeobox-containing homologue; Nr2e3, nuclear receptor subfamily 2 group E member 3; Nrl, neural retina-specific leucine zipper protein; TRβ2, thyroid hormone receptor β2.

Comprehensive cell type classification by massively parallel transcriptional profiling of single cells

In addition to the challenge of conceptually deciding how to define a cell type, technical challenges are also abundant. While cell types have been described based on connectivity, physiology, or morphology, the number of cells that can be cataloged this way is very limited[42]. The state of transcriptional profiling of single-cells is slightly better[43,44]. Using multi-well plates, hundreds of single-neurons can be collected in wells per experiment and sequenced individually. This allows the identification of cellular transcriptional profiles and the elucidation of transcriptional signatures that make cell types. By extension, this methodology allows to identify which proteins cells express, how they communicate with other cells in the brain, how they recognize their neighbors, and how they sense their environment. However, because of the large number of cells in the brain, even if hundreds of cells are captured in a single experiment, this may not suffice to catalog all cell types in any brain area. Even worse is that while the cost of sequencing each cell is not extreme, the cost of extending this methodology to resolve most cell types makes it prohibitively expensive.

Recent advances have revolutionized the field of single-cell transcriptomics. A number of advances in microfluidics have made it possible to process tens of thousands of cells from a single experiment[45]. As cataloging cell types is a numbers game, this has truly ushered a new era for single-cell genomics. The commercial system sold by 10X Genomics has been widely adopted by scientists to transcriptionally profile single-cells. The 10X Genomics system requires a single-cell suspension as starting material, which is often produced by dissociating tissue of interest into single-cells (Figure 3a,b). The suspension is mixed with a set of reagents required to reverse transcribe mRNA and mixed passed through a single capillary in a microfluidic device. Additional capillaries allow the separate flow of oil and hydrogel beads containing bead-specific barcodes (Figure 3b,c). The microfluidic chamber and controlled flow ensure that oil droplets will form continuously - encapsulating liquid and, occasionally, beads. While millions of such droplets form, in some rare cases, the oil will encapsulate both a hydrogel bead and a cell. The low ratio of events where cells are also encapsulated by oil ensures that when this rare event happens, the droplet will contain only a single-cell.

While the droplets are relatively stable, the cells themselves can be easily dissolved in a droplet and their mRNA reverse-transcribed onto poly-T bead barcodes. Each hydrogel bead contains only a single bead-specific barcode - identifying a single cell (Figure 3c). Additionally, each barcode contains another sequence which is unique to each copy of a barcode. When the mRNA of the cell is synthesized onto the bead-specific barcode, it also acquires a sequence unique to that mRNA molecule – a unique molecular identifier (UMI). It is via these sets of barcodes and identifiers that we can later computationally distinguish which molecule belonged to which cells. The bead barcode becomes the unique cell barcode. And, the UMIs allow us to count how many mRNA copies were present in each cell (Figure 3c).

Droplet-based single-cell genomics has proven extremely successful at resolving populations of neuronal cell types. Data sets containing hundreds of thousands, and

even millions of cells, are regularly published (Figure 3d). Their biological scope is also huge: ranging from cataloging cells in model[46–48] and non-model organisms[49], to understanding disease progression[50], aging[32], and development[47,51]. Methods are also continuously developing to allow processing of different materials. Nuclei and fixed cells, more stable than live whole cells, can now be used. The computational tools that analyze these data have also expanded dramatically. It is now possible to not only use these data to identify cell types, but also to infer differentiation pathways, identify evolutionarily-related cell types, and to integrate single-cell data with techniques such as spatial transcriptomics, and epitope (membrane proteins) sequencing.

Taken together, massively parallel droplet-based transcriptional profiling provides researchers with an extremely powerful tool to comprehensively characterize neuronal cell types, identify how and where they change during development including into old-age, and to uncover genetic candidates to investigate how the brain functions.

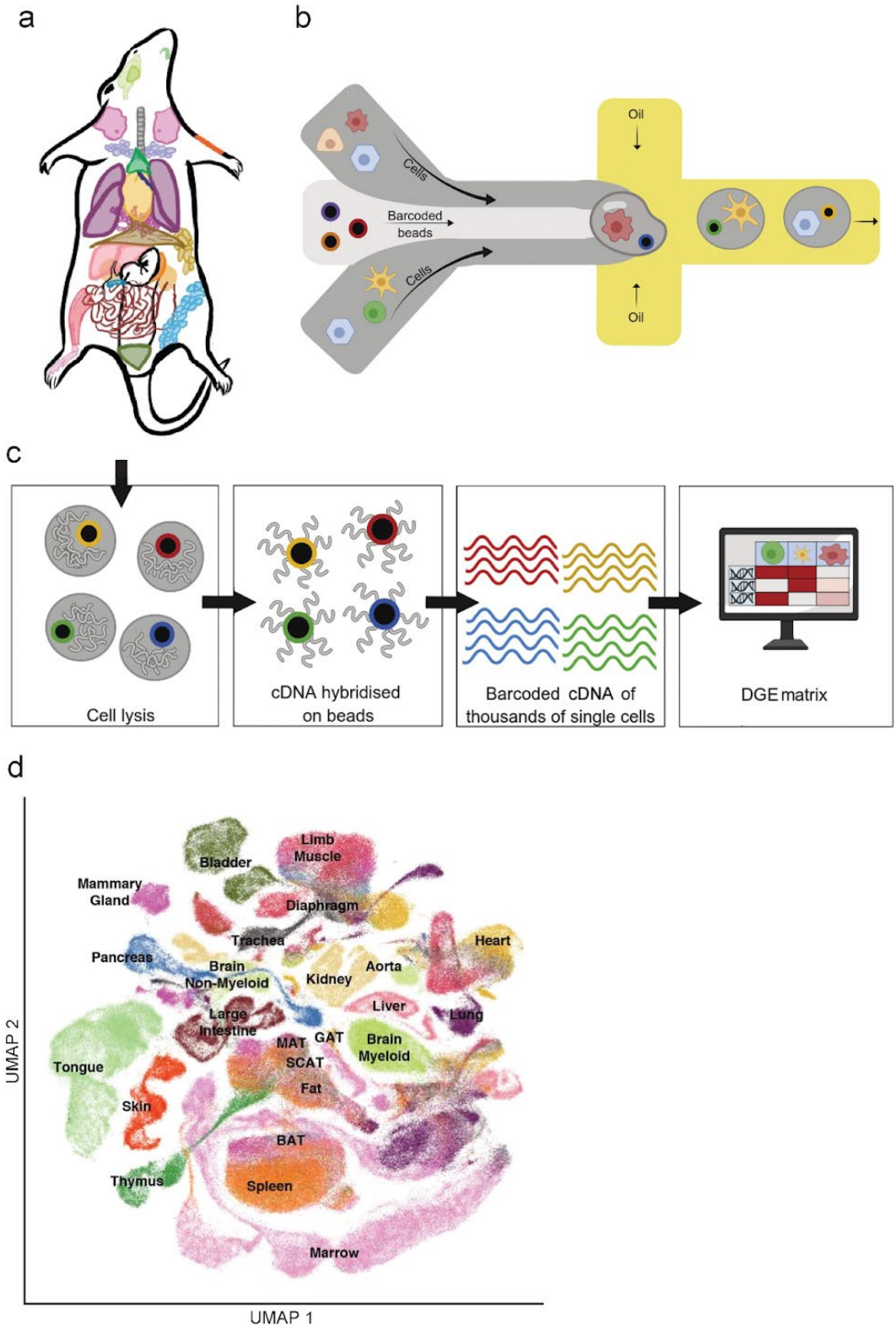

a
b
Oil
Cells
Barcoded beads
Cells
Oil
c
Cell lysis
cDNA hybridised on beads
Barcoded cDNA of thousands of single cells
DGE matrix
d
Limb Muscle
Bladder
Mammary Gland
Diaphragm
Trachea
Heart
Pancreas
Brain Non-Myeloid
Aorta
Kidney
Liver
Lung
Large Intestine
MAT
GAT
Brain Myeloid
SCAT
Tongue
Fat
Skin
BAT
Thymus
Spleen
Marrow
UMAP 2
UMAP 1

Figure 3: Using massively parallel transcriptional profiling of single cells to comprehensively resolve cell types

Overview of workflow and results from the 10X genomics system for transcriptionally profiling single cells (adapted with permission from The Tabula Muris Consortium. Nature, 2020 (a,d); Charlotte Rich-Griffin et al. Trends Plant Sci. 2020 (b,c)). (a) Cells from 23 mouse organs dissociated into single cells for processing using the 10x single-cell transcriptional profiling system. (b) Individual cells are encapsulated in droplets together with a barcoded bead and lysis buffer. (c) Cells are lysed within droplets. Reverse transcription yields cDNA molecules hybridised to probes on beads. Sequencing of cDNA yields a library of transcriptomes of thousands of individual cells. Software is used to count unique reads per gene and per cell yielding a digital gene expression (DGE) matrix. (d) 356,213 individual cells cluster into separate cell classes corresponding to the main mouse tissue types.

Molecular tools used to understand neuronal circuit function and formation in zebrafish

Labeling neuronal populations to explore brain function

Our ability to tag (label) specific neuronal populations is key to characterize, manipulate, and understand the development of specific neuronal populations, the circuits they form, and by extensions – how the brain functions[52].

Most of the work in this respect has been done using transgenic lines, which allow both the control of the expressed protein, the expression pattern, and the time at which the protein is expressed. One of the most widely adopted transgenic expression systems is the GAL4/UAS system from yeast. In this system the GAL4 transcription factor binds to the upstream activating sequence (UAS) to drive expression of the gene downstream of the UAS (Figure 4a). In a transgenic system, both of these can be uncoupled: the transcription factor GAL4 can be controlled to express only in a subset of neurons or brain regions; the UAS can drive the expression of a wide-range of synthetic proteins that can be used to dissect the cells and tissues in which it is expressed. The split system has allowed the rapid integration of transgenic driver lines into an ever growing library. With this system, new genetic elements can be rapidly used to drive the expression of existing tools, while new tools can be quickly used with an existing, ever growing library of transgenic driver lines[53]. Despite its great success, some drawbacks of working with this system have been reported[54]. One point of concern is that different individuals of the same transgenic line can present different levels of transgene expression. Over multiple generations of animals, the transgene becomes variegated by random selection. The genetic elements will become slowly silenced over multiple generations. And, in some animals this will happen faster than in others. To mitigate the

chance of strong variegation requires careful attention in choosing which animals are best suited for an experiment and for maintaining the transgenic line.

The integration of any genetic element into the zebrafish genome has always been a challenge. Over decades of research, multiple approaches have been developed and applied with varying success rates judged by the generation of a stable transgenic animal capable of passing on a transgene to viable fertile progeny. One of the most successful tools for generating transgenic zebrafish has been a teleost transposon-based system. The tol2 system, discovered in medaka, contains two large sequences capable of inserting themselves, and anything in between them, into homologous sequences in the teleost genome at a high success rate[55]. Scientists have used this transposon to integrate transgenes into an animal's genome. While the unique tol2 sequences need to remain unchanged, there is potentially no limit on what can be inserted between them, and subsequently, integrated into the genome. Thus, the tol2 system has offered a flexible tool for generation of transgenic zebrafish. The main drawback of this system is that it is hard to predict where a transgene will be integrated into the genome[55]. Some areas may be very suitable for driving expression of synthetic protein. Other areas may block synthesis altogether. And, in extreme cases, an integration site may be an important gene, causing the death of an embryo. To overcome these challenges, one should carefully select healthy founders and progeny with strong reporter gene expression.

In a way, the large downside of the GAL4/UAS system, is not purely inherent to the system: the increasing variegation over multiple generations, is also a consequence of specific integration sites. Some transgenic lines are very stable over generations. Better control over the integration site, can alleviate transgene silencing and variegation. While this is not possible with a randomly integrating transposon system, the CRISPR/Cas9 system can drive integration into specific, predetermined, genomic sites[56–59]. Random integration, however, can be useful. Randomly inserting GFP into various genetic loci has produced a plethora of enhancer-trap lines – driving expression of GFP in various cellular patterns across the brains[52] (Figure 4b). While screening each enhancer-trap line can be tedious, in absence of clear marker genes for neuronal populations, they can offer unprecedented genetic access into the myriad neuronal populations composing the brain.

Once a transgenic driver line is established, it can be combined with various tools to understand the function of neurons labeled in the transgenic line. Various fluorescent proteins can be used to target these cells for further explorations. Stochastic fluorescent labeling of single neurons can resolve the morphology of neuronal cell types[60,61] (Figure 4c). The cells can also be sorted using fluorescence activated cell sorting (FACS) for further genetic dissection. Cells labeled within a driver line can also be functionally characterized; expression of a genetically encoded calcium indicator such as GCaMP (a fusion protein of GFP and calmodulin sensitive to calcium changes), can allow us to monitor the activity of a neuron in response to various stimuli[62]. Lastly, chemogenetic ablation, by which a cell death can be induced in a specific neuronal population, can allow us to understand how a specific neuronal population contributes to a specific behavior[63].

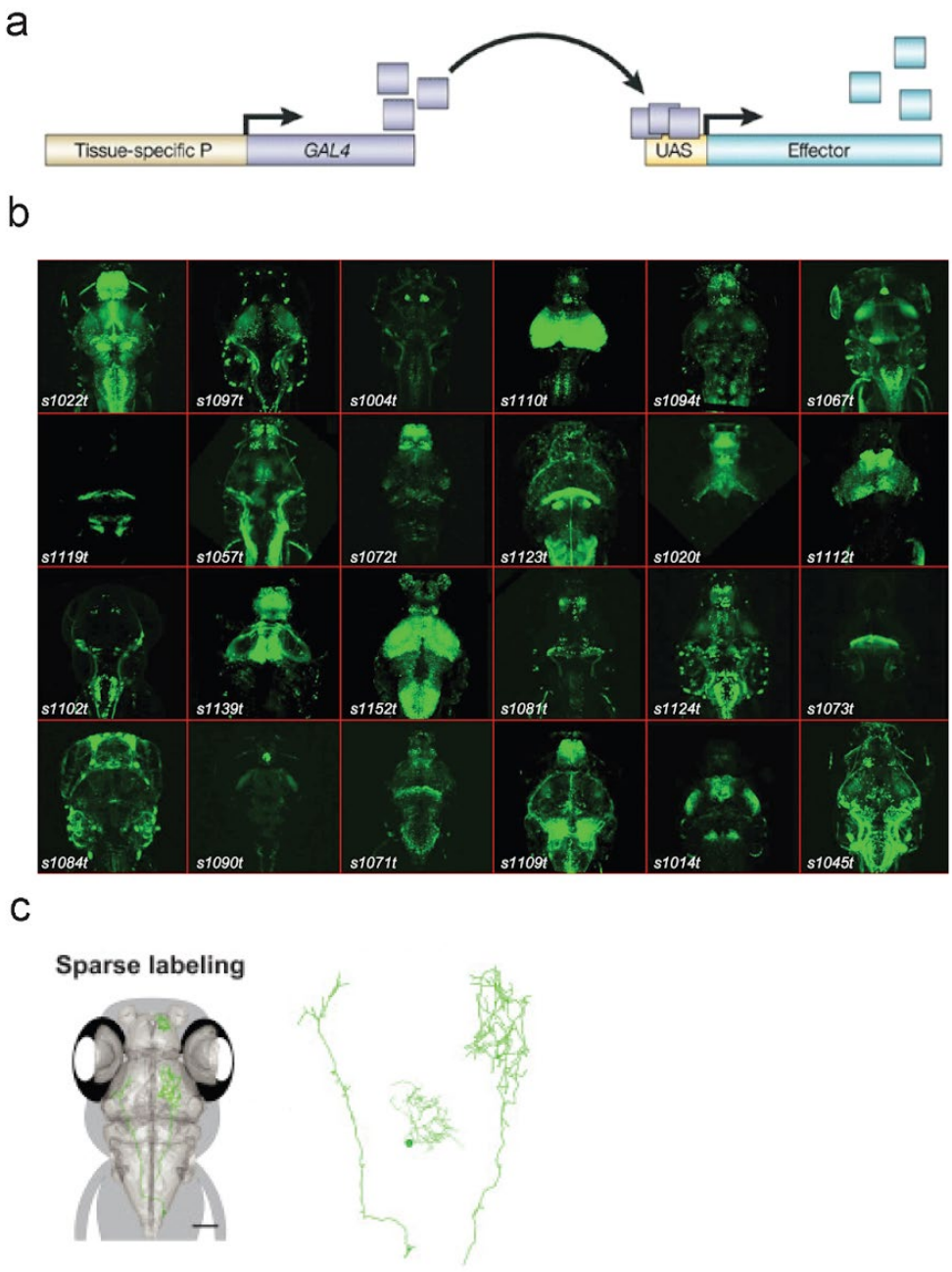


Figure 4: Using transgenic tools to explore neuronal populations

(a) Illustration of the GAL4/UAS system (adapted with permission from Wimmer, E. *Nat Rev Genet* 2003). (b) Expression of the GAL4/UAS system under the control of different enhancer-trap lines in zebrafish (adapted with permission from Scott EK and Baier H, 2009. *Front. Neural Circuits*). (c) Example of sparse labeling of neurons in zebrafish larvae using the GAL4/UAS system. The complete morphology of neurons can be resolved in this way (adapted with permission from Kunst, M. et al. Neuron, 2019).

Using optogenetics to dissect neural circuit function

Optogenetics has been widely adopted as a method of choice to probe neural circuit architecture and the relationship between neuronal populations and animal behavior[21,34,64,65]. In practice, optogenetics refers to a variety of genetically encoded light-gated, ion-selective, transmembrane proteins. They all open or close in response to light allowing or blocking the flow of ions - increasing or decreasing the chance a neuron will fire an action potential. These transmembrane proteins (most commonly channels, but also pumps) are naturally occurring among microbes, algae, and animals, but have been heavily engineered to suit an array of experimental designs.

The zebrafish larva is especially suited for optogenetic manipulation as it is translucent, allowing direct access to most neuronal populations across the entire animal via an external optic fiber or bright LED. Animals can be either restrained, or free-swimming. And, the large number of existing driver lines allows targeting of specific neuronal populations for stimulation. While many tools exist, channelrhodopsin-2 (ChR2) from the green algae *Chlamydomonas reinhardtii*, is most commonly used. The channel is activated by blue light, causing it to open and allowing the flow of cations into the cell causing depolarization and firing of action potentials (Figure 5). Under the right conditions, it is also possible to elicit behavior by optogenetic stimulation. Optogenetics can also be combined with functional calcium imaging using GCaMP to test neuronal connectivity. The use of optogenetics in zebrafish has had profound impact on zebrafish circuit neuroscience and has been used to understand neuronal connectivity, motor circuits, and the sufficiency of neural populations to zebrafish behavior. Taken together, optogenetics is an invaluable tool in larval zebrafish that can reveal how neural circuits generate behavior.

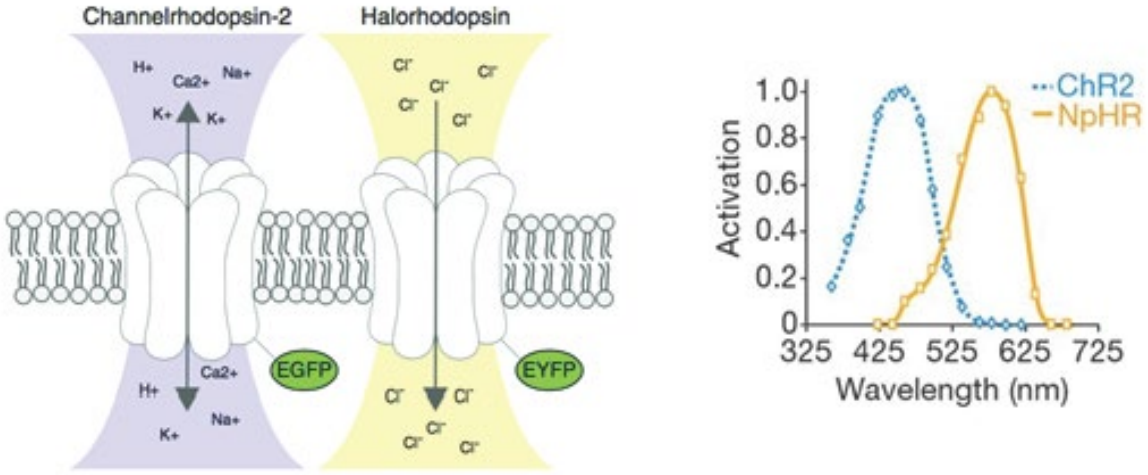


Figure 5: Optogenetic tools can activate and inhibit neuronal activity

(Left) Schematic drawing of two optogenetic tools, channelrhodopsin-2 (ChR2; activator) and halorhodopsin (NpHR, inhibitor). They differ in their activation spectrum,

ion selectivity, and kinetics (adapted with permission from Mutter et al. 2014). (Right) Activation spectrum for ChR2 (peak sensitivity 470 nm) and NpHR (peak sensitivity 580 nm) (adapted with permission from Zhang et al. 2017).

Visualizing gene expression to understand brain structure and development

To understand how the brain is organized, it is important to not only know which genes define cell types, but also where they are expressed and when their expression takes hold. While some cell maintenance genes are always active, others turn on and off in different places and at different times. Early patterning events, for example, are critical for establishing an animal's body plan (Figure 6a). The coordinated expression of genes orchestrating these events is critical during embryonic development. The brain shows a similar segregation of genes where a single gene or a combination of genes contribute to establish the transcriptomic signature of a brain area – generating an environment that is unique and contributes to the establishment of region-specific cell types (Figure 6b).

The brain also continuously changes between birth and adulthood. In zebrafish, the birth of new neurons continues throughout life. New behaviors appear that did not exist in an earlier stage[66] – likely generated by addition of new neurons or modulation of existing neurons and circuits. However, these late changes are heavily influenced by earlier events. Failure to establish early gene-expression gradients can create a cascade of long-lasting effects. Likewise, changes in sensory systems, which inform the brain of its environment, can affect gene expression in the brain[36]. New visual and olfactory experiences routinely cause changes to gene expression. Thus, visualizing where and where certain genes are expressed is critical to understanding how the brain develops, how the identity of brain regions and their cell types is maintained, and how these change over the lifetime of an animal in health and disease.

The most common method to visualize gene expression is RNA in-situ hybridization. In this, a synthetic DNA probe is flushed into cells where it attaches (hybridizes) to complementary mRNA molecules transcribed from a single gene[67,68]. The identification of the location and amount of mRNA can be detected by color generated indirectly by the synthetic DNA probe. The color can then be visualized by either transmitted light or fluorescence (Fluorescent In-Situ Hybridization, FISH). Fluorescence, though, has several advantages over colors visible to the eye. Using a confocal microscope, it is possible to image fluorescence to not only visualize mRNA expression in single-cells, but to also count the number of mRNA molecules. It is also possible to combine multiple probes labeled with different fluorophores (fluorescent dyes) to target different mRNA species (Figure 6b). In this way, we can visualize the expression of different mRNA molecules in the same tissue. One major advantage of zebrafish larvae in this is that FISH probes can be imaged with minimal sample preparation. While most animal brains are large and optically inaccessible, zebrafish larvae have a small, yet complex brain, and they are translucent (allowing light to travel through the brain with minimal

disturbance). FISH can be routinely applied to visualize gene-expression patterns across the entire larval zebrafish brain.

More recently, a number of methods grouped together as spatial transcriptomic methods have made huge strides in increasing both the number of mRNA species that can be visualized within the same tissue and the resolution at which they can be resolved[69–72]. The way in which these methods achieve this is diverse: the Visium spatial transcriptomics solution (10x Genomics) uses spatially-restricted barcodes distributed in cell-sized dots over a large collection area, offering an unbiased spatial-characterization of all mRNAs from tissue slices covering a whole mouse brain (Figure 6c). In-situ sequencing allows targeting of hundreds of mRNA species in tissue by decoding their nucleotide sequences in cells. The most promising techniques are MERFISH and seqFISH. Both use temporal barcoding to target and decode individual mRNA species through multiple hybridization rounds - resulting in individual mRNA molecules being assigned a barcode matching a single mRNA species. Using these various methods, researchers have been able to decode tens of thousands of mRNA species across entire brains.

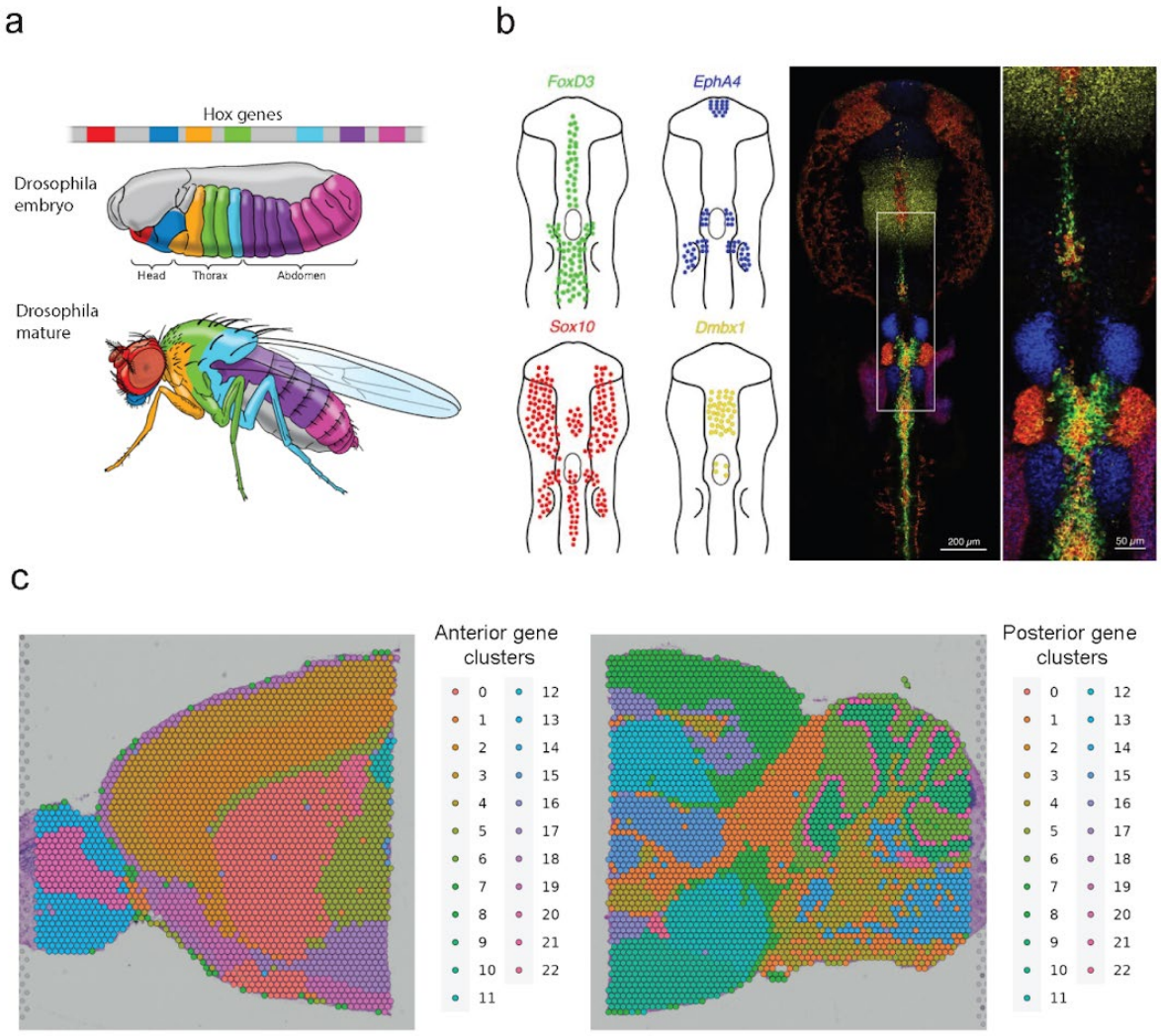


Figure 6: Visualizing gene expression in tissue

(a) Hox gene patterning in Drosophila from embryo to adult (adapted with permission from "Genetic Signaling: Transcription Factor Cascades and Segmentation" Nature Education). Different colors in embryo represent the expression of different Hox genes. Different colors in adult represent corresponding body segments in embryo. (b) Multiplexed mRNA imaging in whole-mount chicken embryos. Left, Expression schematics for four target mRNAs in the head and neural crest: FoxD3, EphA4, Sox10 and Dmbx1. Right, full field image and zoomed-in view of four individual images of fluorescent in-situ hybridization for the same genes as in the left panel (adapted with permission from Choi, Harry MT, et al. Development 2018). (c) Two panels showing anterior (left) and posterior (right) mouse brain slices captured on the 10X Genomics Visium spatial transcriptomics system. The full square shows the capturing area. Colors show where mRNA from tissue was collected and sequenced. Hexagons represent individual capturing areas. Colors show computational clustering of capturing areas based on their transcriptomic signatures. Anterior and posterior colors are independent clustering results and

not interchangeable (adapted with permission from satijalab Seurat tutorial for visium data analysis).

The visual system

The visual system is the collective name for the part of our brain that perceives light coming from the environment: from sensation to perception. The function of the visual system begins with the retina, which detects incoming light, extracts features from it, and relays information to areas in the brain involved in further visual processing. While most of the brain is encased in a skull, making it hard to manipulate and record from, the eye, and the retina within are easily accessible[73]. Injections into the eye have allowed tracing of connections across the visual system and uncover the structure and organization of the visual system as a whole – shedding light on how visual information is processed in the brain[74]. It is thus no coincidence that the visual system is by the far the most studied sensory system in vertebrates.

The structure of the visual system

The retina is the gateway to the visual system. It senses incoming light, processes it into discrete information channels, and transmits those to major visual processing centers in the brain. The retinal structure allows it to preprocess information before transmitting it to brain areas. It is evolutionarily conserved and consists of three main cell layers: the outer nuclear layer (ONL), the inner nuclear layer (INL), and the ganglion cell layer (GCL). Information flows from the outer layer (closest to the inner surface of the eyes) to the inner ganglion layer (Figure 7a). Photoreceptors in the ONL detect incoming photons via specialized molecules and transmit that information to bipolar cells in the INL. The bipolar cells relay information to ganglion cells (RGCs) in the GCL. RGCs are the sole output from the retina and so form a visual bottleneck; they extend their axons (projections) into the brain along optic tracts – forming the optic nerve. In mammals RGCs from each eye project to both sides of the brain; RGCs on the left retina from each eye project to the right hemisphere of the brain and vice versa. In most teleosts, however, the optic nerves are completely crossed - all RGCs from the left eye project to the right side of the brain and vice versa. By the point RGCs relay information to the brain, it is highly processed and transmitted along discrete information channels consisting of: direction of a moving object, its orientation in space, its size and color, and whether it is getting brighter or dimmer[75]. As not all RGCs project to all areas of the brain, different areas receive different types of information and are able to perform selective visual processing[76].

The areas in the brain directly innervated by RGCs are called retinorecipient (receiving retinal input). Different species can have a different number of retinorecipient areas, and

as these are innervated by different RGCs, they also process and combine different aspects of the visual scenery[76–80]. The main retinorecipient areas in all animals include: the pretectum, the superior colliculus (part of the midbrain tectum), the hypothalamus, and the geniculate nucleus of the dorsal thalamus (dLGN), which is the main retinorecipient area in primates. Both the superior colliculus and the pretectum are directly involved in adjusting eye position[20,81–84]. The pretectum processes shifts in the whole visual field correcting the eyes position for small movements – giving a stable image of the world. The superior colliculus process similar inputs, but is involved in head and neck movements to stabilize the visual field and in focusing objects in the center of the visual field.

While the dLGN also processes visual information locally, it also enables conscious visual processing by transmitting visual information to the visual cortex[85] (the main visual processing center in the brain). It is in this area that most of the processes that we intuitively associate with visual processing happen (object tracking, and object and face recognition). The dLGN, or the geniculate pathways, is the main way in which visual information is relayed to the visual cortex. Teleosts, however, do not have a cortex. Hence, most of the visual processing in these animals is done locally in the hypothalamus, thalamus, pretectum, and tectum. The tectum is the main retinorecipient area in teleosts, innervated by >90% of all RGCs[86]. It is in fact so big in teleosts compared to other brain areas that it's termed the optic (visual) tectum.

Zebrafish brains have ten retinorecipient neuropil areas (retinal Arborization Fields, AFs, containing mainly neuronal axons and dendrites) numbered 1-10[76]. The numbers represent the order in which the AFs appear along the optic tract from the ventral to the dorsal side. The different AFs are formed by different combinations of RGCs arborizing in discreet areas; they relay specific processed information from the visual scenery, and contribute to different visual behaviors. AF1, which is part of the suprachiasmatic nucleus, and AF2, part of the preoptic area likely contribute to circadian behavior and visual background adaptation (VBA) - a neuroendocrine response allowing fish to blend with their surroundings by brightening or dimming their skin based on how dark or bright their environment is. AF3 is part of the ventral thalamus, but very little is known about its function. AF4 is also part of the ventral thalamus. Though more anterior than AF3, AF4 likely contributes to phototaxis – a light seeking behavior. AF5 and AF6 are part of the pretectal accessory optic system and are innervated strongly by RGCs responding to directional motion or to large changes in luminance across the visual field. Both are likely involved in the optokinetic response (OKR) and optomotor response – behaviors responsible for gaze and body stabilization. AF7 is located in the parvocellular pretectal nucleus and is involved in prey detection. AF8 belongs to the central pretectal nucleus, innervated by RGCs responding strongly to large expanding objects. Detecting such events is critical for avoidance behavior, from conspecifics (collisions) or predators. Both will create an expanding object on the retina as they physically approach an animal. AF9 is part of the periventricular pretectal (PP) nuclei. It can be split in two: dorsal (PPd) and ventral (PPv). RGCs arborizing in the AF9 PPd mainly respond to brightening while ones terminating in AF9 PPv mainly respond to dimming. AF10 refers to the entire neuropil region of the optic tectum. It is involved in a wide-range of tasks including: localization and identification of objects in the visual field, approaching prey,

and avoiding obstacles. RGCs terminals in the tectum are further mapped retinotopically: RGCs from the nasal retina mainly innervate the posterior tectum and RGCs from the temporal retina mainly innervate the anterior tectum.

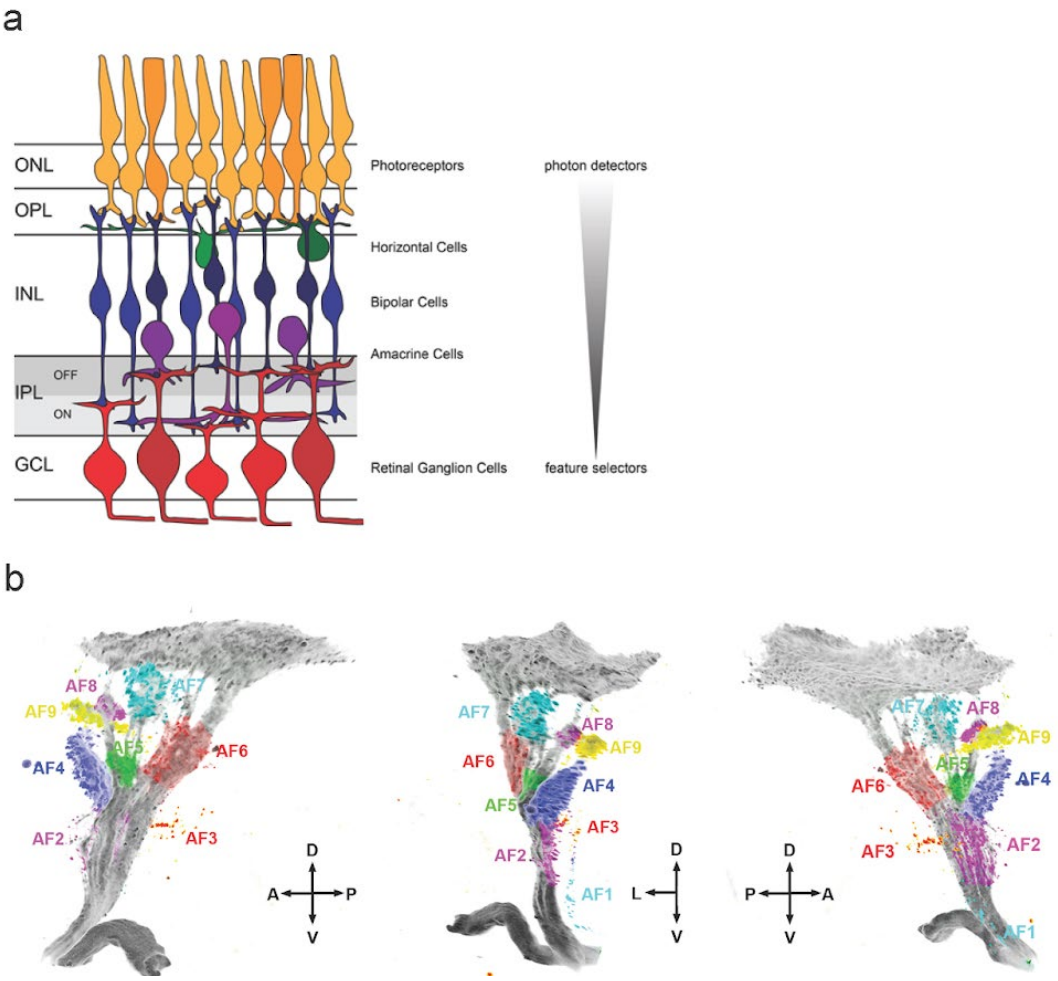


Figure 7: structure of the visual system

(a) Schematic representation of the laminar structure of retina. Each layer harbors different retinal cell classes: PRs sit at the back of the eye in the ONL; HCs, BCs and ACs reside in the INL; and RGCs are located in the GCL. Visual signal is conveyed from PRs to BCs via their synapses in the OPL. There are two groups of BCs: ON BCs encode light increments and terminate in the proximal IPL, whereas OFF BCs relay light decrement signals to the distal IPL. BC input is received by RGC dendrites within the IPL. The vertical, excitatory information flow from PRs via BCs to RGCs is modulated by lateral interactions from inhibitory HCs in the OPL and ACs in the IPL. Intra-retinal circuits formed by diverse cell types extract visual features, which are encoded by RGC types and send to the brain in parallel processing channels. Müller glia and other non-neuronal cells are omitted in this schematic (thesis of Yvonne Kölsch, PhD).

PRs: photoreceptors, HCs: horizontal cells, BCs: bipolar cells, ACs: amacrine cells, ONL: outer nuclear layer, INL: inner nuclear layer, GCL: ganglion cell layer, OPL: outer plexiform layer, IPL: inner plexiform layer. (b) The retinal projectome of larval zebrafish. Robles et al. (2014) created a triple-transgenic zebrafish carrying atoh7:Gal4-VP16 to drive expression of membrane-targeted mCherry (to label all RGC axons) and Synaptophysin-GFP (to label presynaptic terminals). Three views of the confocal image stack, taken at 6 dpf, are shown. AFs are from the left side of the brain (adapted with permission from Baier, H., & Wullimann, M. F. (2021). Journal of Comparative Neurology). A, anterior; P, posterior; D, dorsal; V, ventral; L, lateral.

The development of the visual brain

The visual system involves the retina and optic nerves, and brain regions involved in processing information incoming from the retina. These components build and wire through intricate processes involving the unfolding of transcriptional programs defining a commitment of cells to specific fates (cell types). Depending on the specific cell type, a cell will sense its environment and extend its processes throughout to find appropriate partners with which it will form functional synapses[87,88]. The successful culmination of these processes throughout the brain results in an animal's ability to interpret its visual scenery and generate appropriate behavior.

All major cell types in the brain and retina begin as progenitors. These progenitors will often divide asymmetrically - giving rise to a new progenitor cell and a precursor cell that commits to a neuronal fate[89–91]. Commitment begins with exiting the cell cycle, down-regulation of genes required for maintaining a progenitor state, and up-regulation of genes required for neuronal commitment. Neurons undergoing commitment will constantly up- and down-regulate transcriptional programs (a grouping of many genes working together) until committing to a terminal fate. A fully committed cell will then continuously express genes required to maintain that fate[92].

Concomitant with the transition from a mitotic (cycling) progenitor to a post-mitotic committed neuron, cells also migrate to their appropriate positions within the brain[93,94] (Figure 8a). Progenitors lie only in specific areas of the brain (proliferative zones), unlike their fully committed daughter cells. As a neuron commits to a certain fate, it also commits to a place in the brain away from the proliferative zones. As new transcriptional programs unfold, a cell becomes sensitized to some molecular cues, but not to others[95]. Cues originating from specific locations in the brain form molecular gradients as they diffuse and provide cells with paths and waypoints they can follow to their final location[96]. The unfolding of downstream transcriptional programs further sensitizes a cell to different cues (chemoattractants) while also desensitizing it to others. In this way a cell can constantly follow new migration paths without turning back. Neurons will thus continuously commit to a fate leading away from a progenitor state by making fate decisions that also allow them to reach the appropriate position in the brain.

Neurons will use these same principles to select where their axons and dendrites should project[97]. Cells from the dorsal horn of the spine need to cross its midline to form connections to the brain. To do so they express a netrin receptor which sensitizes them to a netrin gradient originating from the midline (Figure 8b). Cells expressing the netrin receptor will send their axons to the midline following the gradient. To cross the midline, however, they must initiate a new transcriptional program. In its absence, neurons will terminate their projections at the midline, the place where netrin is most abundant. The increasing concentration of netrin causes the upregulation of another receptor, robo, which is sensitive to slit. Slit is also abundant in the midline, but instead of functioning as an attractive force, it functions as a repellent. An axon will initially be much more sensitive to the attractive force of netrin, however, once it reaches the midline, it is much more sensitive to the repellent force of slit. In this way dorsal horn spinal neurons can cross the midline, not turn back, and form appropriate tracts to the brain.

In the retina, the situation is not much different. Retinal ganglion cells (RGCs) will extend their axons to the optic disk to exit the retina (the only exit from the retina). RGCs, as other neurons, express a transcriptional program which sensitizes them to molecular cues leading to exit the eye. Once they do so, they are repelled by it: causing them to move away from the retina (Figure 8c). They then follow additional cues in the brain along predefined optic tracts leading to retinal arborization fields (AFs) where they will form synapses within visual processing centers[98]. As neuronal types in the brain express different genes, they present different targets for RGCs enabling the selection of where to form a synapse[99–101]. It is this first step of molecule recognition and selective synaptogenesis (forming of synapses) that gives the visual system a scaffold on which to build ever more complex neuronal networks to process visual information.

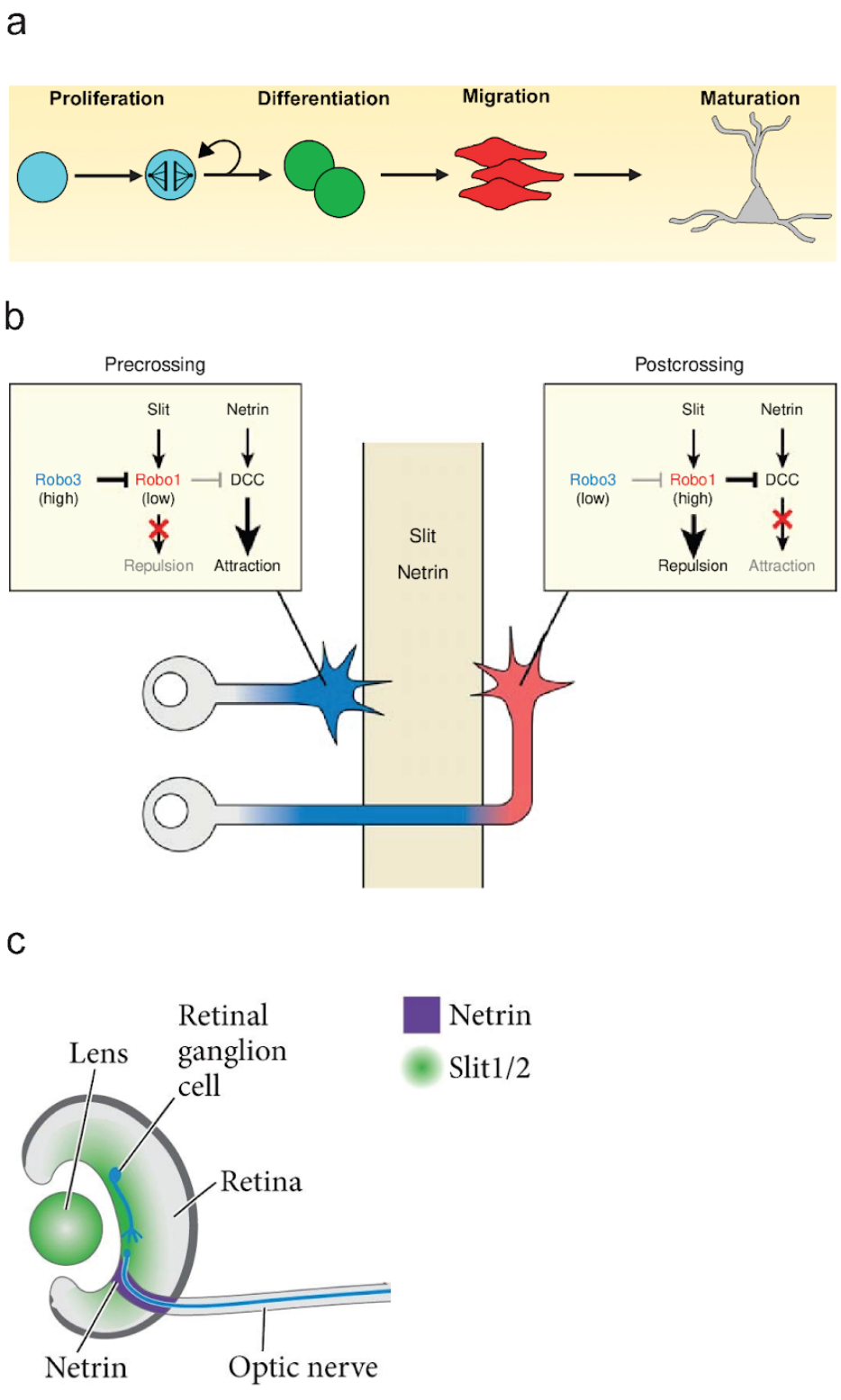


Figure 8: Assembly of the visual system

(a) Neural stem cells (blue) undergo proliferation and either give rise to more stem cells (self-renewal) or daughter cells (green, differentiation. Daughter cells then migrate (red) from

proliferation zones to their terminal positions. Upon reaching their terminal positions, newly born neurons (gray) extend neurites and properly form dendritic arbors (adapted with permission from by Francesca LiCausi and Nathaniel W. Hartman Int. J. Mol. Sci. 2018). (b) Robo3 is a negative regulator of Slit-Robo signaling. In precrossing axons high level of Robo3 is thought to suppress Robo1 mediated repulsion which in turn activate netrin-1 induced attraction. But in postcrossing axons Slit-Robo1 mediated repulsion activated due to the lower level of Robo3 and this event in turn silences the netrin mediated attraction (adapted with permission from Giasuddin Ahmed, 2010). (c) Multiple guidance cues direct the movement of retinal ganglion cell (RGC) axons. Guidance molecules belonging to the netrin, slit, semaphorin, and ephrin families are expressed in discrete regions at several sites along the pathway to direct the RGC growth cones. RGC axons are repelled from the retinal periphery, probably by chondroitin sulfate. At the optic disc, the axons exit the retina and enter the optic nerve, guided by netrin/DCC-mediated attraction. Once in the optic nerve, the axons are kept within the pathway by inhibitory interactions (adapted with permission from oxford university press)

Activity-dependent and activity-independent brain development

While the visual system undergoes the myriad developmental processes resulting in functional assembly, the individual neurons making up the system do not wait in silence; neurons are active throughout development, firing action potentials and communicating with their synaptic partners[102]. Neuronal activity, in fact, plays a pivotal role during neural development. And, in absence of synaptic partners or activity, neurons may eventually die.

Because mammals develop in-utero, visual sensory input plays little role in early development. This extends even further for mammals that are born with eyes closed. However, this does not mean that the retina has no influence over brain development. During uterine development the retina produces internal activity in the form of propagating waves where neurons fire in a spatiotemporal pattern[14,103] (Figure 9a). The waves burst spontaneously, but their propagation across the retina is not. A wave will propagate across the retina - activating neighboring RGCs in sequence. This activity pattern is relayed through RGCs axons to visual processing centers in the brain. Even though the initiation of waves is spontaneous, the correlated activity between neighboring RGCs is not. And, because neighboring RGCs will be activated together much more frequently than RGCs far apart, it instructs the visual centers on the physical map of the retina[104] (which inputs originate from which location on the retina). These types of events, however critical during uterine development, are not unique to mammals[105]. Retinal waves have also been reported in zebrafish, which develop in a translucent egg - exposed to environmental visual cues from fertilization.

Contrary to the zebrafish visual system which is completely crossed (RGCs from each eye project to the contralateral side only), the mammalian visual system is partially crossed (RGCs from the nasal retina project to the contralateral side, while RGCs from the temporal retina project to the ipsilateral side). Simply, the right half of the retina projects to the right side of the brain, and the left side of the retina projects to the left

side of the brain (Figure 9b). This causes a visual challenge by mixing of the left and right visual fields. In the lateral geniculate nucleus (LGN), the major site for retinal innervation in mammals, this challenge is resolved by organizing the input from each eye into alternating layers. And, in some mammals (including primates), this separation is further visible in the visual cortex where the inputs from each eye are organized into columns[106,107] (ocular dominance columns). The process of segregating the inputs from each eye is different from the classical developmental mechanisms of chemoaffinity and repulsion. Instead, it is driven by neuronal activity[106,107]. Inputs arising from one eye fire together in a more correlated manner than with those of the other eye. The eye-specific layers, or columns, form by locally maintaining synaptic inputs that activate jointly. And, also, by removing synaptic inputs that fire independently. In this way, a layer will continuously reinforce inputs from a single eye while rejecting inputs from the other. It is this competition that allows the visual system to separate inputs arising from different eyes into layers (and columns).

Internally generated neuronal activity, however, is just one type of activity. At later stages of development (in some mammals, after eye-opening), visual experience can influence the development of the visual system. Experiments in cats have shown that if specific visual stimuli are deprived from a young animal during a critical period, the animal will not develop the ability to notice these stimuli[2] (Figure 9c). In these, kittens were reared for a prolonged period of time in the presence of shapes in a single orientation, never seeing other orientations - depriving them of a specific visual experience. The kittens never developed the ability to perceive the missing orientation and were essentially blind to objects in that orientation. These experiments, and others, showed that there is a critical period when brains need to be exposed to certain stimuli to be able to process them later. On a cellular level, this means that orientation-selective cells (responding to a preferred orientation) fail to develop properly. Similar experiments have also been conducted in the context of direction-selectivity. Direction-selective cells will respond consistently and strongly only to objects moving in a single direction. Visual experience, however, can change this fundamental property. Direction-selective cells can change their preferred direction (the one they respond most strongly to) upon repeated visual exposure to objects moving in a different direction[10]. And, it is even possible to entrain a preferred direction in cells which show no direction-selectivity[11]. Research into developmental biology has also shown that visual experience can play a role in fate determination and migration of cell types in visual centers[9]. Thus, there exists an abundance of evidence showing the critical role RGCs play in the development of visual processing centers.

In zebrafish, however, it appears that brain development is largely independent of visual experience[108,109]. Zebrafish larvae exhibit a large array of visual behaviors including: optokinetic and optomotor stabilizing behavior, prey capture, and light seeking[110]. Larvae reared in dark or light show no difference in these behaviors, suggesting that visual experience is not required for their development[109,111]. The anatomy of a larval brain lacking all RGC input is surprisingly similar to that of a brain which receives RGC input. This also suggests that cell types in the larval brain undergo fate commitment in a largely activity-independent manner. While it is largely assumed that cell types in the vertebrate visual system depend on RGCs for terminal fate commitment, a systematic

exploration of visual cell-types in absence of visual experience or RGCs altogether is still missing.

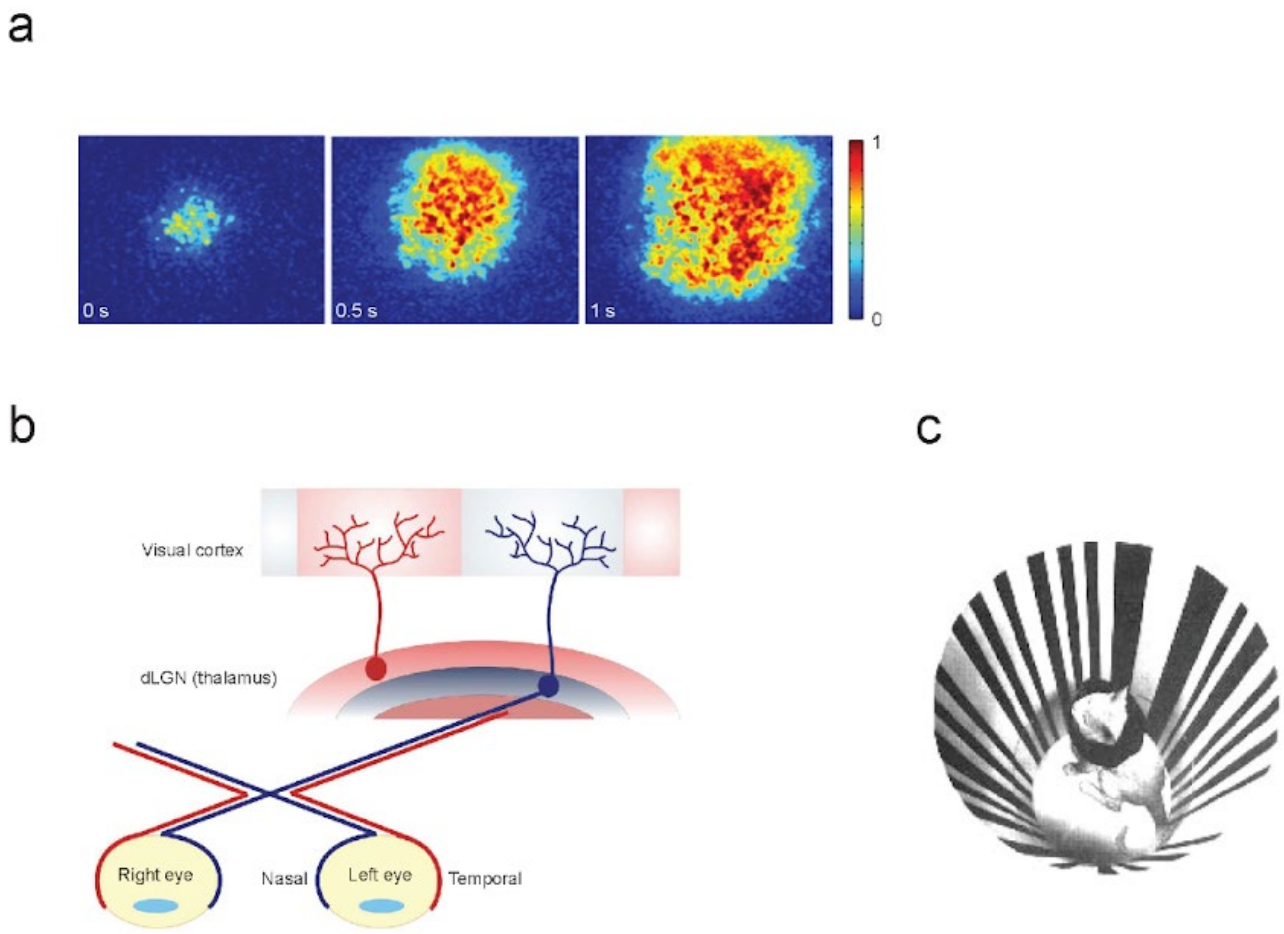


Figure 9: Activity plays a role in the development of the visual system

(a) Fluorescence image of a retina loaded with a fluorescent calcium indicator OGB-1 AM. Sequence of images are pseudocolored to represent the fractional change in fluorescence (ΔF/F) at each pixel following current injection normalized by color to the highest value. Interval between panels is 0.5 seconds (adapted with permission from Kevin J. Ford et al. Journal of Neuroscience 2012). (b) Schematic structure of the mammalian visual system. Retinal ganglion cells from both eyes project to the thalamic dorsal lateral geniculate nucleus (dLGN). Partial crossing of the two pathways occurs in the optic chiasm. Neurons of the lateral geniculate nucleus project to the visual cortex where neurons form eye-specific patches, ocular dominance columns (ODCs). Neurons of the LGN send their projections to the primary visual cortex (V1), where their activity is required for the development of eye-specific patches of neurons in layer IV, i.e., the ocular dominance columns (ODCs; adapted with permission from Adema Ribic, Histocompatibility, 2012). (c) Picture of a cat, reared during its early life in an environment which contained no horizontal stripes, would fail to develop neurons in the visual system which

respond to horizontal edges (adapted with permission from Colin Blakemore & Grahame F. Cooper, Nature, 1970).

THESIS OBJECTIVES

Animals interpret their environment using a number of sensory systems including: visual, olfactory, auditory, mechanosensory and more. The stimuli detected by these systems inform animals about their environment and allow them to behave accordingly. They find food, interact with conspecifics, hide from predators, navigate, survive, and mate. The architecture underlying these systems has been sculpted over the course of evolution. Neurons involved in sensing one feature of the environment, project to specific brain areas. Neurons involved in sensing other features will show different innervation patterns, synapsing with different partners and forming different neuronal circuits. Their development, dictated by genetic programs, follows a predefined trajectory fit to interpret a specific biological niche. Despite this, the sensory system can influence the brain's development. It can inform it of its environment and of changes to it, and so alter the course of development. And so, scientific discussions are ongoing as to the role the sensory surface plays in shaping the structure and function of downstream brain areas.

My thesis aims to tackle this by looking at the development of the visual system. I aim to understand how areas in the brain, which normally receive direct retinal input, develop in complete absence of retinal input. My thesis can be divided into the following parts:

Part I Characterizing cell types in visual processing centers

In this part I used massively parallel transcriptional profiling of single cells to molecularly define cell types across larval zebrafish visual processing centers. Analysis of this data uncovered genetic markers for most visual cell types.

Part II Uncovering the spatial organization of cell types in visual brain centers

In this part, I used the markers uncovered in part I to probe at the spatial organization of visual centers. I was able to uncover the location of putative cell types, to generate an atlas of gene expression patterns, and to uncover combinations of genes that define visual brain areas.

Part III Understanding the inner workings of the "blind" visual system

In this part I explored changes to the visual system of blind mutants lacking all RGCs. Based on my earlier work on the "normal" visual system, I could draw powerful conclusions as to the role that the retina (the sensory surface) plays in the development of downstream visual areas.

EXPERIMENTAL PROCEDURES

Zebrafish husbandry

Adult and larval zebrafish were maintained on a 14:10 hour light:dark cycle at 28°C. Embryos were kept in Danieau's solution (17 mM NaCl, 2 mM KCl, 0.12 mM MgSO4, 1.8 mM Ca(NO3)2, 1.5 mM HEPES). All animal procedures conformed to the institutional guidelines set by the Max Planck Society, with an animal protocol approved by the regional government (Regierung von Oberbayern).

For single-cell experiments one male and one female were placed in individual breeding tanks with a divider the evening before spawning. Dividers were removed at 9.00 am the next morning and fish let spawn until 10.00 am. Eggs were collected shortly after and mixed together. *Tg(HGn12C:GFP)* larvae were sorted for GFP expression 24 hours after fertilization. All single-cell experiments were performed at 6 days after fertilization. For WT *Tg(HGn12C)* experiment, pigmented larvae were used from *Tg(HGn12C)* pigmented males or females crossed with WT Nacre (*mitfa* -/-). For *lakritz* and sibling WT experiments, *Tg(HGn12C:GFP) lakritz* heterozygote adults were crossed with adult *lakritz* heterozygotes. For experimetns with *lakritz* larvae, only *lakritz* were collected, for other experiments they were discarded. For *Tg(1026t)* experiment, *Tg(Gal4s1026t)* were crossed with *Tg(UAS:GFP)*. All larvae were reared at a maximum density of 60 individuals in 10 cm petri dish.

Cell-dissociation for single-cell RNA sequencing

Six-day-old larvae expressing GFP under *Tg(HGn12C)* or *Tg(Gal4s1026t)* were used to label neuronal cell types in the forebrain. Ames medium (Sigma A-1420) was used throughout and prepared according to the manufacturer's guidelines. Before animal handling, 300 ml Ames medium was oxygenated at room temperature for one hour. Buffer pH was adjusted to 7.2-7.3 using HCl and filtered through a standard Steritop filter. Buffer pH was checked again after filtration for a desired range of 7.4-7.5. Oxygenated Ames medium was then placed on ice. Larvae were anesthetized in oxygenated Ames ice slush and rapidly decapitated. Brain material from a maximum of 110 larvae was collected using a broad glass pipette and transferred into chilled oxygenated Ames in a 2 ml tube on ice. Tube medium was replaced with fresh Ames after every material transfer. In addition to transgenic larvae expressing fluorescent protein, 30 non-transgenic larvae were used to adjust FACS gates. Cell suspensions from both samples were prepared in-parallel.

Tissue was dissociated into single-cells using a modified protocol from (Kölsch et al. Neuron 2021). Papain solution [25 U/ml final] was prepared by mixing 4810 µl of oxygenated Ames with 89.3 µl papain stock 42.8[mgP/ml], 32.7[U/mgP], 50 µl DNaseI [13K U/ml] (Sigma D-4527, 40K Units), and 50 µl L-cysteine [152.2 mM] (Sigma C-1276). Papain solution was then placed for 15 minutes in a tabletop spinning incubator preheated

to 34°C spinning at 10 RPM. The solution was then examined: though initially milky, the papain solution becomes transparent when activated. Papain solution was not filtered pre- or post-activation. Ames buffer was removed from sample tubes and replaced with 2 ml activated papain solution. Samples were placed in the same tabletop incubator at 34°C for one hour spinning at 10 RPM. After 20 minutes, samples were carefully triturated five times with a narrow glass pipette flamed at the tip to avoid shearing. After one hour, the sample was placed shortly on a bench to let the biological material sink to the tube's bottom. Papain solution was removed and replaced with 1 ml papain inhibitor solution (4450 µl oxygenated Ames, 500 µl ovomucoid stock, 50 µl DNaseI [13K U/ml]). 10x ovomucoid stock was prepared as follows: 150 mg BSA (Sigma A-9418), 150 mg ovomucoid (Worthington LS003087), 10 ml Ames buffer; pH adjusted to 7.4, filtered and then stored at -20°C. After resuspension in inhibitor solution, tissue cells were completely dissociated by triturating a maximum of 30 times with a p1000 pipette (not broad-end) set to 850 µl. A good indicator of a successful dissociation was that there were no observable white particles (brain matter) in solution. Intact eyes were a good indicator that the dissociation was sufficiently gentle to allow high cell-survival. After mechanical dissociation, 1 ml of inhibitor solution was added to each tube. Samples were passed through a pre-wet 30 um filter (Sysmex). Wetting the membrane with Ames buffer allows liquid and small particles to smoothly pass through the mesh filter. Samples were pelleted in a centrifuge pre-cooled to 4°C for 10 minutes at 300 g. Supernatant was removed and the pellet resuspended in 2 ml Ames with non-acetylated BSA (4.5 ml oxygenated Ames, 500 µl 4% non-acetylated BSA, 0.5 µl DNaseI [13K U/ml]). The resuspended solution was filtered through a pre-wet 20 um filter into a new 2 ml tube and short spun to get all liquid through the filter. 2 µl of calcein blue [1 µl/ml] was added to the solution to stain live-cells. Calcein blue was not added to control samples. Suspensions were kept on ice and processed further by FACS.

FACS

BD FACSAria III was used to sort cells. FACS gates were set after 500,000 recorded events to sort live single-cells expressing GFP. Similar gates were used across experiments. Cells were sorted using a 100 um nozzle (~20 PSI) into 2 ml protein LoBind Eppendorf tubes (Eppendorf 0030108132) placed in a cooling holder. Tubes were treated pre-FACS for one hour with 2% BSA in Ames while spinning. Before FACS all liquid was removed. The combination of treatment with LoBind plastic ensured cells will not adhere to the collection tube after FACS. Collection tube was filled before cell-sorting began with 500 µl FACS collection medium containing: 750 µl oxygenated Ames, 250 µl non-acetylated BSA (stock 4%), 0.1 µl DNaseI [13K U/ml]. In total, we aimed to collect 200,000 cells, as in our hands this would fill a 2 ml tube completely. After FACS completed, cell-suspension was centrifuged at 4°C for 5 minutes at 300 g. Centrifuge was not pre-cooled. We found it helpful to memorize the tube's orientation, as often it was difficult to visualize the pellet. After centrifugation, the pellet was resuspended in 60 µl of Ames with non-acetylated BSA diluted 1:10 in oxygenated Ames (0.04% non-acetylated BSA, final). The medium was slowly let slide over the pellet multiple times, and then the tube was gently tilted back-and-forth to push the suspension over the location of the pellet

a few more times. The cell suspension was placed over ice until used for single-cell sequencing.

Calculating cell-suspension density

13 µl oxygenated Ames was supplemented with 2 µl trypan blue (stain for dead cells) and 5 µl cell suspension (1:4 suspension dilution). 20 µl were loaded into a Fuchs-Rosenthal chamber (NanoEnTek, DHC-F01). A minimum of four large squares were counted for live and dead cells using either a dark-field or DIC microscope. Cell density was routinely between 500-1000 cells/µl with viability at ~90% (live cells / live+dead cells).

Single-cell RNA sequencing

Droplet RNA sequencing experiments using the 10X chromium platform were performed according to the manufacturer's instructions with no modifications. Single-chromium chip channels were loaded routinely with 8000 cells aiming to capture 5000 cells with a doublet rate <5%. In our hands, we noticed that loading 8000 cells would usually result in capturing of 3000-3500 cells. For experiments with WT *Tg(HGn12C)*, 17 replicates were collected across 4 experiments. For *lakritz Tg(HGn12C)*, 10 replicates were collected across 4 experiments. For WT *lakritz* siblings *Tg(HGn12C)*, 8 replicates were collected across 3 experiments. For *Tg(Gal4s1026t)*, 4 replicates were collected from 1 experiment. The cDNA libraries were sequenced on an Illumina HiSeq 2500 to a depth of ~50,000 reads per cell.

Alignment of gene expression reads and initial cell filtering

Initial preprocessing was performed using the cellranger software suite (version 3.1.0, 10X Genomics) following standard publisher guidelines. Reads for each channel were aligned to the zebrafish reference genome (GRCz11.98). Further analysis was performed as described below using the Seurat R package (Satija et al., 2015) on the filtered cellranger output matrices.

Initial analysis using Seurat

Output from cellranger was loaded into Seurat allowing for 200-4000 genes/cell, 400-8000 UMIs/cell, and a maximum detection of mitochondrial genes of 12% of all transcripts detected in cell. Unless otherwise stated, batch correction was performed using harmony on experiment and genotype. For analysis of marker genes, the data was first clustered and separated into glutamatergic and GABAergic neuronal datasets, and then batch-corrected and reclustered. For each dataset we recalculated the 2000 most variable

genes (“vst”). We used 18 PCs to cluster our glutamatergic dataset and 20 to cluster our GABAergic dataset. We used a clustering resolution of 1.6 for our glutamatergic dataset and 1.8 for our GABAergic dataset after determining the best resolution using clustree (Zappia et al. GigaScience 2018). Marker genes were extracted using the command FindAllMarkers(…, only.pos = TRUE, logfc.threshold = 0.75), filtered for adjusted p-value < 0.05, and inspected individually using the FeaturePlot visualization tool.

Integration of all three genotypes followed the same pipeline, except that in addition to batch correcting using harmony, we also tested batch correction using the Seurat anchoring method. In short, each combination of experiment and genotype was SCTransformed independently. For each transformed dataset, the 3000 most variable genes were selected as possible integration features. We followed the standard integration pipeline using SelectIntegrationFeatures, PrepSCTIntegration. We used the unrelated WT dataset as our integration space, as it was our largest dataset. We integrated our datasets finally using FindIntegrationAnchors, and IntegrateData. Using this pipeline, we came to the same conclusion as when using Harmony – that we cannot see a difference in cell types between samples.

Independent clustering of datasets from different genotypes

Data from three genotypes was loaded into Seurat and independently processed and clustered. Separation of the GABAergic and glutamatergic datasets was done as described before, but independently for each genotype. Batch-correction was performed for each dataset using Harmony on experimental replicates. For clustering of all cells, 14 PCs were used for all datasets with a resolution of 1. For the glutamatergic dataset, 18 PCs were used with a resolution of 2. For the GABAergic dataset, 22 PCs were used with a resolution of 2.5.

Matching of clusters between datasets was performed based on unrelated WT marker expression identified using the Seurat command: FindAllMarkers(…, only.pos = TRUE, min.pct = 0.25, logfc.threshold = 0.75). For each cluster, we visualized the best marker genes on the UMAP. Upon confirmation that they were descriptive of the cluster (or additional clusters in the case of over clustering), we visualized the expression in the other datasets and assigned all clusters a name based on the expression of a single marker or multiple markers. In the case where no markers could be verified for a cluster, we looked for the nearby clusters on the UMAP, and merged the cluster with another cluster that showed the fewest distinguishing markers.

In some cases, there were clusters we could not assign to all datasets. For the glutamatergic datasets these were: *tubb5/chd4a, nhlh2/zic2a, pvalb6, pax6a, dlx5a, bhlhe23/grm2b, cabp5b, foxb1a/tfap2e*. For the GABAergic datasets these were: *aldocab, BX088, adarb2, CR34551, emx2, gata2a/nxph1, gyg1b, irx1b, crhbp, otp, onecut1, penkb, pnocb, txn/nr4a2a, uncx4.1*. These differences were often the result of the larger cell numbers in the unrelated WT dataset, allowing for the clustering algorithm to more accurately separate cell types. In the cases where a cell type is a rare one, the

cells were often merged with larger nearby clusters. Some notable exceptions: *adarb2* cells clustered strongly in lakritz and its siblings, but not in the larger unrelated WT dataset. *Otp*, neurons which are mainly glutamatergic appeared in the unrelated WT GABAergic dataset. *Uncx4.1* formed a significant cluster in the unrelated WT dataset, but was largely missing from the *lakritz* and sibling datasets, as was the same for the glutamatergic *dlx5a* cluster. However, there was never a case where cells expressing these markers were completely missing from all other datasets. For closer inspection, see Supplementary Fig. S6, S8.

Determination of confounding batch-effects

WT and *lakritz* datasets were first analyzed independently. To determine whether there are batch effects within each dataset across experiments, each dataset was first analyzed independently. Normalization, variable feature selection, scaling, and principal component analysis were performed with either WT or *lakritz* cells. The DimPlot visualization tools (Seurat) was used to inspect different combinations of the first 5 PCs where cells were color-coded by experiment. For the WT dataset, where we could detect a batch effect, we tested whether batch-correction using Harmony can correct the effect. We did not test this for the *lakritz* dataset, as we could not detect such a batch effect. To test whether WT and *lakritz* datasets were integrated properly also in absence of batch-correction, we followed standard procedure, but omitted a batch-correction step.

To test whether *lakritz* cluster structure is preserved following integration with WT cells, we first clustered the *lakritz* data independently and saved the results as a large matrix containing cell and "*lakritz*-only" cluster identities. We then followed standard procedure to cluster WT and *lakritz* cells together. Lastly, we plotted *lakritz* cells according to their new UMAP coordinates (embedded with WT) and color-coded the cells according to their "*lakritz*-only" clusters.

Analysis of globally differentially expressed genes between genotypes

For analysis of genes differentially expressed globally between genotypes, datasets were batch-corrected using Seurat's anchoring pipeline as described before. We used the FindMarkers command with default parameters comparing either unrelated WT with *lakritz* or unrelated WT with sibling WT. Globally differentially expressed genes were filtered using adjusted p-values < 0.05. We highlighted genes for which the log fold change was either larger than 1 or lower than -1, or where the ratio of expressing cells across groups was either higher than 2 or lower than 0.5. We also ran this analysis using Harmony for batch-correction, but this led to no differentially expressed genes.

Alignment of PCs between genotypes

Datasets were batch-corrected using Seurat's anchoring pipeline as described before. PCs were calculated using the Seurat RunPCA function for the first 15 PCs. For each comparison, only genes that appeared across PCs in both groups were compared. PCs were compared by calculating the cosine similarity between all PCs. PCs were compared be calculating the cosine similarity and extracting the highest value for each PC generating a distribution of values for most similar PCs. We generated our null distribution, by randomly dividing our WT dataset into two groups similar in ratio to *lakritz* / unrelated WT. We repeated this ten times to generate a null distribution for PC similarity across different datasets. We used a Wilcoxon signed-rank test to compare the different distribution.

Comparing neighborhood embedding between genotypes

For each cell, the neighborhood was determined by extracting the 19 nearest neighbors' genotype composition and calculating the cosine distance between the observed neighborhood genotype composition (unrelated WT / sibling WT / *lakritz*) and the expected neighborhood genotype composition. The null distribution was generated by randomly shuffling the genotype labels and recalculating the distance between observed and expected neighborhood genotype composition. From this, we extracted the standard deviation and calculated a Z-score for each cell in the original dataset. We color-coded our UMAP based on the Z-score to uncover areas in the UMAP (potentially clusters) that show significantly altered neighborhoods in *lakritz*.

In-silico cell type ablation analysis

Cell identity and cluster information were imported from a completed clustering analysis. Iteratively, *lakritz* cells belonging to a single cluster were removed from the count matrix. After removal of the cells, the matrix was processed following the standard Seurat pipeline without batch-correction. To determine whether we could observe an area missing in *lakritz* cells, we visualized all three genotypes together in UMAP space. Upon determination of a missing area, we opened an image where cells in the cluster missing *lakritz* cells were highlighted. This either confirmed our suspicion (if the cell type were missing, we could detect the effect), or denied it. We applied this analysis separately to datasets containing glutamatergic or GABAergic neurons for all clusters.

Finding clusters with altered transcriptomes in absence of RGCs

The analysis was applied as reported in (Sharma et al. Nature 2020) on datasets batch-corrected using Seurat's anchoring points pipeline as reported before. In short, for each

cluster we measured the ratio between WT and *lakritz* for the set of marker genes and for all genes detected in the cluster. We then compared these two distributions using a Wilcoxon signed-rank test and corrected the p-values for multiple testing (depending on the number of clusters) using Bonferroni's correction. We extracted the WT identities belonging to the cluster and projected them into the formerly analyzed WT dataset in which we performed our marker analysis. We color-coded the cells based on their p-value for transcriptome alteration. From this, we extracted a list of markers expressed in clusters predicted to be the most altered in *lakritz* compared to WT. We performed this analysis also by applying the same processing pipeline starting with only glutamatergic or GABAergic cells, essentially applying normalization within the separated dataset. We also tested this pipeline with Harmony batch-correction.

Comparing cell type proportions

For each experiment, we calculated the relative proportion of cells in each cluster as a fraction of the total number of cells captured across replicates in a single experiment. We compared the distribution of proportion of each cluster between WT and *lakritz* using a Wilcoxon signed-rank test and corrected the p-values for multiple testing using Bonferroni's correction based on the number of clusters.

Developmental trajectory analysis

The identities of cells belonging to the neural precursor cluster (expressing *ngn1* and *ascl1b*) were extracted and highlighted. Marker genes for this cluster were calculated using the FindMarkers Seurat function. Each gene was visually inspected using the FeaturePlot visualization tool to determine whether it is expressed only in the cluster (transient expression), or also in other clusters (continuous expression). For each gene we also summarized whether it is expressed in other cell classes (GABAergic neurons, glutamatergic neurons, habenula neurons, and progenitors).

HCR fluorescent in-situ imaging

HCR stains were performed according to the manufacturer's instructions (Molecular Instruments) with no modifications. Larvae used for HCR stains were reared in PTU from 24 hours to 6 days. At 6 days, larvae were fixed following instructions. All larvae were Tg(HGn12C:GFP) and stained for a maximum of two transcripts labeled with either/both Alexa546 and Alexa647. All probes were purchased from the manufacturer. Imaging was performed on a commercial Zeiss confocal (LSM780)

For WT and *lakritz* comparative stains, larvae were generated from parental *Tg(HGn12C:GFP*; *isl2b:RFP*, *lakritz*+/-) crossed with Tg(*lakritz*+/-). *Lakritz* (no RGCs) phenotype was selected for, or against, using *isl2b:RFP* expression. Comparative stains

were performed only on siblings from the same clutch stained in-parallel for each gene. Imaging in *lakritz* and WT was performed using the same imaging parameters in both samples when comparing the expression of the same gene. Images were collected from a minimum of two replicates for each condition. For each individual either one or two tiles were imaged.

Morphological registration using ANTs

Brain registration using ANTs was performed as described before (Kunst et al. Neuron 2019). In short, a standard brain was generated using 12 HCR stained Tg(HGn12C:GFP) larvae using ANTs. All HCR stains were registered to the Tg(HGn12C:GFP) standard using Tg(HGn12C:GFP) background stain present in all HCR stacks. In case of drift during image acquisition, the transmitted-light channel was used to correct drift using the MultiStackReg ImageJ plugin. Once the Affine transformation was saved, it was applied to all other channels.

For comparative WT and *lakritz* stains, individual confocal stacks were registered first to the Tg(HGn12C:GFP) standard and then differences were inspected by eye. In the case where differences were detected, the original stacks were then inspected to validate differences.

Anatomical masks from the MapZeBrain atlas were registered to the *Tg(HGn12C:GFP)* standard using a bridging *Tg(HuC:H2b-RFP)* from the atlas. *Tg(HuC:H2b-RFP, HGn12C:GFP)* larvae were independently registered to the *Tg(HGn12C:GFP)* standard to generate an H2b-RFP stain in the standard morphological space. Transgenic fluorescence was successfully preserved in *Tg(HuC:H2b-RFP)* by modifying the HCR fixation protocol as reported in (Lovett-Barron et al., Nature Neuroscience 2020).

HCR image analysis

Individual confocal stacks of HCR stains were aligned to a standard *Tg(HGn12C:GFP)* expression pattern as described before. Color-coded depth projections were generated using an ImageJ script written by Kota Miura (miura@embl.de). For analysis of HCR signal, individual confocal stacks were binarized using the ImageJ “RenyiEntropy” algorithm applied to individual slices using each slice’s histogram. In slices where the algorithm produced sudden spikes in the total number of pixels compared to surrounding slices, pixels were instead binarized using the “maxEntropy” algorithm. This step improved the binarization process where the “RenyiEntropy” algorithm was sensitive to sudden changes in noise. However, both algorithms were sensitive to different kinds of noise which allowed to generate a combined but smooth binarization mask. The binarized pixels were then multiplied by the image intensity in each pixel. The morphological signal analysis was performed on the product of these images. Registered anatomical masks were generated as described before. For total expression analysis, the sum of the

intensity across the binarized pixels was calculated for each anatomical mask. This analysis was performed for replicates of the same stain and averaged across replicates. The final heat matrix was calculated by normalizing the expression of each gene to the highest value for the specific gene. For background normalization, the background was defined as the signal intensity in pixels inside the anatomical mask that are not HCR positive binarized pixels. For predictions of best matches between brain areas and gene expression, the background-normalized signal values were normalized in the heatmap to the highest value in each brain region.

Optogenetic stimulation and behavioral tracking

Optogenetic stimulation and behavioral tracking were performed as reported in (Wu et al. Neuron 2020). In short, after embedding and removal of agarose close to the eyes, an optic fiber was placed on top of the fish targeting the pretectum. In order to track the eyes while in dark, larvae were illuminated from the bottom using an 850 nm infrared LED. Light emitted by optogenetic illumination was filtered out by an IR filter in front of the recording camera (Thorlabs absorptive filter, ND = 1.0). For focal optogenetic activation with ChR2, a 50um optic fiber (AFS50/125Y, Thorlabs) was used to shine blue light (473 nm, 20-40 mW/mm2, Omicron Lighthub) onto the right or left pretectum. In each experiment, larvae were presented with a phase of stationary gratings, followed by moving gratings (40s), followed by stationary gratings, followed by blue light illumination (60s) and stationary gratings. For optogenetic activation of the *lakritz* pretectum Tg(*mitfa* (-/-), *isl2b*:GFP, Gal4s1026t, UAS:Chr2-mCherry lakritz (+/+)) larvae were used. For control, the same larvae lacking ChR2-mCherry expression were used. Larvae were identified as *lakritz* via *isl2b*:GFP expression.

For eye tracking, the angle of each eye was calculated relative to the body midline. During visual stimulation (gratings moving), the eyes of a fish almost exclusively saccade in a single direction intermittent with a smooth movement in direction of motion. When there is no visual stimulation (stationary gratings), the eyes of a fish will saccade in one direction and then in the opposite direction. Hence, saccades are a reliable readout of the optokinetic response (OKR). To calculate an OKR index the saccades in one direction were subtracted from the saccades in the other direction for each eye – producing higher values during OKR. The average of both eyes was taken as the OKR index.

Validation of *lakritz* genotype

For larvae, *lakritz* mutants were identified by extreme pigmentation. When nacre (*mitfa* -/-) or PTU treated larvae needed to be identified, either *Tg(Isl2b:RFP)*, or *Tg(Isl2b:GFP)* were used. In WT *Tg*(*isl2b*) drives expression in RGCs, trigeminal ganglia, and spinal neurons. In *lakritz*, there are no RGCs, but *Tg*(*isl2b*) still drives expression in the trigeminal ganglia and spinal neurons. In adults, fin clips were used to determine carriers of the *lakritz* mutation as described in (Kay et al. Neuron 2001). In short myTaq extract-

PCR kit (Bioline) was used to amplify a 300 bp fraction of the *ath5* gene containing the *lakritz* mutation using the following primers: *Fw_ccggaattacatcccaagaac, Rv_ ggccatgatgtagctcagag*. The amplified product was digested using the StuI restriction enzyme over night at 37°C. Carriers of the *lakritz* mutation would show three products on an agarose gel at: 300 bp, 200 bp, and 100 bp. WT fish show two products at: 200 bp and 100 bp.

RESULTS

Characterizing cell types in visual processing centers

Preparation of a single-cell suspension

To transcriptionally profile neurons in zebrafish visual processing centers, I first identified the transgenic line Tg(HGn12C:GFP) in the Kawakami library of transgenic zebrafish lines (www.ztrap.nig.ac.jp) as one in which neurons in most visual processing centers were labeled by a GFP reporter. After carefully imaging larvae from this transgenic line, I concluded that it labels the entire thalamus (homolog of the LGN, the largest retinorecipient area in mammals), the entire pretectum (homolog of the accessory optic system in mammals), a large percentage of the tectum (the largest retinorecipient area in zebrafish), and additional small areas not involved in visual processing (Figure 10a, left). Because the expression in this transgenic line labeled most cells involved in visual processing, I reasoned that it could be used to characterize almost the complete set of visual cell types in the brain without including too many unrelated cells. In the entire brain, these cells would account for a much lower fraction of all cells. In this transgenic line, however, they account for the majority.

Our ability to transcriptionally profile single-cells depends on their separation from one another into a single-cell suspension. Most animal cells, however, are never in a suspension (except for blood cells and spermatozoa). Cells are normally held together by the interactions of many membrane-bound and extracellular proteins. In the case of neurons, they are also held together by their synaptic partners in a mesh encompassing the entire brain. To process brain tissue into a single-cell suspension, I used a protocol established in (Kölsch Y et al. Neuron. 2021) involving protease digestion followed by fluorescent-activated cell-sorting (FACS). In this paradigm, the tissue is digested by the protease papain - cutting extracellular proteins in tissue and releasing cells into a suspension (Figure 10a). Cells can then be sorted into a tube by their fluorescence (whether they contain GFP, or not). While I was able to use this protocol to generate a single cell suspension, I had to modify the protocol to increase the quality of cells I collected by FACS (see experimental procedure). The protocol was initially optimized to process RGCs from the retina and needed further optimization to suit processing neurons labeled in the HGn12C transgenic line.

Using FACS, I could discriminate between debris, single-cells, and multiple cells still adhering to each other (Figure 10b). Using calcein blue, the FACS could sort for live cells, and using GFP, it could sort for neurons of interest (Figure 10b). Using my optimized protocol, I could process 110 larvae in a single experiment, and collect 200,000 live, single-cells using FACS. Concentrating the suspension using a centrifuge resulted in a 60ul suspension with a cell density of 500-1000 cells/ul. In this suspension, cells would maintain a >90% viability (live cell / (live cells + dead cells)) for 3 hours. At this point I was able to proceed to use the enriched cell-suspension for single-cell sequencing using the 10X Genomics system.

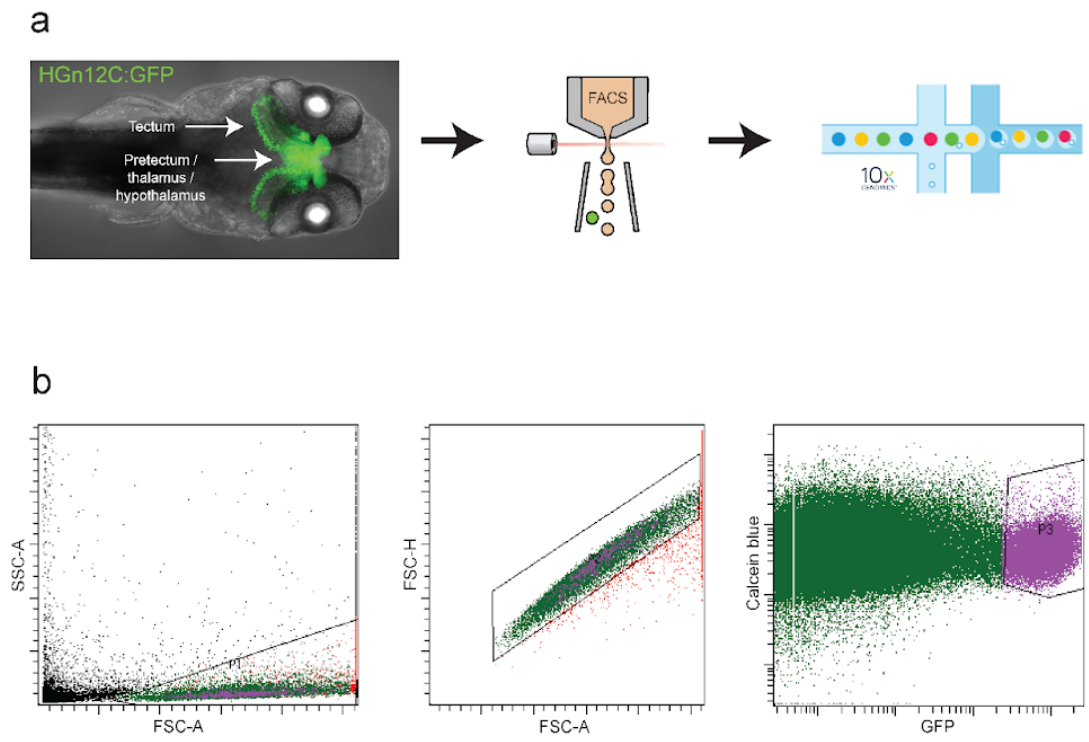


Figure 10: Generation of a single-cell suspension from a transgenic line

(a) Schematic workflow to generate a single-cell suspension from a transgenic line for transcriptional profiling using the 10X Genomics commercial system. Left, a maximum Z projection acquired on a confocal stack from 6 dpf larva carrying the transgene Tg(HGn12C:GFP). In green is GFP expression labeling the visual brain areas: tectum, pretectum, and thalamus. (b) Three panels from FACS session, sorting for single, live, GFP-labeled neurons from 6 dpf Tg(HGn12C:GFP) larvae. The gates set by FACS from left to right sort for subsets of cell population. Left shows gate P1 sorting for cells over debris. The gate is set by measuring the forward scattering over the side scattering. Side scattering is generated by sharp angles normally absent in intact cells. Middle shows gate P2 set to distinguish single-cells from multiples by measuring the height of the forward scatter curve over its area. The relationship should be a linear relationship between the height and area as cells change in size. This relationship is broken when multiple cells adhere to each other. Right shows get P3 set to distinguish GFP labeled cells and live cells over the rest. The upper right corner, where the gate is set, matches cells that show high GFP fluorescence and high Blue fluorescence (live cells). Gate P3 is a subset of gates P2 and P1 and contains 3% of all events detected by FACS.

Processing single-cells sequencing data

I submitted the genetic libraries generated on the 10x platform for illumina sequencing. Once sequencing was complete, I used the 10X Genomics pipeline to align resulting sequences to the zebrafish reference genome and to perform initial cell filtering based on the number of unique molecular identifiers (UMIs). For all experiments, the algorithm determined that 500 transcripts (UMIs) are sufficient to distinguish cells from other particles. In total, I was able to sequence 123,224 cells.

In order to understand how the sequenced cells organize into cell types, I continued to process my data using the Seurat R package. In the first part of my analysis I used additional filters to discriminate single-cells from unhealthy or stressed cells, or multiple cells sequenced together. I decided to identify cells as ones that contain between 400-8000 transcripts (UMIs), and if these arise from 200-4000 genes. To distinguish between live-healthy cells and dead or dying cells, I used the fraction of mitochondrial genes present in each cell (a common metric used in the field of single-cell sequencing). The reasoning behind this is that as cells are separated from their tissue environment they can rupture - causing the cytoplasm to leak. As mitochondria are large, they often remain inside the cell as they cannot leak through small ruptures as opposed to free RNA molecules. The result of this is that the fraction of mitochondrial transcripts increases over the fraction of cytosolic transcripts. However, as different tissues have different amounts of mitochondria, a threshold based on mitochondrial transcript percentage needs to be applied carefully. Healthy neurons maintain a mitochondrial transcript percentage at around 12.5%[112]. I thus discarded from further analysis cells where mitochondrial transcripts accounted for more than 12% of all transcripts. After applying all filters, I was left with a total of 95,122 transcriptionally profiled cells.

I next used the Seurat package to normalize the datasets. The ability to compare different cells in the dataset depends heavily on normalization. Cells can be larger or smaller, but still be the same overall cell type. This is why we cannot simply compare the absolute number of transcripts across cells, but they must first be normalized to the size of each cell. For convenience, the numbers are converted to logarithmic scale after normalization. The next step involves scaling of the expression across all cells. In essence, this step shifts the average expression across all cells to 0. This is done because the absolute number of transcripts in each cell is not highly informative; where a cell lies in relation to all other cells is incredibly informative. Shifting the average to 0, makes it much easier to understand which of all genes shows extreme expression in every cell (how far the cell's value is from the average value).

To understand which cell-types exist in the data, one would potentially need to compare all genes in all cells to all genes in every other cell. However, with a large number of cells, and a large number of genes, comparing everything to everything becomes computationally prohibitive. It is generally agreed upon that a few shortcuts can be taken in the number of comparisons needed. For example, not all genes need to be compared. Many genes show very little variability between cells (the opposite of genes that are differentially expressed between cell types). For analysis of my data, I decided to use between 2000-3000 genes whose expression varies the most between cells in order to separate cells into cell types (zebrafish neurons express ~25,000 genes; Figure

11a). After reducing the number of genes used to infer cell types, I performed principal component analysis (PCA) on my data (Figure 11b,c). Finding principal components (PCs) in this data is similar to finding meta-features that separate cells into different groups. While it is not easy to understand what these meta-features are, I noticed that the first PC calculated from my data was separating cells roughly along the lines of neurons to progenitors (Figure 11b).

In the last step of the pipeline, I used the Seurat package to calculate for all cells which cells are most similar to them (nearest-neighbors). Based on this graph map of transcriptional similarity, it was possible to find highly interconnected hubs of cells (neighborhoods) and to draw boundaries between these and other neighborhoods - drawing boundaries for cell types in the data. In the final step, I embedded the data in 2D UMAP (uniform manifold approximation and projection) space (Figure 11d). This last step, while not statistically informative, allows the visualization of high-dimensional data using a low number of dimensions. For zebrafish transcriptionally-profiled cells, this means that instead of representing a cell by the expression of all genes (25,000 genes = 25,000 coordinates), they can be represented by a lower number of dimensions (x = UMAP1, y = UMAP2) and easily visualized.

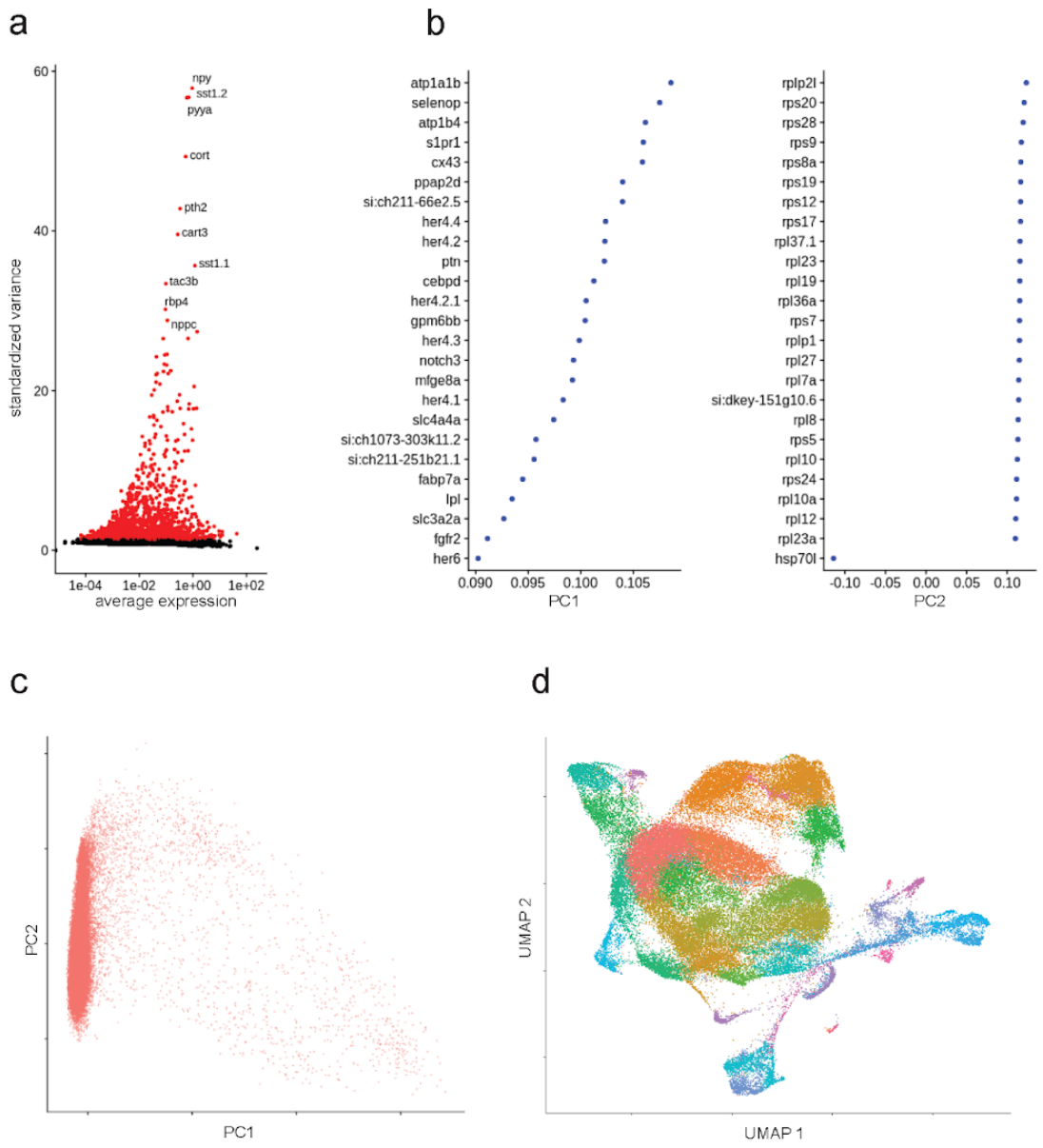


Figure 11: Initial processing and clustering of single-cell RNA-seq data

(a) Identifying the most variable genes across the sequenced cellular population. Genes are scores on their variability across cells. The most highly variable genes (HVGs) are used to construct principal components (PCs). (b) Constructing PCs from HVGs. Each PC is essentially a vector with a number of dimensions equal to the number of HVGs. Each cell than receives a score for each PC. This reduces the number of dimensions for each cell from the number of HVGs to number of PCs required to explain most of the variability between cells. Here are presented the top genes for PC1 (left) and PC2 (right), both explaining the highest share of the variance between cells. (c) Embedding of cells in PC space. Each cell has a score for each PC (a coordinate for each PC). Each red dot is a cell with a coordinate in a space composed of PC1 and PC2 (shown in b). (d) Embedding of cells in UMAP space (a low-dimensional

representation of PC space). Each dot is a cell with coordinates (UMAP1, UMAP2). Color-coding is the result of clustering of cell types (drawing boundaries between highly interconnected neighborhoods).

Transcriptionally profiling single-cells

To begin to understand what defines clusters (putative cell types), we can look at the genes expressed in each cluster and compare them to genes expressed in other clusters. In this way it is possible to identify genes that are differentially expressed in each cluster, and by extension across cell types. For this, I used the Seurat R package to iteratively run a Wilcoxon Rank Sum test between all clusters to identify cluster-specific differentially expressed genes. Once I completed running all tests, I was able to determine which genes distinctly mark each cluster (marker genes) from other clusters.

By initial visualization of the UMAP embedding of all cells and where certain genes are expressed, I was able to conclude that my data was clustering along super-clusters, or cell classes. In total, I could identify five cell classes: progenitors, and habenula-, GABAergic-, glutamatergic-, and *otp* neurons (figure 12). I identified a class of progenitors by the expression of known progenitor markers such as: *pcna*, *her4.1*, and *fabp7a* (figure 12e). The other classes separate from the progenitor class by the expression of *elavl3* and *elavl4*, both classical neuronal markers. Neurons originating from the habenula, a small distinct structure in the forebrain, express *gng8*, *kiss1*, *tac3a*, and *nwd2* (figure 12d). *Otp* neurons are defined by the expression of orthopedia homeobox transcription factors: *otpa* or *otpb* (figure 12a). The largest classes of neurons I identified were excitatory glutamatergic and inhibitory GABAergic neurons from visual processing centers (figure 12b,c). Glutamatergic neurons could easily be identified by the differential expression of *vglut2a* or *vglut2b* responsible for loading glutamate into vesicles for excitatory neurotransmission (figure 12b). GABAergic neurons could be identified, easily as well, by the expression of either gad1b or *gad2*, enzymes involved in GABA synthesis critical for inhibitory neurotransmission (figure 12c).

One common problem with clustering all cells together is that cell type identity is obscured by the large differences between cell classes. For example dopaminergic and serotonergic cell types did not cluster independently when I clustered all cells together. Instead, they mixed together with other inhibitory cell types. This is probably because inhibitory cells are very similar when compared to habenula neurons and progenitors. To expose the true cell type heterogeneity in my dataset, I subsetted the glutamatergic and GABAergic neurons into individual datasets - separated from all other cell classes. When I reclustered these datasets independently, I was able to identify 40 glutamatergic (Figure 13) and 37 GABAergic (Figure 14) putative cell types. The list of marker genes drawn from these clusters was even larger as in a large number of cases a single marker gene was insufficient to distinguish a single cluster from all other clusters. In other cases, multiple markers could be used equally well in separating a cluster from the rest.

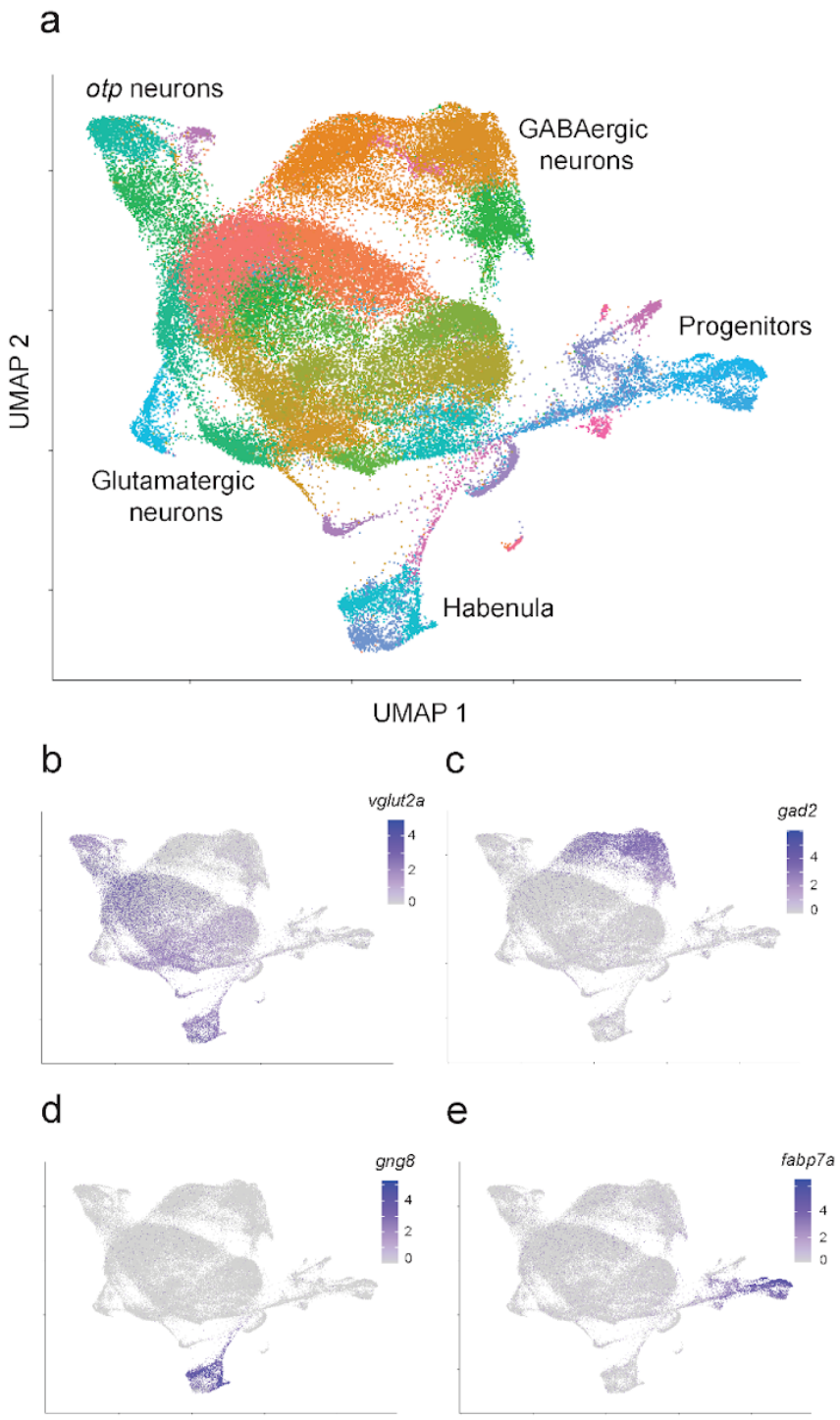


Figure 12: Identifying transcriptomic classes

(a) UMAP embedding of all sequenced cells. Color coding represents different clusters. Text labels adjacent cell classes. (b-e) Gene expression plots of cells embedded in UMAP space (vglut2a, glutamatergic neurons; gad2, GABAergic neurons; gng8, habenula neurons; fabp7a, progenitors). (e,f) Markers for glutamatergic (top, e) and GABAergic (bottom, f) clusters. Color shade represents the average level of marker expression in a cluster (average expression). Dot

size represents the percentage of cells expressing the marker in a cluster (percent expressing). UMAP space in all plots is the same as in (a).

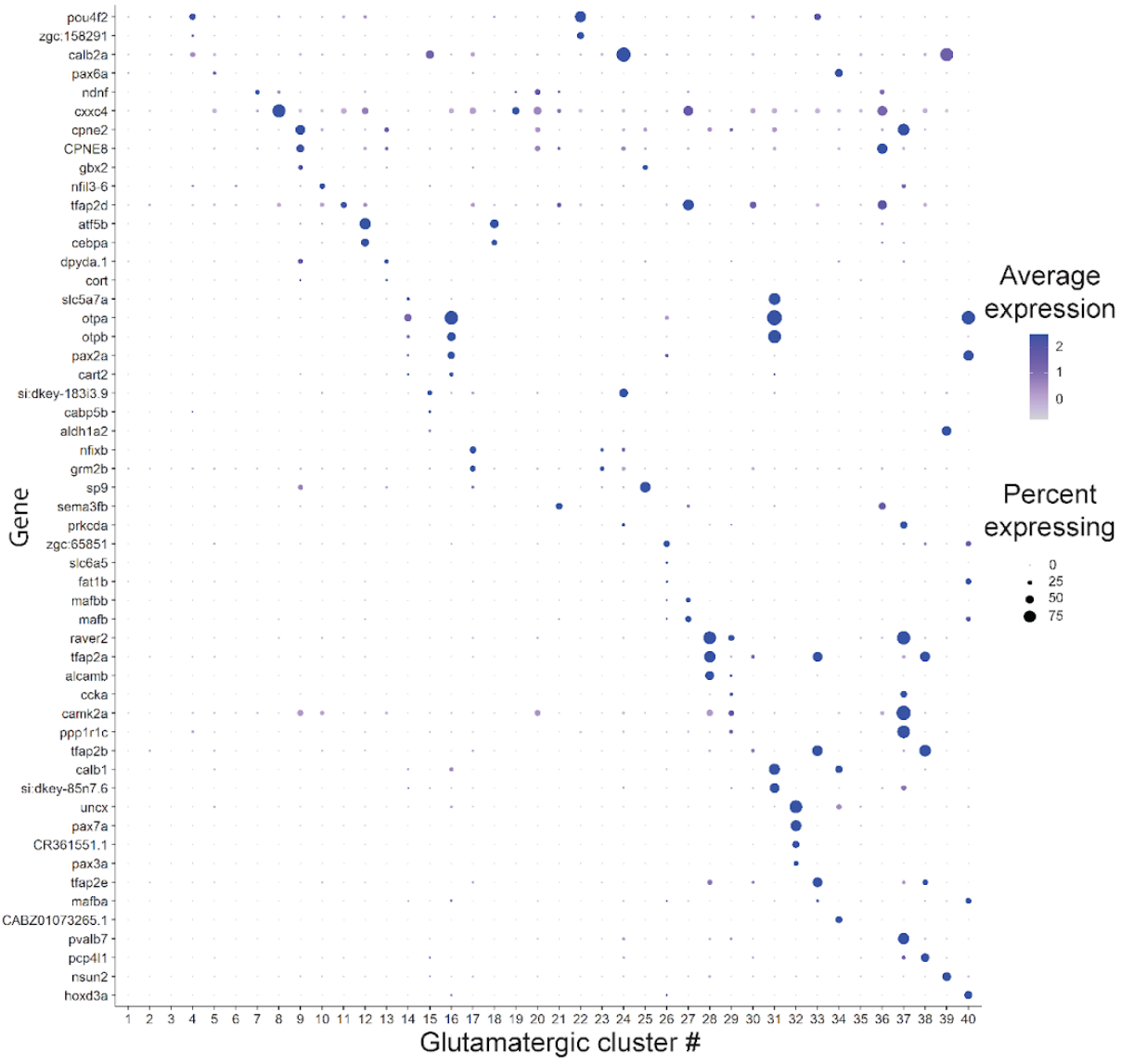


Figure 13: Identifying markers for glutamatergic clusters

Markers for glutamatergic clusters. Color shade represents the average level of marker expression in a cluster (average expression). Dot size represents the percentage of cells expressing the marker in a cluster (percent expressing). UMAP space is the same as in figure 12.

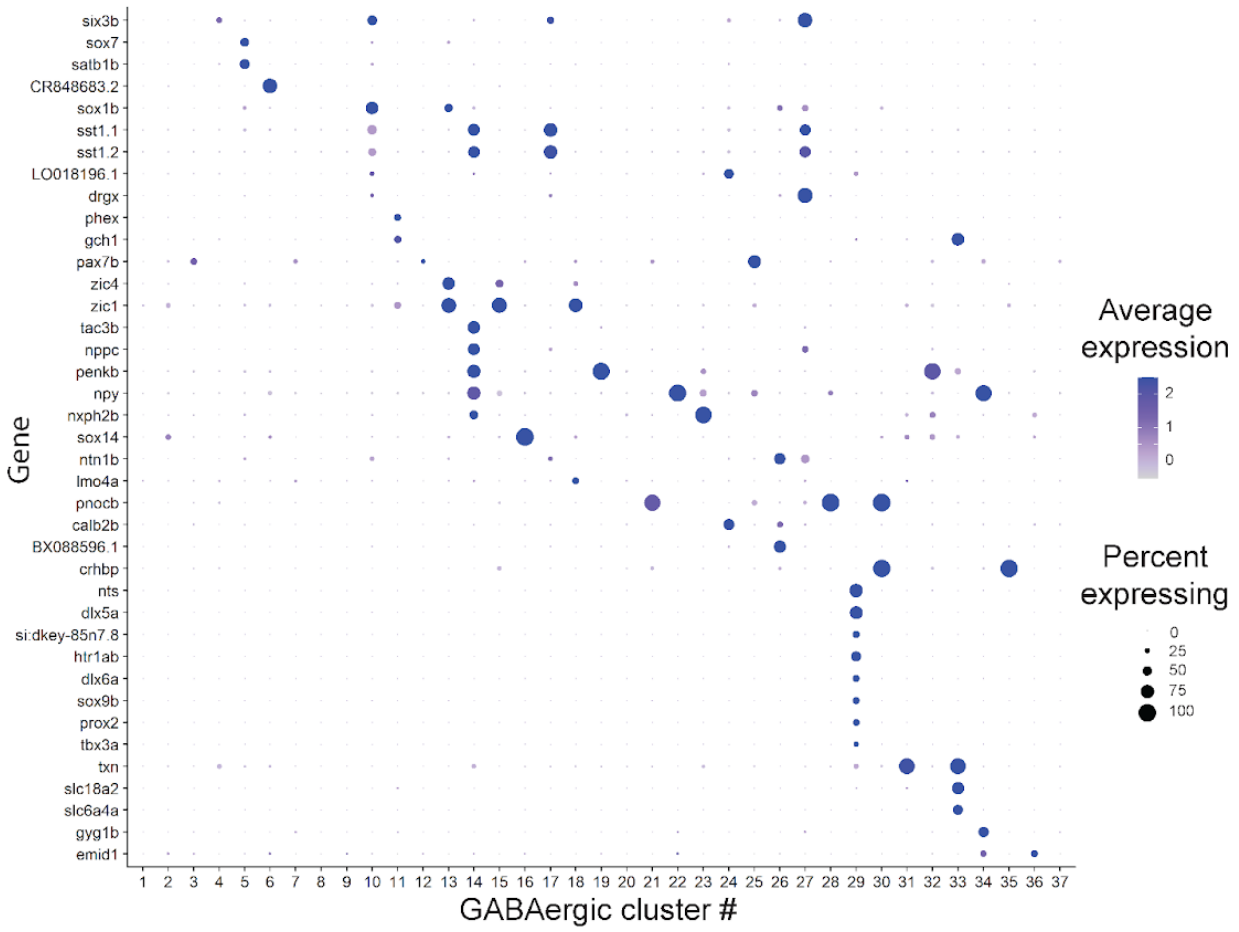


Figure 14: Identifying markers for GABAergic clusters

Markers for GABAergic clusters. Color shade represents the average level of marker expression in a cluster (average expression). Dot size represents the percentage of cells expressing the marker in a cluster (percent expressing). UMAP space is the same as in figure 12.

Uncovering the spatial organization of visual cell types

To understand which cell types make the zebrafish visual processing system, I transcriptionally profiled its neurons by single-cell RNA sequencing. However, in doing so, I had to dissociate cells from their native environment - losing all information pertaining to their 3D organization within the brain. Without knowing how these cell types are organized in tissue, it would be impossible to understand the cellular organization of the visual system. To overcome this, I began to reconstruct the 3D organization of cell types in visual centers using the list of GABAergic and glutamatergic markers I acquired by single-cell sequencing. In this way, I used these lists as a genetic access window to reconstruct the visual brain's cellular architecture.

To target specific markers, I used a modified version of fluorescent in-situ hybridization (FISH) termed hybridization chain reaction FISH (HCR-FISH). This allowed me to visualize the 3D tissue architecture of cell types in the brain. HCR-FISH targets many probes to bind to transcripts of the same gene. Pairs of probes are designed in a way that, only together, they form a binding site for read-out fluorescent probes - minimizing the chance that the wrong transcript would be read-out. Once read-out (amplification) sites are formed by binding of primary probe pairs, read-out fluorescent probes are flushed into the tissue. These will, in-turn, bind to the amplification site and begin a chain reaction, by which more fluorescent probes will be recruited to the amplification sites. As each transcript will bind multiple pairs, and each pair will recruit multiple fluorescent read-out probes, each transcript generates a very bright fluorescent dot. Depending on the number of transcripts, a cell can present many or few fluorescent dots for each gene. And, as multiple fluorescent dyes can be differentiated from one another, the expression of multiple genes can be visualized simultaneously in the same cell.

Using HCR in-situ, I was able to spatially characterize the distribution of all cell types I identified using single-cell sequencing. What became immediately clear was that most cell type markers labeled many cells across different parts of the brain, including non-visual areas. However, this was expected as other studies have shown that it is the expression of a number of genes that defines most cell types. But, within the context of a specific brain area, usually a single marker gene could be used to identify a cell type. The tectum is the largest retinorecipient area in zebrafish, but it was under-represented in the HGn12C transgenic line used for single cell sequencing. Despite this, many of the marker genes I imaged also labeled different cell populations in the tectum. Most tectal patterns often spanned the entire length of the tectum, but diverged mainly along the tectal layers. *Npy* for example, had sparse labeling along the entire tectal cell body layer, while the gap junction protein *gjd2b* sparsely labeled only the neuropil region of the tectum (figure 15k). The transcription factor *sox1b* mainly labeled cells along the tectal ventricle (https://www.biorxiv.org/content/10.1101/2021.11.15.468630v1, Supplementary Figure 2). *Onecut1* and *ccka* both label specific layers further away from the ventricle. *Pax7a/b*, *pou4f2*, *calb2a/b* labeled cells spanning almost the entire tectal cell body layer, with the exception of *calb2b*, which also labels cells in the tectal neuropil (figure 15k).

In the pretectum there was no clear axis along which cell types could be separated. Pretectal cell types usually showed a more diffuse pattern when compared to other areas. A single gene in which I could detect an axis was *zic1* (figure 15h), which labeled the entire anterior half of the pretectum. One anatomical difference that did differentiate pretectal cell types was that some markers labeled cell types around the pretectal neuropil, but not others. The genes *calb2a*, *foxb1a*, *cabp5b*, and *aldh1a2* label cells around the AF7 neuropil (figure 15g,h). *Npy* and *grm2b* were specific markers of pretectal cell types, but did not label any cells around the pretectal neuropil (figure 15i). *Pax* expression appeared to broadly label pretectal cell types (figure 15f).

The most interesting results appeared in the thalamus (homolog of the major retinorecipient area in mammals). A study characterizing thalamic cells in mice did not find the existence of cell types[113]. Instead, it reported that gene expression gradients define cell types along their thalamic axis. What I discovered shows that this is not true for zebrafish. Not only was I able to uncover the existence of thalamic cell types, but I also observed that cell types in the thalamus showed some of the most specific expression patterns in the visual system; thalamic cell types formed in many cases nuclei. *Cort*, *crhbp*, *crhb*, *sst1.2*, and *cckb* all showed expression in small thalamic nuclei (figure 15b-d). I was also able to confirm the expression of *pth2* in a thalamic cluster, expanding on earlier published work[114] (figure 15b). The transcription factor *sp9* was more broadly expressed in the thalamus, but not expressed in other visual areas - marking the boundary between the pretectum and thalamus (figure 15f).

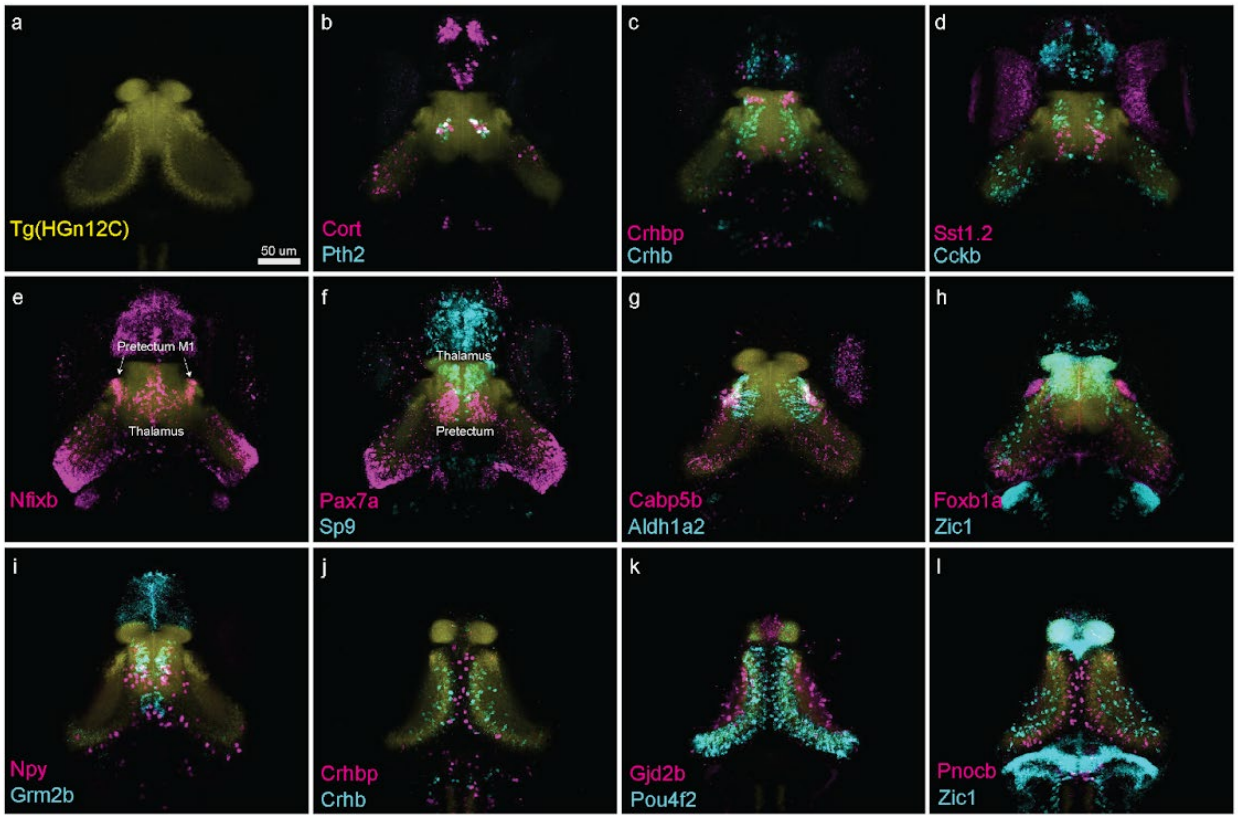


Figure 15. HCR-FISH uncovers the spatio-molecular organization of visual-processing centers

Marker genes validated by HCR in-situ labeling cells in discrete visual-processing areas. (a-i) Sub-stack maximum Z projections of registered HCR-FISH stains. In yellow, Tg(HGn12C:GFP) background stain used for registration. (b-f) selected thalamic markers. From left to right: cort, *pth2, crhbp, crhb, sst1.2, cckb, sp9*. (e-i) selected pretectal markers. From left to right: *nfixb*, *pax7a, cabp5b, aldh1a2, foxb1a, zic1, npy, grm2b*. (j-l) selected tectal markers. From left to right: *crhbp, crhb, gjd2b, pou4f2, pnocb, zic1*.

Building a gene expression atlas

Visualizing the expression of single marker genes is informative for the superficial characterization of tissue. It is also sufficient to validate the expression of marker genes in the brain and their location. However, transcriptomic clusters are usually defined by the expression of multiple genes. And, even specific clusters can be composed of multiple subtypes sharing the expression of some genes, but differentially expressing others. That is why it is critical to visualize the expression of multiple genes in the same tissue. HCR in-situ hypothetically allows to visualize the expression of up to five genes simultaneously. In my single-cell dataset I identified tens of marker genes. Using HCR in-situ to visualize all combinations of gene expression patterns would be prohibitively time consuming and expensive. One way in which this can be circumvented, is by using similar patterns across stains to align all images into a single space - a gene expression atlas.

To align all stains into a single gene expression atlas, I decided to create an average template of the Tg(HGn12C) transgenic line. I used the advanced normalization tools (ANTs) package to calculate the average expression pattern of the transgene from a total of 12 stains. This pattern describes the general recurring pattern in individual larvae and offers a general template to which their individual HGn12C pattern can be aligned. ANTs will calculate for each larvae the mathematical transformation needed to morph it into the shared atlas place. Once this is calculated, it can also be applied to other stains in the same animal. In this way, I was able to use ANTs to create a gene expression atlas of >70 genes across hundreds of animals. Using this atlas I was then able to visualize any combination of gene expression patterns in a putative larval zebrafish brain (https://www.biorxiv.org/content/10.1101/2021.11.15.468630v1, Supplementary Figure 2).

Using a gene expression atlas to understand tissue architecture

An HCR atlas containing many different gene expression patterns is a powerful tool to understand tissue architecture. But, to truly use it to understand the structure of the brain requires an anatomical component - defining which pixels belong to which brain area. I included anatomical brain annotation in my HCR atlas using brain area annotations from the mapzebrain atlas[61]. Using a bridging stain (a panneuronal cell body stain), I was able to use ANTs to transform all brain area annotations from their mapzebrain space into my HCR atlas space.

Having individual pixels associated with specific brain areas allowed me to assign gene expression values to each area, and to computationally compare the expression between different genes and different areas. The task of deciding for each pixel if it indeed contains information is a tricky one. Confocal images always contain noise and the amount of noise changes based on: tissue, imaging depth, and detection color. It is therefore wrong to simply sum all pixel values into an expression score within an anatomically defined area. Doing this for the tectum and retina (the largest anatomical areas), for example, will always result in a high expression score. Even small amounts of noise will sum into large values over large areas.

I designed a comprehensive pipeline using the imageJ software and R to extract expression information from my HCR atlas for individual gene expression patterns and then to extract comparative expression statistics for all stains for all annotated brain areas. In this pipeline, pixels containing signal were excluded from background for every slice in every confocal stack - binarizing pixels across a whole image stack. While this created a number of false positives, I reasoned that by multiplying the binarized pixels with the original signal, most of the final values will arise from pixels containing biological signal (for more information see experimental procedure).

After processing all of the atlas stacks, I decided to look at three matrices that I believed are informative to understand the relationships between gene expression profiles and brain areas. I first calculated the total intensity in the processed data that appeared in each brain area. This meant that values summed from pixels that fall within a certain anatomical area. For each gene I then normalized all expression values to the highest value of that particular gene. I found this metric informative to understand where a gene is most abundantly expressed and whether it is differentially expressed in certain areas. For a second metric I decided to normalize the intensity to the background intensity within each area, before normalizing across a gene. This metric offers a more conservative estimate of intensity and takes into account differences between samples. Third, I normalized the earlier data across brain areas, instead of the specific gene expression. I reasoned that this offers an estimate of specificity and can guide researchers looking for candidates to experimentally target specific brain areas. Lastly, I summarized all metrics in a single table presenting for each gene the top areas in which it is expressed (https://www.biorxiv.org/content/10.1101/2021.11.15.468630v1, Supplementary Figure 3 and Supplementary Table 1).

Understanding the inner workings of the “blind” visual system

After characterizing the cell types and tissue architecture that make the zebrafish visual system, I decided to ask how these change in absence of retinal input. *Lakritz* mutants lack the expression of the *Ath5* transcription factor and so do not develop any RGCs[115] (Figure 16). This means that not only are *lakritz* mutants blind, but in these animals there is never retinal input to visual centers, functional or molecular. I reasoned that characterizing cells in the *lakritz* visual system should uncover if the retina does play any role in cell fate determination of visual neurons in the brain. I crossed transgenic Tg(HGn12C) fish with *lakritz* carriers to generate blind animals in which I could characterize visual cell types that never received retinal input. Because the *lakritz* mutation is recessive, I decided to use the heterozygote siblings (akin to WT) as a way to control for differences between the unrelated WT and blind animals that do not stem from the lack of RGCs. Altogether, I was able to transcriptionally profile 20,221 cells from *lakritz* mutants and 25,687 cells from its WT siblings. After processing the data in the same way I described earlier (unrelated WT), I was left with 17,029 *lakritz* cells and 18,443 WT siblings cells.

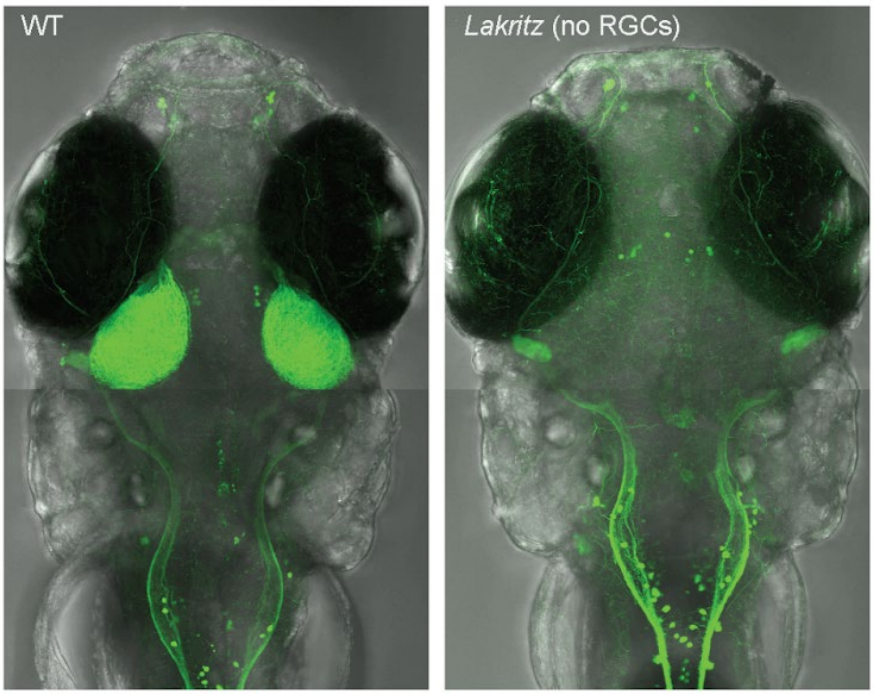


Figure 16. *Lakritz* mutant brains receive no retinal input because RGCs are absent

Images of WT (right) and *lakritz* (left) expressing the RGC labeling transgene Tg(Isl2b:GFP) in green. The transgene also drives expression in spinal neurons, trigeminal ganglia, and other rare populations. In WT all populations are present and show green expression. In *lakritz* there are no RGCs and so the RGC tectal arbors are completely missing. Other populations are unaffected and clearly visible in *lakritz*.

Integration of single-cell datasets and identification of cell type differences

While there is no consensus on how different single-cell datasets should be analyzed and compared, there exist many tools for integrating and comparing different samples[116,117]. I began my analysis by integrating all three samples (unrelated WT, WT sibling, and *lakritz*) and visualizing differences in 2D UMAP. I reasoned that this simple analysis could directly lead to identification of putative cell types that fail to develop without retinal input. A missing cell type would present as a cluster missing in the *lakritz* dataset, but not in either the unrelated or sibling WT datasets. Upon visualization of the resultant 2D UMAP, I did not observe any differences between the samples - all visual cell types that develop in WT larvae also develop in *lakritz* (Figure 17a). I continued to analyze the datasets for differences by bioinformatic separation of the glutamatergic and GABAergic cell types. In my earlier analysis of the WT dataset, it was only by bioinformatically separating GABAergic and glutamatergic cell types that I was able to uncover the full heterogeneity of visual cell types. This, however, did not uncover new differences: all cell types were present across all samples (Figure 17b,c).

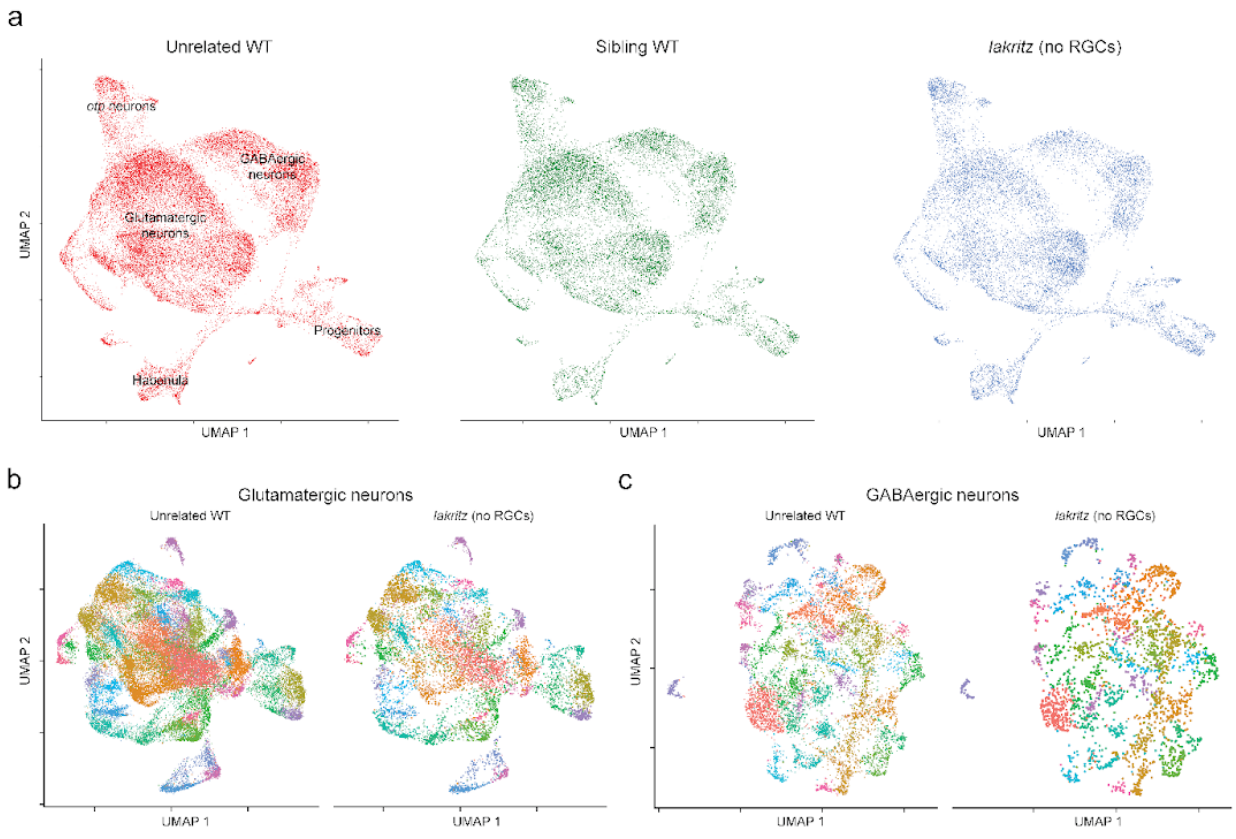


Figure 17. Forebrain cell-type diversity emerges in absence of retinal input

(a) Embedding of different genotypes in the same 2D UMAP space. Text labels adjacent cell classes. (b,c) Clustering of glutamatergic (left) and GABAergic cells (right). Presented side-by-side are WT and *lakritz* cells of the same clusters.

I reasoned that one possibility the samples could appear similar while hiding a difference, is the large dissimilarity in the number of cells between samples. This could theoretically result in a smaller sample being forced into a space dominated by another sample - forcing it to look the same in the process. Batch effects and their correction could also affect the successful integration of the different datasets; across replicates they can introduce variability that obscures biological effects. Using batch correction while integrating could also obscure biological differences. To address these, I began by looking for batch effects within replicates of the same datasets. While the *lakritz* datasets presented no batch effects across replicates, the WT dataset did (Figure 18a,b). I observed, however, that batch correction resolved the small effect present among WT replicates (Figure 18a). I next decided to ask if without batch correction, *lakritz* and WT datasets would show differences. If yes, there could either be a technical batch effect between the datasets, or there could be a biological difference between datasets obscured by batch correction. Fortunately, even without applying batch-correction, I could not see a difference between samples (Figure 18c). To exclude the possibility of the larger WT dataset “forcing” the smaller *lakritz* dataset to look more similar, I clustered the *lakritz* cells independently and plotted these clusters in the space integrating both datasets. I could clearly see that even when *lakritz* was integrated with all WT cells, the clusters were still preserved in this dataset (Figure 19). The similarity between datasets is thus likely the result of true, biological, similarity between datasets.

Visualizing multiple samples in UMAP space is an easy way to identify cell type differences between them. However, there exists difficulty in judging cell type differences based only on UMAP projection. Clusters do not always separate from other clusters neatly on a UMAP. Cells belonging to a single cluster may be hidden by cells belonging to other clusters. This is because it may not be possible to properly describe the high-dimensional distances between cells and clusters using a 2D UMAP projection. While most clusters can separate neatly on a UMAP, it is possible that clusters missing in *lakritz* are also the ones that appear diffuse (unclustered) in a 2D UMAP. In this case, we would wrongly conclude that all samples share the same clusters. I therefore decided to create a pipeline I termed “in-silico cell-type ablation” to see how severe this issue might be - for how many clusters, if they were missing, would I be able to conclude that they are missing on a 2D UMAP.

In-silico cell-type ablation involved iteratively deleting *lakritz* cells belonging to a single cluster from the original count matrix (similar to if they were never sequenced because they did not exist in a *lakritz* brain). After processing the datasets, I visualized the UMAP and identified where I thought an effect lay: where *lakritz* cells might be missing. After hypothesizing where an effect was present, I would uncover the true location of the missing cluster. I would then count it as a positive or negative run, depending on whether my hypothesis was correct or not (Figures 20, 21). Overall, I was able to identify missing clusters in 87% of the glutamatergic clusters and 71% of the GABAergic clusters. Together, these clusters accounted for 90% of the cells I sequenced. This suggested that if there was a difference between lakritz and WT that I missed, that this only involved a small fraction of all cells and represented a small effect for retinal input on cell fate determination. Upon inspection of clusters I could not identify their

absences, I realized that they were mostly dispersed throughout the UMAP, lacking clear clustering.

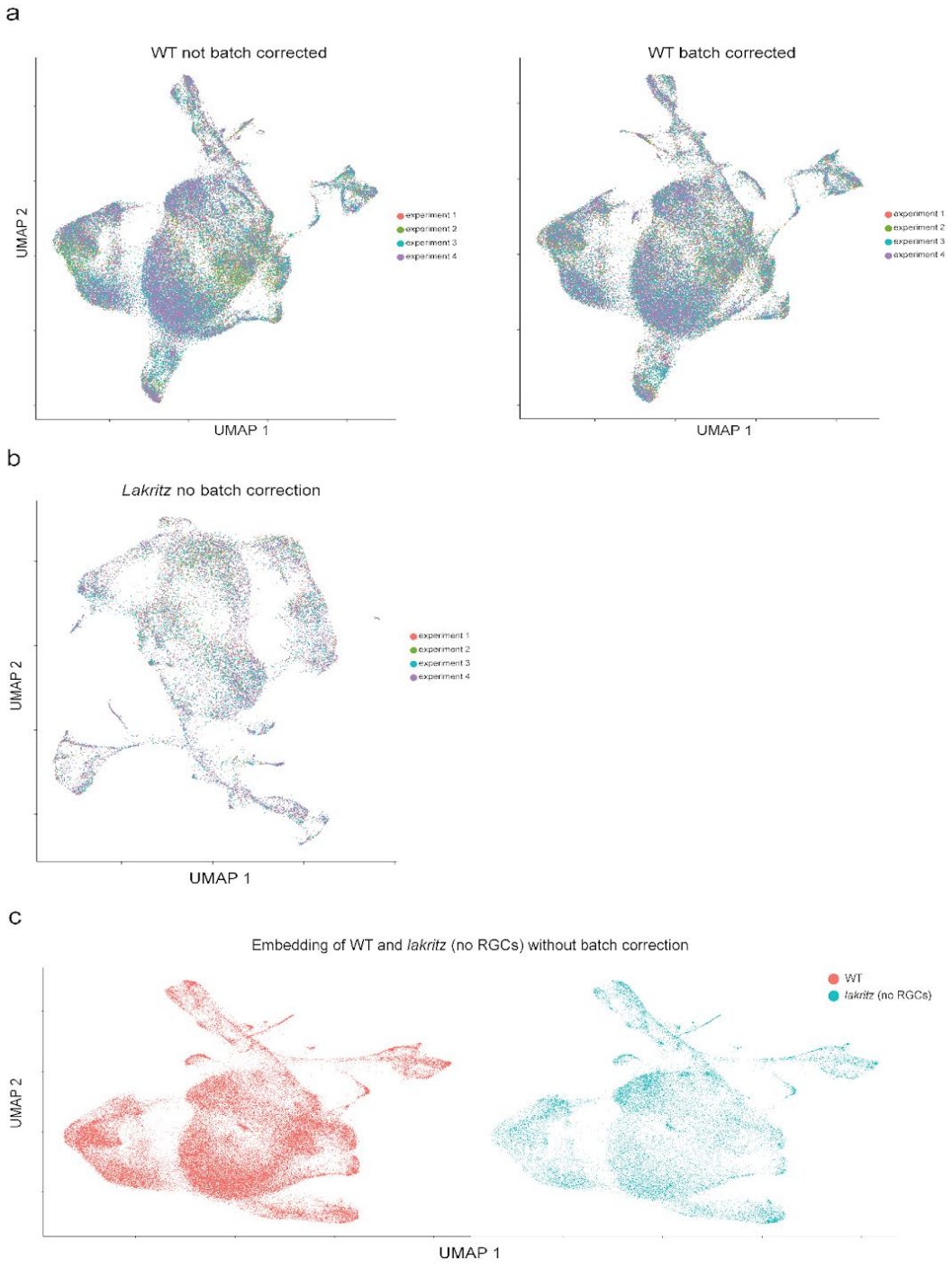


Figure 18. *Lakritz* (no RGCs) and WT cells can be fully integrated with no observable confounding effects

(a) UMAP embedding of all WT cells before (left) and after (right) batch correction. (b) UMAP embedding of all *lakritz* (no RGCs) cells without batch correction. (c) Embedding of WT and *lakritz* (no RGCs) cells in the same UMAP space without batch correction.

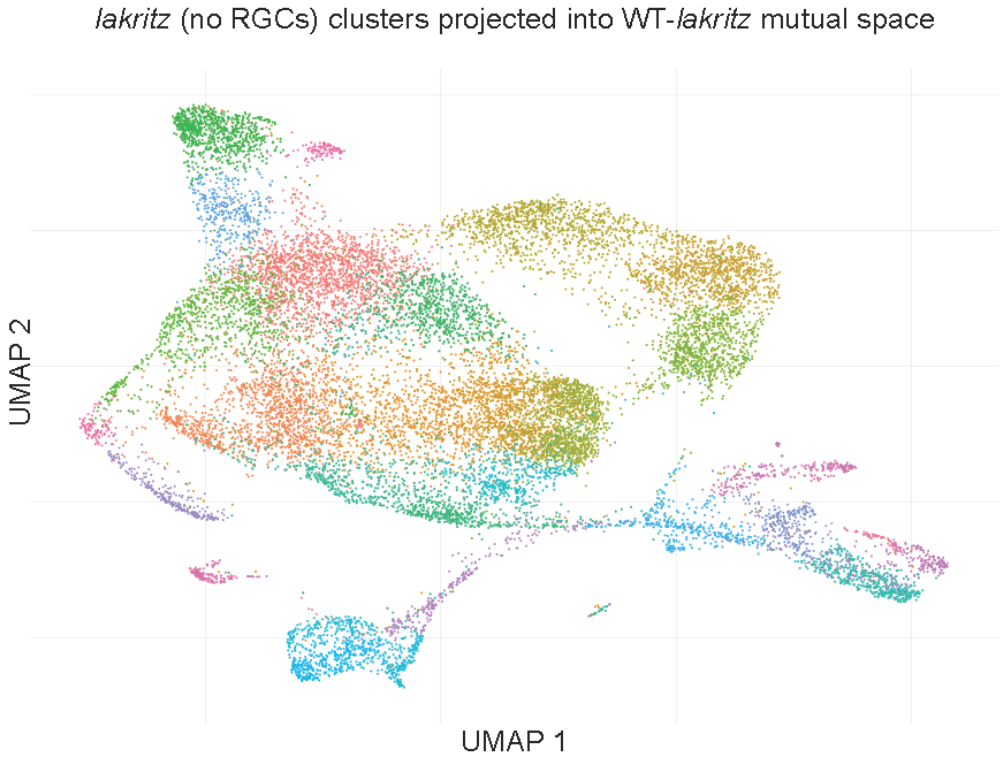


Figure 19. Larger WT population does not coerce lakritz (no RGCs) cell-type identities to resemble WT identities

Lakritz (no RGCs) cells embedded with all other cells. Color-coding is drawn from cluster identity following clustering of only *lakritz* (no RGCs) cells. Shown are only *lakritz* cells in WT-*lakritz* mutual UMAP space.

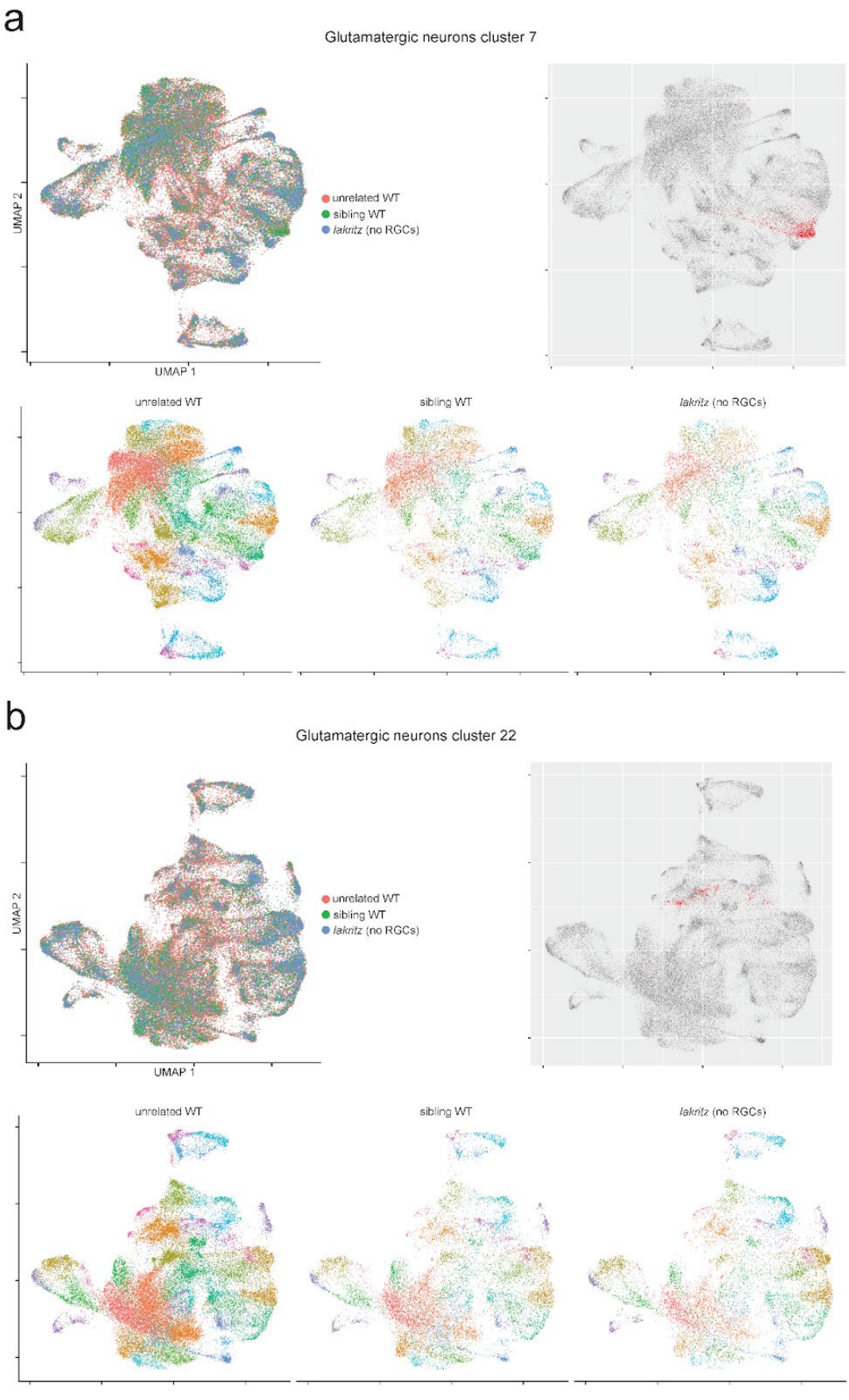

a
Glutamatergic neurons cluster 7
UMAP 2
UMAP 1
unrelated WT
sibling WT
lakritz (no RGCs)
unrelated WT
sibling WT
lakritz (no RGCs)
b
Glutamatergic neurons cluster 22
UMAP 2
UMAP 1
unrelated WT
sibling WT
lakritz (no RGCs)
unrelated WT
sibling WT
lakritz (no RGCs)

Figure 20. In-silico cell type ablation of a subset of glutamatergic neuronal clusters

A select number of glutamatergic clusters processed via an in-silico cell ablation pipeline. (a) Example of a glutamatergic cluster where ablation could be identified. (b) Example of a glutamatergic cluster where ablation could not be identified

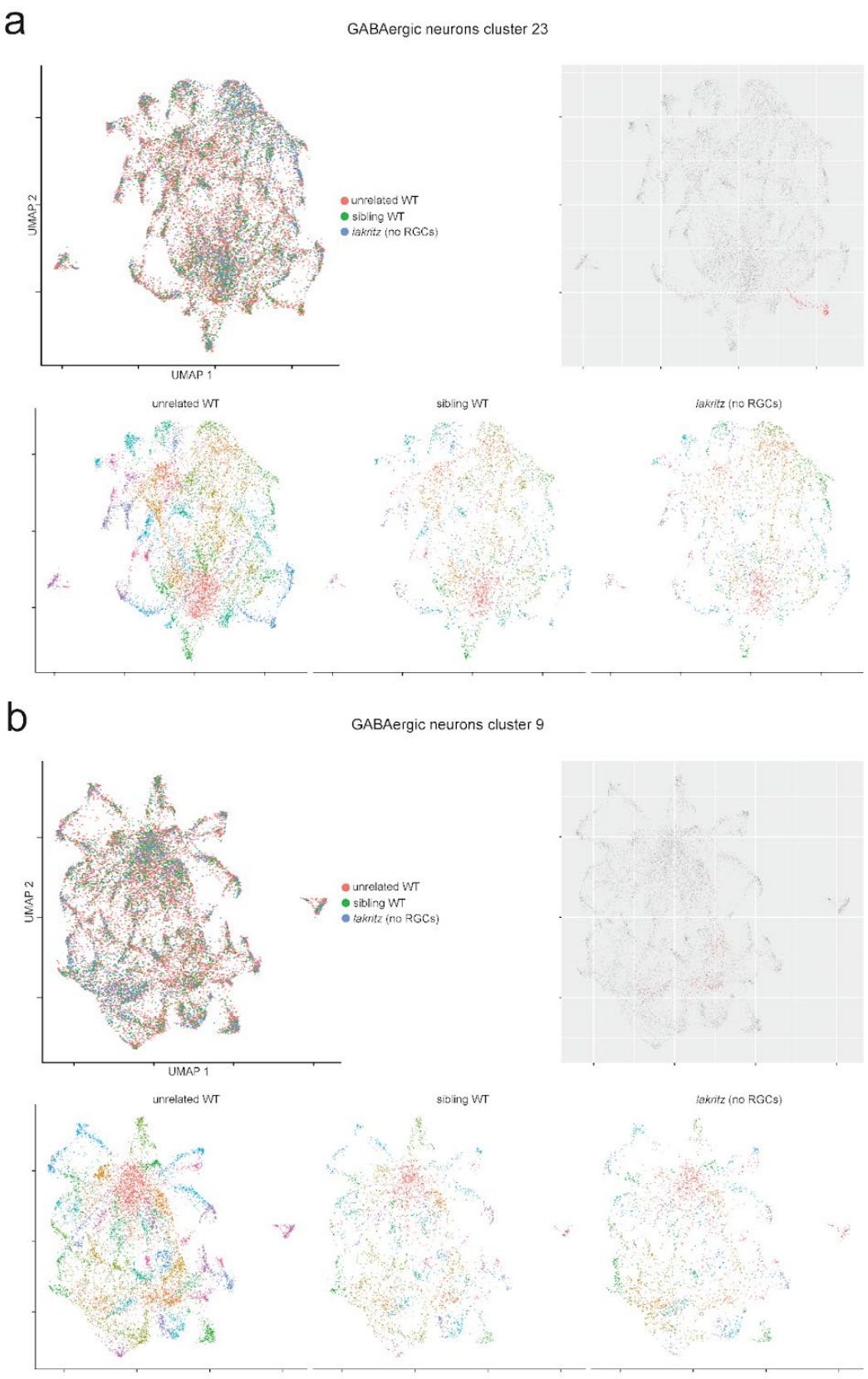

a
GABAergic neurons cluster 23
UMAP 2
UMAP 1
unrelated WT
sibling WT
lakritz (no RGCs)
unrelated WT
sibling WT
lakritz (no RGCs)
b
GABAergic neurons cluster 9
UMAP 2
UMAP 1
unrelated WT
sibling WT
lakritz (no RGCs)
unrelated WT
sibling WT
lakritz (no RGCs)

Figure 21. In-silico cell type ablation of a subset of GABAergic neuronal clusters

A select number of GABAergic clusters processed via an in-silico cell ablation pipeline. (a) Example of a GABAergic cluster where ablation could be identified. (b) Example of a GABAergic cluster where ablation could not be identified.

Lastly, I decided to cluster the three datasets independently. In this way it would be possible to compare clusters while completely avoiding confounding effects potentially resulting from UMAP embedding. Clustering datasets independently to compare cell types, however, is difficult. While there are a number of methods that can compare clusters, they have not been widely applied and still require human intervention[118,119]. Clusters depend heavily on pre-selected semi-arbitrary parameters. And, finding different parameters that lead to the exact same clusters across datasets may not actually be possible. I thus decided that I will manually align the clusters by inspecting the combinations of markers defining each cell type. Similar clusters should share similar marker “hotspots”. In the case where there are fewer cells from a particular cluster in one dataset, resulting in under-clustering, it would still be possible to identify these hotspots based on visual inspection of the markers across datasets.

To align clusters across datasets, I began by clustering all cells in each dataset (unrealted WT, sibling WT, and *lakritz*). I calculated for each dataset independently how many principal components (PCs) should be used and which clustering resolution should be applied. I then applied the highest values across all datasets. I found that visualization alone of the UMAP, in this case, was mostly sufficient to match groups of clusters across datasets. Upon finding marker genes for all clusters, I was able to manually annotate all clusters and align most of them across datasets (Figure 22). In cases where clusters did not appear across all datasets, visualization of marker gene expression patterns could confirm the presence of the cluster in all datasets. For example, some clusters in the *lakritz* and WT sibling datasets appeared merged with other clusters, while in the larger unrelated WT sample they appeared as two distinct clusters. I was thus satisfied that *lakritz* and WT indeed share the same cell type population - that retinal input was not necessary for cell fate determination of visual cell types (https://www.biorxiv.org/content/10.1101/2021.11.15.468630v1, Supplementary Figure 6-9).

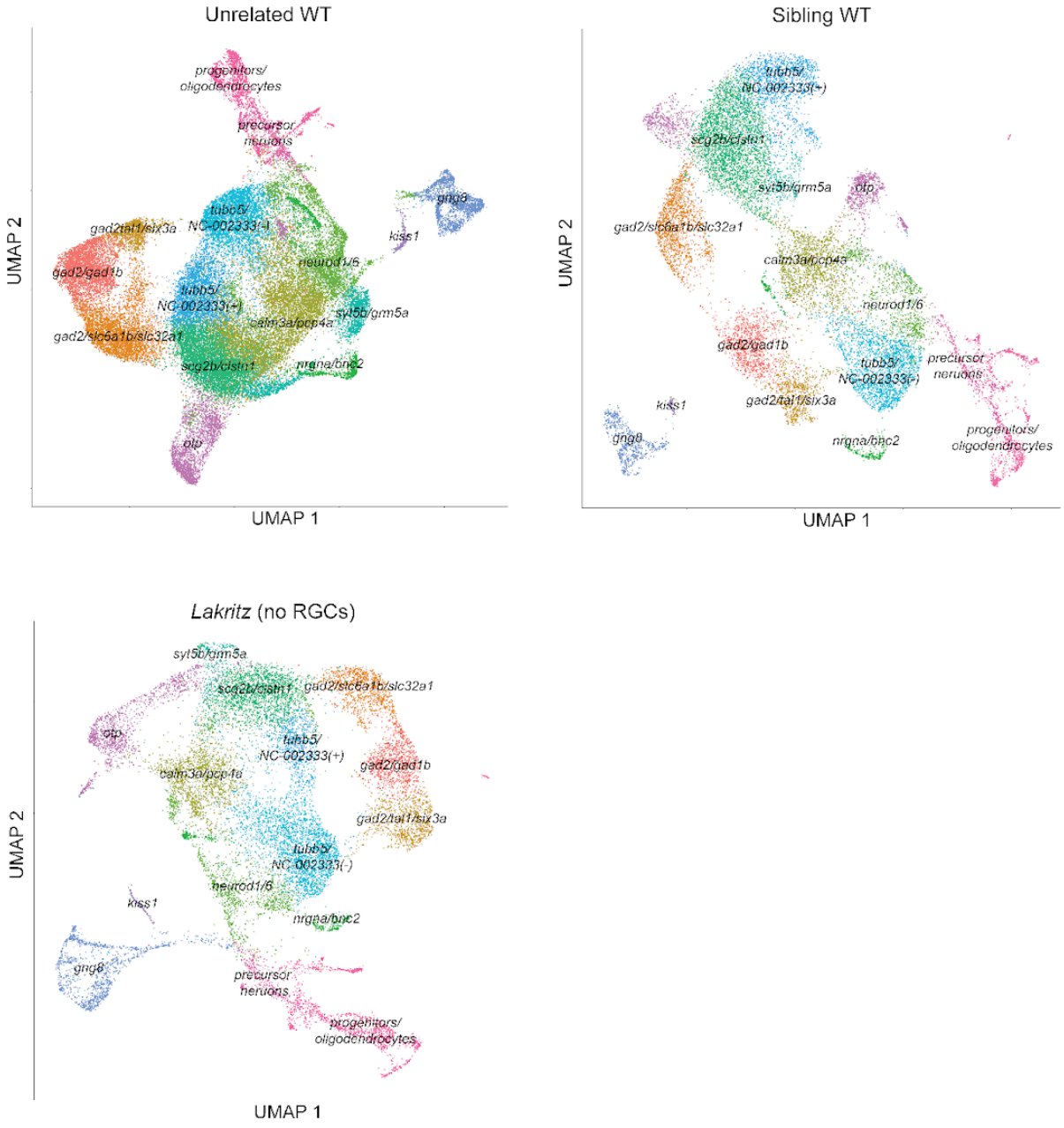


Figure 22. Independent clustering uncovers similar clusters across samples

UMAP embedding of three different samples (unrelated WT, sibling WT, *lakritz*) after independently processing and clustering each sample.

Identification of transcriptomic differences unrelated to cell fate determination

While I was able to show that *lakritz* mutants develop the full repertoire of WT visual cell types, it did not exclude the possibility that there are transcriptomic differences between them. Single genes between samples can show significant changes without largely affecting the composition of cell types; a cell type that's sufficiently different from other cell types can still be identified by the expression of a few marker genes, even if others are altered. It is also possible that while terminal selection of cell types is unaltered that earlier programs are - leading to differences in the abundance of cell types, without leading to change in the repertoire of cell types. I therefore decided to explore whether I could identify transcriptional changes in absence of RGCs that do not directly relate to cell fate determination.

I first asked whether I could identify global differences between samples. As differences between samples can also be technical or arise from genetic background, I performed a three-way comparison between samples to find differences caused by the absence of RGCs. A difference caused by the lack of RGCs should exist between *lakritz* and both WT samples and should be smaller or not exist between WT samples. I could find more globally differentially expressed genes between *lakritz* and WT than between the two WT samples. However, only a few of the genes dysregulated in *lakritz* passed a significance threshold reserved for marker genes, and in all cases they did not appear to have any clear implication in neuronal function (Figure 23a). I found that this small difference in global gene expression also extended to principal component (PC) space: while the most important PCs were present across all samples, in some cases their order was switched in *lakritz* (Figure 23b).

I next looked at the composition of local neighborhoods between WT and *lakritz* cells. The local neighborhood is composed of highly similar cells. If samples are largely similar then a cell's neighborhood should contain cells from all samples in proportions relative to their size. If not, there should be a bias for one sample over another. In this case, if *lakritz* cells were indeed different from WT cells, I would expect the *lakritz* neighborhoods to reflect this; I would expect that the neighborhoods of *lakritz* cells will have an unusually high number of other *lakritz* cells. For the two WT populations I would expect that the number of neighbors for each cell would simply be the ratio of unrelated WT to sibling WT cells. I would further expect that if only some *lakritz* cell types were very different from WT cells, then altered neighborhoods should mainly exist in UMAP hotspots where clusters of these cell types exist. In my data, I could not find such an effect (Figure 23c). There existed altered *lakritz* neighborhoods, but they were more distributed throughout the UMAP space. This suggests that effects are not in specific cell types, but rather present a global drift. Moreover, the magnitude of difference between *lakritz* and WT could also be found between both WT samples, although I detected fewer of those neighborhoods overall. These analyses in the end failed to uncover large and specific differences caused by the absence of RGCs in *lakritz*.

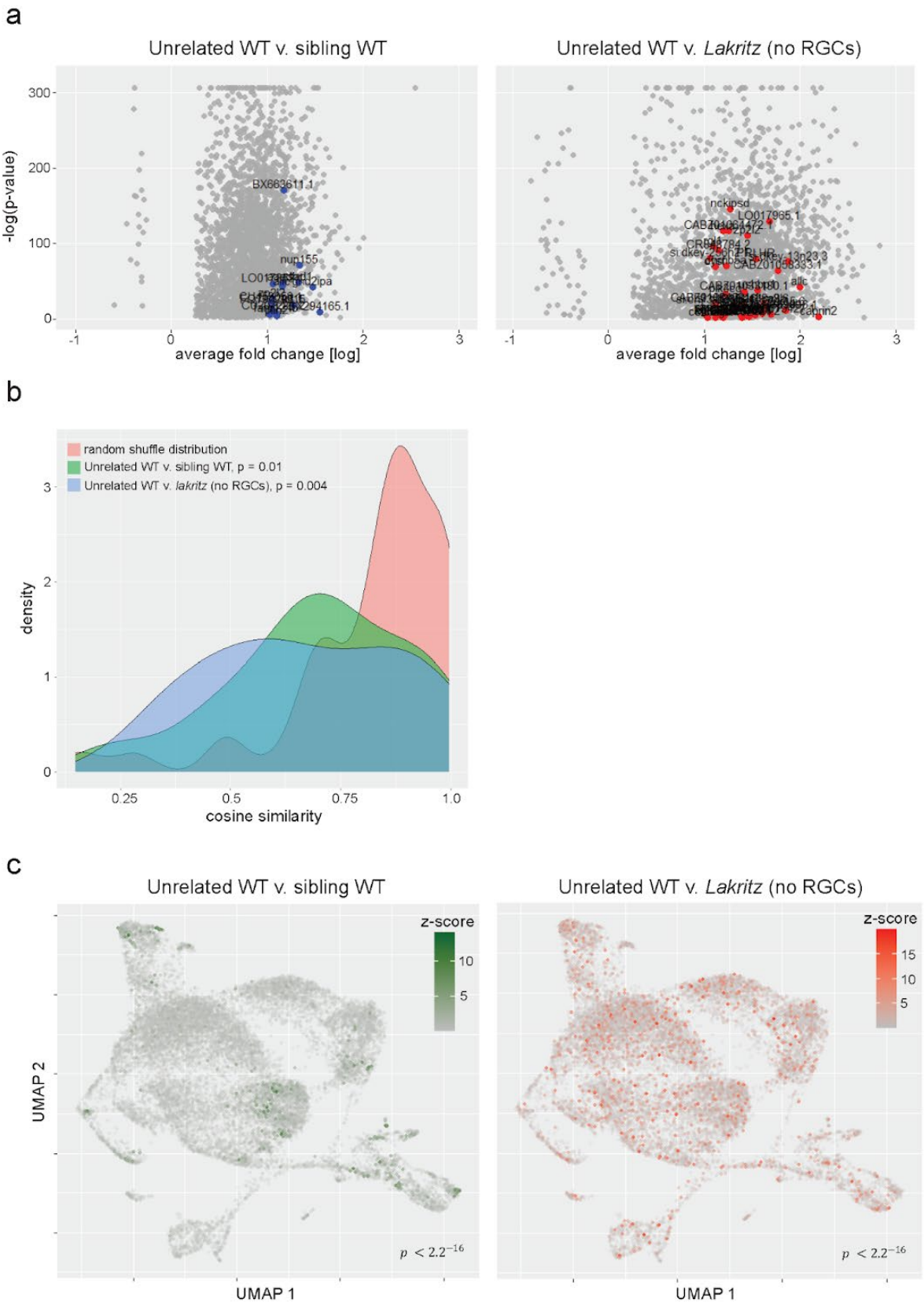

a
Unrelated WT v. sibling WT
Unrelated WT v. Lakritz (no RGCs)
-log(p-value)
average fold change [log]
BX663611.1
nup155
b
random shuffle distribution
Unrelated WT v. sibling WT, p = 0.01
Unrelated WT v. lakritz (no RGCs), p = 0.004
density
cosine similarity
c
Unrelated WT v. sibling WT
Unrelated WT v. Lakritz (no RGCs)
z-score
UMAP 2
UMAP 1
p < 2.2^−16

Figure 23. *Lakritz* (no RGCs) population shows a global transcriptomic drift from WT population

(a) Volcano plots of global markers (gray dots) detected between unrelated WT and WT (left) and *lakritz* (no RGCs) and unrelated WT (right). Colored dots with text represent marker genes differentially expressed both in their level and in the proportion of cells expressing the gene (blue, control; red, *lakritz*; no RGCs). (b) PC similarity across groups. Red, iterative random shuffle of control cells into two groups generating a null distribution. Green, unrelated WT compared with WT. Blue, unrelated WT compared with *lakritz* (no RGCs). P-values calculated using a wilcoxon signed-rank test. (c) UMAP embedding of all cells. Color-coding shows cells with altered nearest-neighbor neighborhoods. Color intensity shows Z-score for neighborhood alteration. P-values calculated using a wilcoxon signed-rank test comparing either WT (left, green) or *lakritz* (no RGCs; right, red) to a random label shuffle distribution.

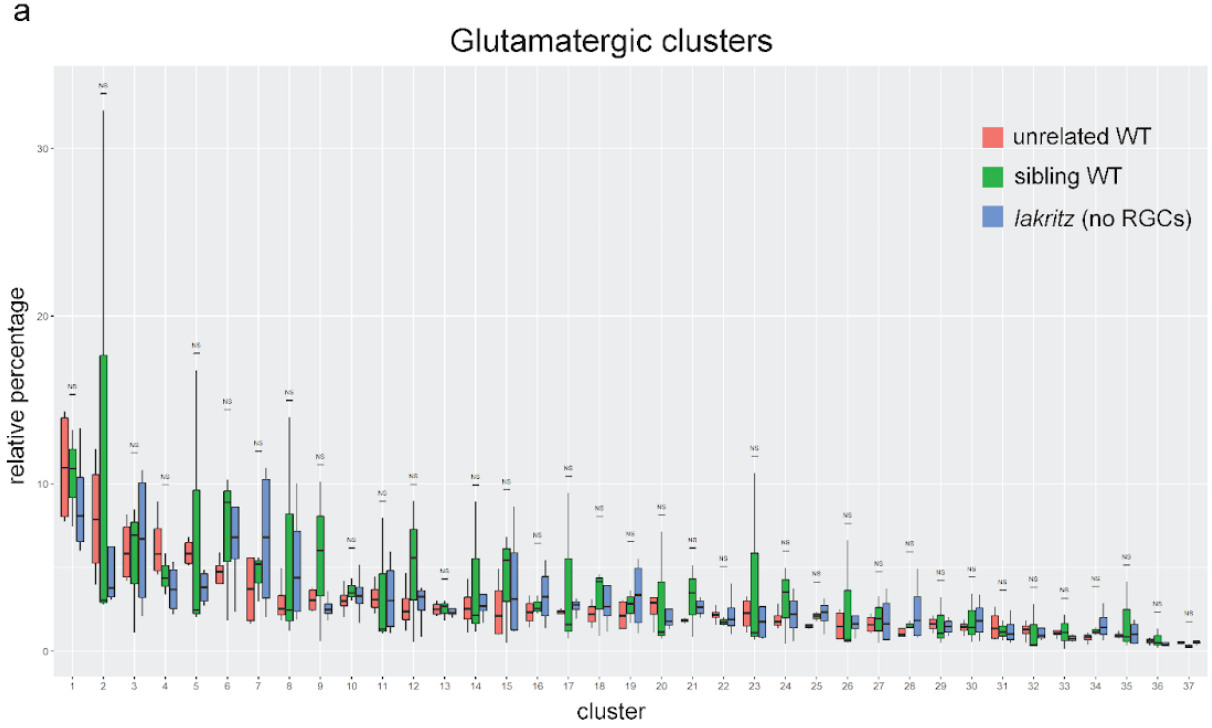
In absence of strong global effects for the absence of RGCs, I decided to explore whether specific clusters were affected more than others by a lack of retinal input. I first tested whether there were differences in the abundance of specific cell types. While the cell types could develop normally in absence of RGCs, there could be a change in the reinforcement of transcriptional programs - leading to a change in the ratio of cell types. However, I was unable to find statistically significant changes in the abundance of cell types (Figure 24). I next tested whether some clusters showed more severe alteration to their transcriptomes than others. While it is clear that all cell types are present in *lakritz*, it is possible that some types could experience alterations in marker gene expression but bioinformatically remain similar to the corresponding WT cluster. For this, I applied an analysis described in (Sharma N et al. Nature 2020). Conceptually, for each cluster the expression of marker genes is measured and compared to that of its WT counterpart. For most genes, the ratio of WT to *lakritz* gene expression will be equal to 1 (a large number of genes are expressed in most cells and involved in general cell maintenance). For marker genes, if RGCs are important for cell fate determination, then they will show severe changes in a cell type-specific manner. Using this method I was able to rank clusters in their level of alteration and calculate a list of marker genes that should label severely altered cell types.

To confirm that transcriptomes in specific cell types were altered in absence of RGCs, I performed comparative HCR stains between *lakritz* and their WT clutch siblings. For this, I created a cross between *lakritz* heterozygotes expressing the transgene Tg(HGn12C:GFP) with fish carrying the genetic elements Tg(Isl2b:RFP). Normally *lakritz* mutants can be identified by darker pigmentation. For HCR stains, however, they must be either Nacre mutants, lacking pigmentation, or reared in PTU which chemically blocks the biosynthesis of melanin. In both of these scenarios, it would be impossible to visualize the differences between *lakritz* and WT. Using the Tg(Isl2b:RFP) transgene, it is possible to identify *lakritz* mutants using a fluorescent microscope: the transgene drives the expression of RFP in both RGCs and the trigeminal ganglia (a sensory nerve bundle in the peripheral nervous system). In *lakritz*, which lacks all RGCs, the ganglia

will still express RFP. In this way I was able to identify *lakritz* mutants lacking pigmentation.

After sorting a sufficiently large number of pigmentless larvae from *lakritz* and WT clutch siblings, I performed HCR stains in parallel for a selected number of candidate marker genes. I then registered all acquired images to my HCR atlas and identified differences in the level and location of expression. Almost all of the candidate marker genes, which we hypothesized to show cell type specific transcriptome alterations showed absolutely no differences. Out of more than thirty candidate genes, only four showed an observable difference (Figure 25). The genes *aldh1a2* and *cabp5b* both showed down regulation in the pretectum, while the gene *calb2b* showed downregulation across the entire tectum. In all three cases the expression was still detectable. The gene *crhbp* was the only gene for which I could confirm that the expression in a specific cell type was gone. *Crhbp* is expressed in a large number of cells across the fore-, mid-, and hindbrain. However, in a small thalamic nucleus in *lakritz* it was completely abolished. While most single-cell clusters could be easily aligned between *lakritz* and WT, one of the *crhbp* clusters could not be easily aligned between samples. Together, this points to a cluster of thalamic *crhbp* cells as the only cells that I could identify as potentially missing in *lakritz*. In the larger context, however, the *lakritz* brain cell-typome is almost indistinguishable from that of the WT.

a


Glutamatergic clusters
unrelated WT
sibling WT
lakritz (no RGCs)
relative percentage
cluster

b

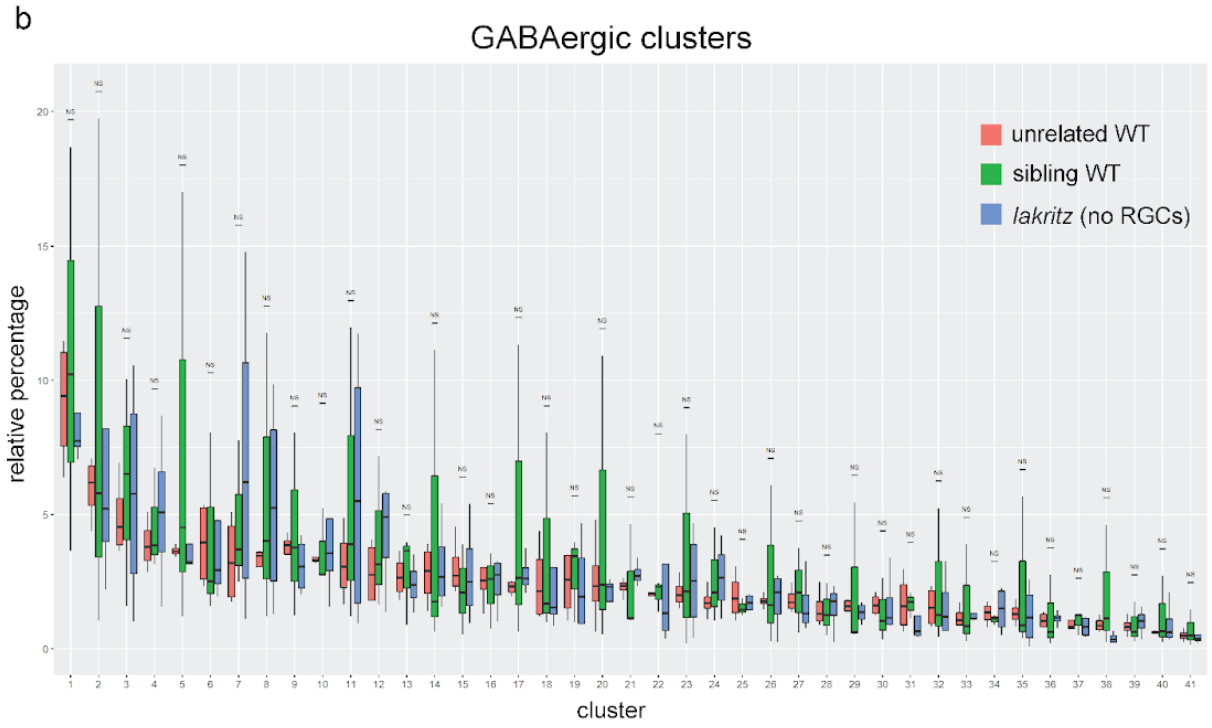

GABAergic clusters
unrelated WT
sibling WT
lakritz (no RGCs)
relative percentage
cluster

Figure 24. There is no significant change in relative cluster proportions in absence of RGCs

Bar plots showing for each cluster the variance in cluster's relative percentage across replicates (a, glutamatergic clusters; b, GABAergic clusters). For each cluster, the variance is shown for all genotypes (red, unrelated WT; green, sibling WT; blue, *lakritz* (no RGCs)). P-values calculated using a wilcoxon signed-rank test and corrected for multiple testing using the Bonferroni correction.

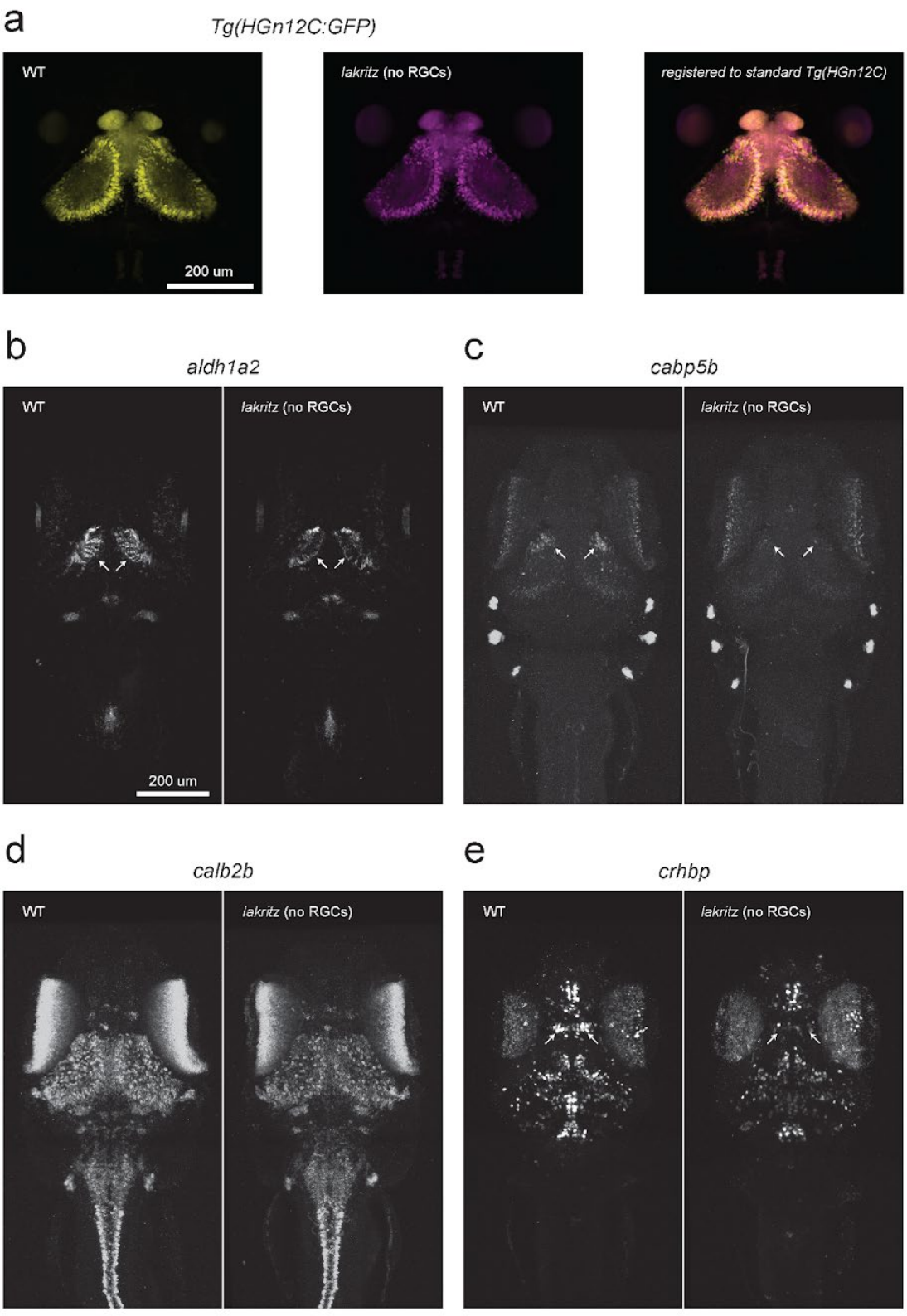

a
Tg(HGn12C:GFP)
WT
200 um
lakritz (no RGCs)
registered to standard Tg(HGn12C)
b
aldh1a2
WT
lakritz (no RGCs)
200 um
c
cabp5b
WT
lakritz (no RGCs)
d
calb2b
WT
lakritz (no RGCs)
e
crhbp
WT
lakritz (no RGCs)

Figure 25. A small subset of markers shows morphologically-specific altered expression in absence of RGCs

(a) Registration of Tg(HGn12C) expression patterns from WT (left, yellow), and *lakritz* (no RGCs; middle, magenta). Right panel shows registered expression patterns to standard HGn12C pattern. (b-e) Comparative stains between WT (left), and *lakritz* (no RGCs; right). Arrows (b,c,e) show areas where expression is altered in absence of RGCs. In (d) the tectum shows down-regulation.

Identifying and comparing differentiation trajectories in single-cell data

In addition to driving GFP expression in neurons, the Tg(HGn12C:GFP) transgenic line also drives expression in a small population of progenitor cells. These populations, however, are not discrete: progenitors eventually exit the mitotic cycle and commit to a neuronal fate. I noticed that in my data I was able to capture these transitions. I had transcriptionally profiled cells that appeared to be neither progenitors nor neurons. As I continued to transcriptionally profile even more cells, I noticed that these transitioning cells formed a continuum from progenitors to neurons. Embedding cells from all samples: unrelated WT, WT sibling, and *lakritz*, further increased the number of rare transitioning cells I could analyze together. This allowed me to resolve the differentiating pathways progenitors take as they gradually transition from uncommitted, mitotic progenitors to fate-committed post-mitotic neurons (Figure 26a).

Upon clustering my data, I could see that differentiating progenitors formed their own unique cluster. I identified markers that defined this differentiating cluster using the Seurat R package and then visually inspected each of the markers for annotation. I observed that markers generally fell in one of three categories: progenitor associated markers, transient markers, or cell-type markers (https://www.biorxiv.org/content/10.1101/2021.11.15.468630v1, Supplementary Table 2). I identified progenitor-associated markers as genes that undergo down-regulation in response to neuronal fate commitment. Essentially these genes are markers of the progenitor state, and their expression diminishes as commitment to a neuronal fate unfolds. On the other side of the spectrum were cell-type markers that are constitutively expressed in cell-types as they undergo terminal differentiation. For example, vglut2a and gad1b are markers of glutamatergic and GABAergic cell types, respectively. Cells that commit to a GABAergic fate, will up-regulate the expression of gad1b and will constitutively express it throughout their life. They may then undergo further differentiation into more specific GABAergic neuronal types. The most interesting markers, however, were transient markers that are expressed temporarily as a cell makes fate decisions.

Transient markers were the most abundant category of marker genes in the differentiating progenitors cluster (neuronal precursor cluster). My analysis was able to identify many such putative markers (https://www.biorxiv.org/content/10.1101/2021.11.15.468630v1, Supplementary Table

2). However, to confirm that these markers indeed were involved in fate selection is a daunting task and extends beyond the scope of this thesis. I decided thus to look at whether I could identify canonical differentiation markers. One of the most striking features of the precursor cluster is that it clearly separates into different trajectories: a bottleneck of progenitors turns into two precursor paths leading to the GABAergic and glutamatergic neuronal clusters. Concomitant with this split precursor along the GABAergic path express *ascl1b* while the ones on the glutamatergic path express *neurog1*. *Ascl1b* and *neurog1* are both canonical, evolutionarily conserved, markers defining progenitors differentiating into GABAergic and glutamatergic neuronal types[120,121]. They are also mutually exclusive and are only expressed during a short period. *Neurod1* is expressed along the same trajectories as *neurog1*, except that it's expressed later on in the developmental pathway - turning on after *neurog1* and turning off as precursors become more differentiated[122,123]. The excitatory precursors split into two different trajectories - one leading into the habenula super-class, and the other leading to excitatory neurons of the thalamus, pretectum, and tectum. The differentiation into a habenula excitatory neuron is coupled to the expression of *cxcr4b* along with the more general markers *neurog1*, and *neurod1*. The other excitatory neurons do not express *cxcr4b*, but rather *neurod2* along with *neurod1*, and *neurog1* (https://www.biorxiv.org/content/10.1101/2021.11.15.468630v1, Supplementary Table 2).

I next asked how similar the developmental trajectories between *lakritz* and WT are. While it was clear that most, if not all cell types develop properly in *lakritz*, it is possible that visual cell types could develop along altered developmental trajectories in absence of RGCs. In addition to their functional role, RGCs also secrete different morphogens. For example, RGCs secrete the protein sonic hedgehog. In the retina, the secretion of this protein by RGCs results in the suppression of further RGC differentiation[124]. However, I could not see a difference between *lakritz* and WT in their developmental trajectories (Figure 26b,c). One effect that was clear was that the ratio of neurons to progenitors was significantly different between *lakritz* and WT. In WT there were 19 neurons for every progenitor and in *lakritz* there were only 9. This means that in *lakritz* there is an accumulation of progenitors, or that differentiation is slower, albeit along the same trajectories. To validate this effect, I performed semi-quantitative comparative HCR stains for *lakritz* and WT clutch siblings (Table 1). I stained for genes that are expressed in progenitors and early precursors (*cyp19a1b, pcna, her4.1, fabp7a, gfap, p27, s100b*), and genes that are expressed in differentiating and late neuronal-committed precursors (*ascl1b, neurog1, sox2, neurod1*).

In *lakritz*, I could visualize that *pcna* (a marker of cycling progenitors) and *s100b* (a marker of early precursors) are both downregulated in retinorecipient areas. The largest effect I observed was in the downregulation of *fabp7a* (a marker of early glial precursors) across multiple retinorecipient areas. All other markers showed largely no effects. I interpreted these results to mean that in *lakritz*, in absence of RGC effectors, the cell cycle is altered. This, coupled with dysregulated maturation of a population of early and glial precursors, likely underlies the increase in the ratio of progenitors-to-neurons.

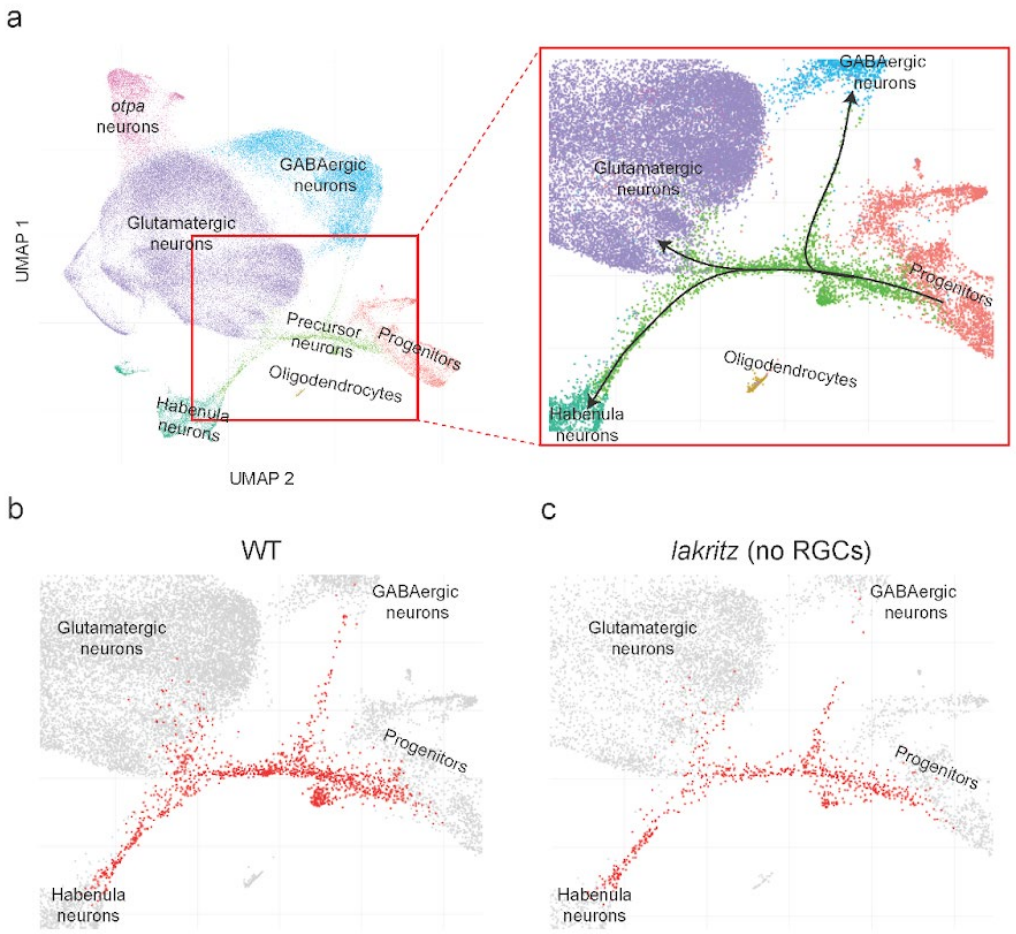


Figure 26. Differentiation trajectories are conserved in absence of RGCs

(a, left) UMAP embedding of all cells color-coded by cell identity. Text labels adjacent cell classes. (right) Enlarged area (red square, left) showing a class of neuronal precursor cells differentiating into the major cell classes. (b,c) same enlarged area as in (a, right), but showing only WT cell (b), or lakritz (no RGCs) cells. In red are highlighted all neuronal precursor cells showing splits along the same differentiation trajectories.

gene	Anatomical mask	expression fold change (log2)	adjusted p-value
ascl1b	cerebellum ventricle	0.1585	1
ascl1b	diencephalon ventricle	0.0329	1
ascl1b	tectum midline ventricle	-0.2134	1
ascl1b	tectum proliferative ventricle	-0.1663	1
ascl1b	telencephalon ventricle	0.1389	0.5772
ascl1b	cerebellum	-0.0522	1
ascl1b	dorsal thalamus	-0.0415	1
ascl1b	pretectum	-0.0399	1
ascl1b	ventral thalamus	0.1261	1
ascl1b	tectum cell layers	0.1359	1
ascl1b	telencephalon	-6.197	0.0603
cyp19a1b	cerebellum ventricle	0.286	1
cyp19a1b	diencephalon ventricle	0.055	1
cyp19a1b	tectum midline ventricle	0.0697	1
cyp19a1b	tectum proliferative ventricle	0.1825	1
cyp19a1b	telencephalon ventricle	0.0728	1
cyp19a1b	cerebellum	0.3001	1
cyp19a1b	dorsal thalamus	0.0211	1
cyp19a1b	pretectum	-0.0542	1
cyp19a1b	ventral thalamus	0.0811	1
cyp19a1b	tectum cell layers	0.055	1
cyp19a1b	telencephalon	-0.7202	1
fabp7a	cerebellum ventricle	-0.1872	1
fabp7a	diencephalon ventricle	-0.769	2.00E-04
fabp7a	tectum midline ventricle	-1.194	1.00E-04
fabp7a	tectum proliferative ventricle	-0.7334	3.00E-04
fabp7a	telencephalon ventricle	-0.5334	0.0067
fabp7a	cerebellum	-0.5665	1
fabp7a	dorsal thalamus	-0.9196	1.00E-04
fabp7a	pretectum	-1.0535	0
fabp7a	ventral thalamus	-0.5531	0.0053
fabp7a	tectum cell layers	0.459	1
fabp7a	telencephalon	0	1
gfap	cerebellum ventricle	0.139	1
gfap	diencephalon ventricle	-0.1779	1
gfap	tectum midline ventricle	0.1191	1
gfap	tectum proliferative ventricle	0.0876	1
gfap	telencephalon ventricle	-0.002	1
gfap	cerebellum	-1.6311	1

gfap	dorsal thalamus	-0.1381	1
gfap	pretectum	-0.1429	1
gfap	ventral thalamus	-0.09	1
gfap	tectum cell layers	-1.1658	1
gfap	telencephalon	-0.6973	1
her4.1	cerebellum ventricle	0.1146	1
her4.1	diencephalon ventricle	-0.0469	1
her4.1	tectum midline ventricle	-0.0801	1
her4.1	tectum proliferative ventricle	-0.0958	1
her4.1	telencephalon ventricle	0.1141	1
her4.1	cerebellum	-0.4754	1
her4.1	dorsal thalamus	-0.0245	1
her4.1	pretectum	-0.1049	0.3022
her4.1	ventral thalamus	0.1551	1
her4.1	tectum cell layers	0.0257	1
her4.1	telencephalon	3.0613	1
neurod1	cerebellum ventricle	-0.532	1
neurod1	diencephalon ventricle	-0.462	0.888
neurod1	tectum midline ventricle	-0.3831	1
neurod1	tectum proliferative ventricle	-0.4937	1
neurod1	telencephalon ventricle	-0.3803	1
neurod1	cerebellum	-1.4612	1
neurod1	dorsal thalamus	0.0438	1
neurod1	pretectum	-0.1774	1
neurod1	ventral thalamus	0.1316	1
neurod1	tectum cell layers	0.8349	1
neurod1	telencephalon	-24.0788	1
neurog1	cerebellum ventricle	0.4766	0.0181
neurog1	diencephalon ventricle	0.3271	0.9556
neurog1	tectum midline ventricle	0.1882	0.2377
neurog1	tectum proliferative ventricle	0.3252	0.6699
neurog1	telencephalon ventricle	0.5117	0.0957
neurog1	cerebellum	-1.296	1
neurog1	dorsal thalamus	0.1227	1
neurog1	pretectum	0.0683	1
neurog1	ventral thalamus	0.2692	0.0789
neurog1	tectum cell layers	0.1761	1
neurog1	telencephalon	-0.2668	1
p27	cerebellum ventricle	-0.0438	1
p27	diencephalon ventricle	-0.0939	1
p27	tectum midline ventricle	-0.0255	1

p27	tectum proliferative ventricle	-0.1487	1
p27	telencephalon ventricle	-0.1036	1
p27	cerebellum	0.2801	1
p27	dorsal thalamus	-0.1158	1
p27	pretectum	-0.108	1
p27	ventral thalamus	0.0358	1
p27	tectum cell layers	0.3089	1
p27	telencephalon	-0.7311	1
pcna	cerebellum ventricle	-0.5393	1
pcna	diencephalon ventricle	-0.6033	0.1603
pcna	tectum midline ventricle	-0.1061	1
pcna	tectum proliferative ventricle	-0.5537	1
pcna	telencephalon ventricle	-0.4541	1
pcna	cerebellum	-1.1997	0.9679
pcna	dorsal thalamus	-0.6093	0.0321
pcna	pretectum	-0.4451	0.459
pcna	ventral thalamus	-0.4499	1
pcna	tectum cell layers	-0.2186	1
pcna	telencephalon	3.3507	1
s100b	cerebellum ventricle	-0.1168	1
s100b	diencephalon ventricle	-0.2923	0.2019
s100b	tectum midline ventricle	-0.698	0.0215
s100b	tectum proliferative ventricle	-0.189	1
s100b	telencephalon ventricle	-0.1815	1
s100b	cerebellum	-0.2281	1
s100b	dorsal thalamus	-0.1772	1
s100b	pretectum	-0.3712	0.2793
s100b	ventral thalamus	-0.1702	1
s100b	tectum cell layers	-0.0835	1
s100b	telencephalon	-2.3256	1
sox2	cerebellum ventricle	0.0779	1
sox2	diencephalon ventricle	0.158	0.6389
sox2	tectum midline ventricle	0.1187	1
sox2	tectum proliferative ventricle	0.0232	1
sox2	telencephalon ventricle	0.1669	1
sox2	cerebellum	1.5038	1
sox2	dorsal thalamus	0.0361	1
sox2	pretectum	0.0601	1
sox2	ventral thalamus	0.2146	0.531
sox2	tectum cell layers	0.4726	1
sox2	telencephalon	-3.0421	1

Table 1. Dysregulation of some progenitor markers in areas lacking retinal input.

Table of markers for critical transitions during neuronal differentiation. Marker expression was quantified in key brain areas that receive retinal input and some that do not. Expression fold change shows whether a gene is up- or down-regulated in *lakritz* mutants. The p-values were calculated using a Wilcoxon signed-rank test and corrected for multiple testing using the Bonferroni correction.

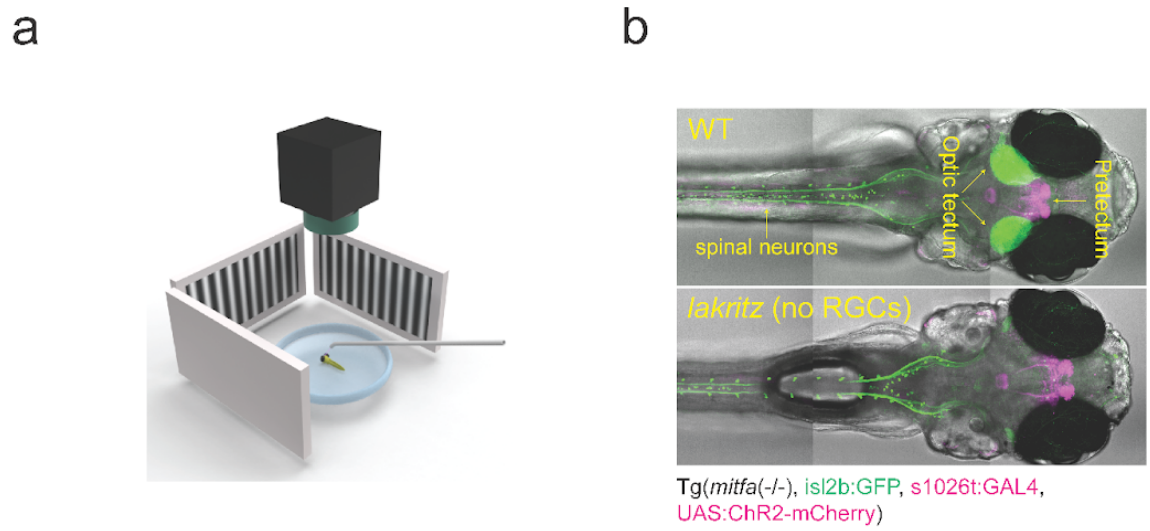
Formation of visual behavior circuits in absence of retinal input

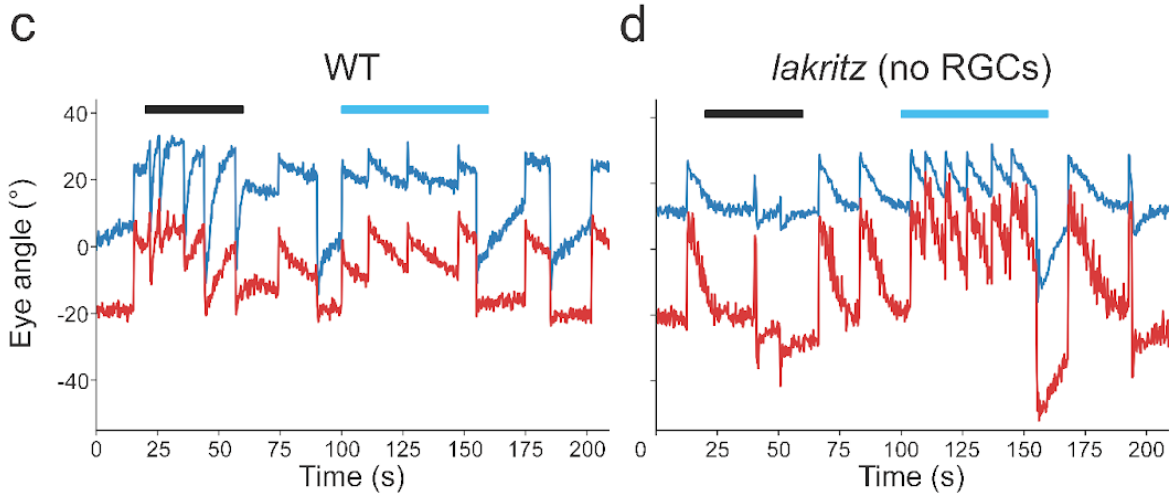
Because all of the cell types, visual and non-visual appear present in *lakritz*, I decided to ask whether these can assemble into circuits capable of driving visual behavior. A small population of pretectal neurons in zebrafish larvae can normally drive optokinetic reflex (OKR) in response to optic-flow[20,21]. It is also possible to optogenetically activate these neurons to elicit OKR[20,21] (Figure 27a). I thus decided that optogenetically activating the same population of neurons in *lakritz* can shine light on whether a visual behavior circuit can properly wire in absence of retinal input. If the pretectal OKR circuitry properly assembles in absence of such input, optogenetic activation should elicit OKR in *lakritz*.

One of the challenges of executing this experiment was producing transgenic larvae expressing multiple elements and harboring multiple mutations. I was able to eventually generate triple transgenic larvae Tg(*Gal4s1026t, UAS:ChR2-mCherry, isl2b:GFP*) on a *nacre* and *lakritz* background through multiple generations of outcrossing individual lines. The combination of Gal4 and UAS transgenes allowed to optogenetically target the specific pretectal population driving OKR, while the *nacre* mutation ensured no pigment could absorb excitation light. The isl2b:GFP transgene ensured that I would be able to identify *lakritz* mutants in absence of pigmentation (Figure 27b).

I embedded transgenic 6 dpf larvae in agarose and later released their eyes so that their movement can be recorded. My experiment first elicited OKR by showing moving gratings to larvae (visual stimulation) followed by optogenetic activation using an optic-fiber. In response to visual stimulation, WT control larvae would always perform OKR with their eyes pursuing the moving gratings and resetting with a fast saccade. *Lakritz* blind mutants never performed OKR in response to visual stimulation, as expected (Figure 27c-e). When I stopped the visual stimulation, WT larvae would also immediately stop performing OKR. After a short break, I optogenetically stimulated the pretectal population driving OKR labeled by the Tg(Gal4s1026t) transgene. In both WT and *lakritz* larvae, I could robustly elicit OKR-like behavior (Figure 27c-e). In *lakritz* larvae lacking ChR2 expression, I could never elicit OKR. And, in *lakritz* larvae expressing ChR2, OKR never appeared in absence of excitation light, or if areas other

than the pretectum were targeted. The fact that *lakritz* can perform OKR, means that the visual circuitry underlying this behavior can properly assemble without retinal input. My findings match well a separate study which managed to elicit prey capture behavior (another visual behavior) by targeting a different pretectal population[125]. Together, these findings support that circuits underlying visual behavior can develop and wire up in complete absence of any retinal input.


a
b
WT
Optic tectum
Pretectum
spinal neurons
lakritz (no RGCs)
Tg(mitfa(-/-), isl2b:GFP, s1026t:GAL4,
UAS:ChR2-mCherry)


c
WT
d
lakritz (no RGCs)
Eye angle (°)
40
20
0
-20
-40
0
25
50
75
100
125
150
175
200
Time (s)

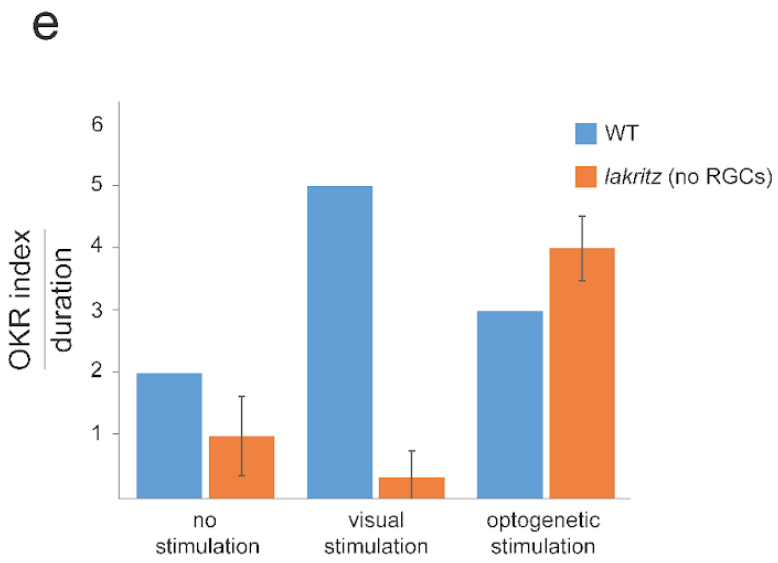

e
OKR index
duration
6
5
4
3
2
1
WT
lakritz (no RGCs)
no
stimulation
visual
stimulation
optogenetic
stimulation

Figure 27. The pretectal OKR circuitry assembles to generate behavior without retinal input

(a) Illustration of experimental setup. Larvae embedded in agarose on a small transparent dish facing screens showing moving gratings (generating optic-flow). A camera records eye movement and an optogenetic fiber can shine light into the brain. (b) Selected images of larvae used for experiment. Maximum Z-projection of either WT (top) or *lakritz* (no RGCs; bottom) 6 dpf transgenic larvae Tg(*mitfa* (-/-; transmitted light), *isl2b:GFP* (green), *Gal4s1026t*, *UAS:Chr2-mCherry* (magenta)). GFP expression was used to phenotype *lakritz* (no RGCs) mutants. (c,d) Eye movement traces (blue, left eye; red, right eye) of larvae in (b) either WT (c), or *lakritz* (no RGCs; d). Black line shows interval of visual stimulation (gratings moving). Cyan line shows interval of optogenetic stimulation. (e) OKR index for the three experimental phases (no stimulation, visual stimulation, and optogenetic stimulation). Increase in OKR index indicates a shift from spontaneous eye movement to repetitive synchronized eye movement (for *lakritz* n=3)

Optogenetic activation of the pretectal population underlying OKR has only been successfully achieved in the transgenic line Tg(Gal4s1026t). The cells I transcriptionally profiled across the larval visual centers, however, were all from the Tg(HGn12C:GFP) transgenic line. I wanted to confirm that the cells I optogenetically activated in the Tg(Gal4s1026t) transgenic line were also labeled in the Tg(HGn12C:GFP) line. I began by registering the Tg(HGn12C:GFP) expression pattern to the mapzebrain atlas, to which the expression pattern of the Tg(Gal4s1026t) was already registered. Both patterns were largely similar. The HGn12C contained areas labeled only partially by the 1026t line. But, all parts of the pretectum labeled by the 1026t line, were also labeled by the HGn12C line. I next transcriptionally profiled cells from the 1026t line to confirm that it overlaps with the HGn12C line. I crossed Tg(Gal4s1026t) fish with fish carrying the Tg(UAS:GFP) transgene and processed the larvae in the same way as those of the Tg(HGn12C:GFP) line (see experimental procedure). I transcriptionally profiled 8775 cells from the 1026t line and aligned them to existing data from the HGn12C line. I could not identify a striking difference between the neuronal or progenitor types labeled in either line (Figure 28). I could, however, identify a large difference between the two lines: the 1026t line labels a much larger population of glia than the HGn12C line does. As the neuronal population of the 1026t is completely labeled within the HGn12C line, I concluded that the population of neurons driving OKR is present in both WT and *lakritz*.

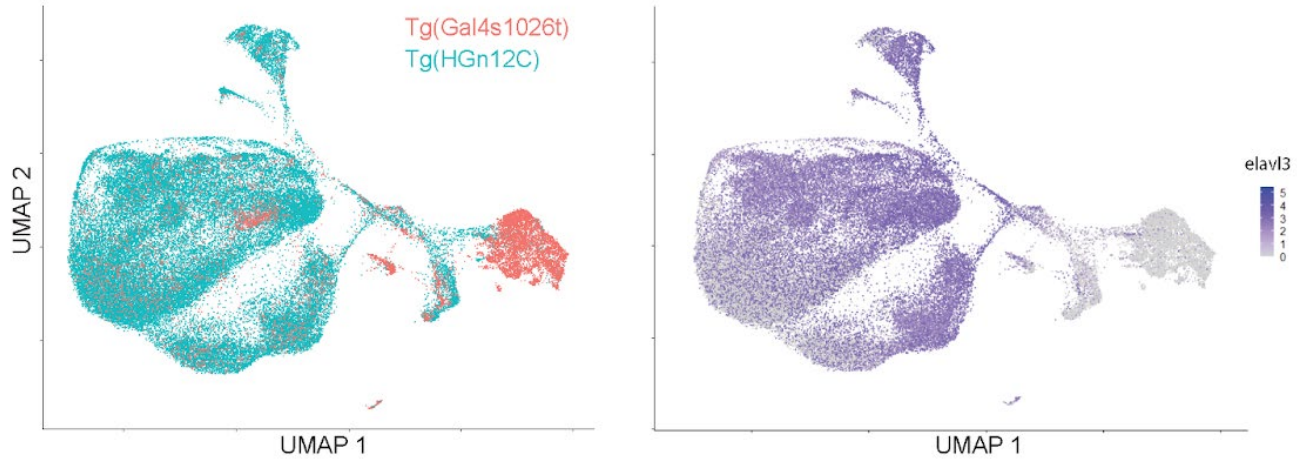


Figure 28. The Tg(HGn12C) pattern encompasses Tg(s1026t) pretectal neurons

(a) UMAP embedding of scRNA-seq results collected from Tg(s1026t; red, left) together with all cells from WT Tg(HGn12C; cyan, left). On the right, color coding of *elavl3* expression level across all cells in UMAP.

DISCUSSION

Summary of key findings

In this thesis, I transcriptionally profiled cells from visual processing centers of the larval zebrafish and used this to generate a gene expression atlas. I next transcriptionally profiled visual cells from a blind mutant lacking all retinal ganglion cells (RGCs). I developed a computational pipeline to compare the two samples and gained insight into how the brain's visual system develops in absence of retinal input. Lastly I used optogenetic manipulation to activate a visual circuitry in this blind mutant. The results of my thesis provided the following advances in the field of neuroscience:

My work on transcriptionally profiling cell types has improved on existing cell dissociation protocols. The improved dissociation procedure, presented in this thesis, offers a significant improvement over older protocols in quickly processing zebrafish larvae for fluorescent-activated cell sorting (FACS) followed by single-cell sequencing[59]. The dataset of transcriptionally-profiled cells I generated offers the first full accounting of cell types in any vertebrate retinorecipient area. The full characterization of all thalamic neuronal types is of special interest, as it is the central retinorecipient site in most, if not all, mammals. My work has uncovered hundreds of putative marker genes for cell types in these areas - offering a swath of genetic access points to further explore the structure and function of the zebrafish visual brain.

The gene expression atlas I generated, using the most informative marker genes, is the first resource of its kind available to zebrafish neuroscientists. While other gene expression atlases do exist, they are different. The Allen Brain Atlas, the most famous vertebrate gene expression atlas, was generated using physical sections across the entire mouse brain[126]. In order to bring sections from different animals into a common atlas space, researchers relied on heavy computational tools to infer gene expression across the entire brain. Brain atlases also exist in zebrafish, but they mainly aimed to align different transgenic lines[61,127]. However, transgenic lines often do not recapitulate the pattern of gene expression. In most cases, random insertions of a reporter gene will drive expression under unknown enhancers; the resultant expression pattern unlike that of any single gene. The Thisse atlas is by far the most elaborate attempt at a zebrafish gene expression atlas[128]. The atlas involved the characterization of roughly ten thousand genes throughout development. While very good for visualizing early patterning events, the colorimetric stains are mostly uninterpretable in later stages in understanding regionalized gene expression patterns. The atlas I generated in this thesis has all information needed to understand where in the brain a gene is expressed. It has single cell resolution. And, it is aligned to the mapzebrain MPIN atlas allowing the computational use of anatomical annotation. The atlas and its analysis offer a step forward in understanding brain regionalization via gene expression. It will, in no doubt, offer other researchers in the field a resource they can use to validate their own single-cell expression experiments. It is also highly likely that researchers will use this resource as a genetic access window to create future transgenic lines driving

expression in specific brain areas - allowing unprecedented access for circuit investigation.

The last part of my thesis involved understanding how a part of the brain involved in sensory processing develops in absence of input or contact from the sensory surface. Specifically for the visual system, perturbation experiments have shown a wide range of changes if visual input was altered, or if it was absent altogether[2,129,130]. Changes in the visual system included: changes in the layering of the visual system[9], poor refinement of retinotopic maps[131], and altered visual behavior[2]. Studies have also shown that some cell types are absent from their appropriate location in the brain following perturbation[9]. These findings drove a sentiment that cell types are likely to be altered or completely absent when an animal's visual experience is perturbed. My work in this thesis has shown that this is not true; that most, or all visual cell types, develop their cell type-defining transcriptomic signature in complete absence of retinal input. This finding conclusively shows that retinal input has little effect on cell type determination. In driving visual behavior in *lakritz* blind mutants (lacking all RGCs), I was able to further show that the genetic programs unfolding in the developing, plastic, vertebrate brain are sufficient to drive the assembly of circuits capable of driving visual behavior. While this finding may be surprising in the context of mammalian cortical research, it is less so in light of years of zebrafish research[109,111,132]. Earlier works have already shown that zebrafish larvae, reared in dark (and thus deprived of visual experience), behave indistinguishably from their light-reared siblings. While the mammalian cortex is likely an exception in how neural circuits develop, the main retinorecipient sites are conserved across all vertebrates. Thus the rules governing neural development, including cell-type fate determination, in these areas in zebrafish likely extend to mammals and other vertebrates.

Resolving cell-type heterogeneity in visual processing centers

Large-scale, unbiased, and systematic identification of cell types has been instrumental in resolving the plethora of cell types in different animals, tissues, and states. In this work, I identified a transgenic line which labels the major visual-processing centers in zebrafish and used it to systematically characterize cell types in those areas.

To take advantage of the 10X Genomics system for transcriptionally profiling single-cells, I improved on an existing protocol established previously[59] to dissociate retinal ganglion cells (RGCs). The end result offered a significant improvement in tissue handling, repeatability, survival of cells, and overall quality of the data. In total, I created a resource of 100,000 cells spanning the major retinorecipient sites in zebrafish – offering a treasure trove of genetic markers to explore the molecular architecture of the visual system.

While the architecture of the visual system has been explored in other works, it has seldom been done systematically. Most works focused on a handful of markers, which were mostly visualized independently. To avoid this, I created an atlas of all genetic markers I identified by single-cell sequencing. Using only a single bridging stain, I was able to visualize any combination of genetic markers. Exploring this atlas, I could clearly see that most markers labeled specific cell types in different brain areas. *Npy* for example, labels a subset of cells in the tectal cell body layer, while the gap junction protein *gjd2b* labeled a population of tectal superficial interneurons. Genetic markers also show how visual centers differ in molecular organization from one another. For example: pretectal cell types show a diffuse pattern whereas thalamic cell types routinely form nuclei.

While unexplored in my work, the molecular architecture is likely related to the functional properties of different brain areas. Cell types in the tectum segregate along different tectal cell-body layers. As different layers mediate different functions, it is likely that a combination of genetic markers could be used to define a functional tectal cell type. In the pretectum are cells around AF7 that respond selectively to prey, but attempts to find genetic markers for these cells have fallen short. Pvalb6, for example, is expressed in these cells, but also in many other brain areas. In my work I have uncovered multiple genes which show differential expression in cells occupying the same niche. Future studies could use these markers to understand how the prey-capture circuitry develops and functions. Lastly, in the thalamus each of the nuclei I uncovered could potentially play a role in a different set of behaviors. One of these markers, *pth2*, has been identified as marker for cell types involved in social behavior [114]. Further mining of this atlas will surely lead to insight into cell types underlying visual processing.

Development of visual centers without retinal input

What is the role of innervation in growth and development? Sensory surfaces (organs involved in sensing the external environment) develop in parallel to brain areas specialized to interpret activity from them. Both, together, form specialized sensory systems and their development must be tightly coordinated for an animal to properly respond to environmental cues.

Surgical manipulation experiments of sensory inputs revealed the importance of normal growth and development of axonal targets. Removal of the olfactory placode in frogs, for example, arrests the development of the olfactory bulb[133]. In the auditory system of chicks, removal of the cochlea leads to loss of activity in auditory processing centers (specifically, the nucleus magnocellularis) and eventually, cell-death[134]. Is the altered course of development a result of loss of neuronal activity or lack of molecular signals?

The understanding that neural activity can influence growth and development is not new. Early experiments in cats showed that monocular deprivation leads to deficient vision and atrophy of parts of the LGN associated with the deprived eye[135,136]. Spontaneous-structured retinal activity in mice during prenatal development propagates firing of action potentials to the LGN. Blocking this activity inhibits the separation of the LGN into eye-specific layers[137,138]. Neuronal activity can also affect gene-regulation in the neurons themselves as well as post-synaptic partners. *cFos*, as well as other immediate early genes, are routine indicators of neuronal activity. They are also heavily up-regulated after seizures in brain areas ranging from the cortex to the limbic system[139,140]. Both NGF and BDNF are also up-regulated by neuronal activity and are critical for survival. They are also secreted and can influence the survival of neuronal partners. Thus these factors mediate a paradigm where groups of neurons can influence each other's survival by intrinsic activity[141,142].

There also exist strong evidence that secreted molecules transported along neuronal process can influence the development of synaptic partners. In the visual system of insects, in-growing retinal axons influence the development of the target tissue through secretion of signaling molecules. In flies, for example, retinal axons induce precursor cells in the lamina to undergo terminal cell division, the first step required to produce neurons from progenitors[143]. Retinal axons play this critical role by transporting sonic hedgehog (SHH) and the EGF receptor ligand Spitz to the lamina where they act together to activate the EGF pathway to induce differentiation[144–146]. Further experiments in rats and chicks have demonstrated that this is not unique to the insect visual system[147–149].

While all of these show the strong influence the sensory surface has on the development of downstream brain areas, they do not address the basic question of their influence on the emergence of cell-type heterogeneity in these areas. In my thesis I decided to address this by characterizing the full repertoire of cell types in zebrafish visual brain areas with and without retinal input. In *lakritz* mutants, retinal axons fail to develop because of a single truncated transcription factor that is required only for RGC cell-fate determination. This mutant offered a clean manipulation to understand whether any of the contributions of RGCs (via activity or molecules) to the development of the visual system is important for cell-type fate determination. It is striking that I was unable to identify a class of cells, or even a single cell type, which was absent in these congenitally blind mutants.

Blind *lakritz* mutants do show up- and down-regulation in a large number of genes. However, so do it's phenotypically WT siblings (heterozygotes or null) when compared to larvae from an outgroup. Mutant cells do show altered neighborhoods in PC space – meaning that *lakritz* cells resemble more closely other *lakritz* cells than WT cells, and more than what we can expect by random chance. These altered neighborhoods, however, are apparent throughout the UMAP space and are not localized to specific hotspots – resembling a global transcriptomic drift. While *lakritz* cells may show a global drift from WT ones, overall relationships between cell types are conserved; within the *lakritz* brain, cell types maintain a similar transcriptomic distance. The PCs are also remarkably similar between *lakritz* and WT: each top PC in the WT population correspond to a single PC in the *lakritz* population. Furthermore, visualization of

individual marker genes in *lakritz* and WT failed to unravel a difference in the presence of specific cell types. Some of the differences in the expression of specific genes I could verify by in-situ hybridization were in genes involved in calcium signaling. As neural activity is tightly linked to fluctuations in calcium, it is maybe not surprising that in *lakritz* there is down-regulation of proteins responsible for calcium buffering.

Using optogenetics I was also able to show that in *lakritz* a circuit involved in visual behavior can properly assemble in absence of any retinal input. The OKR circuitry, involved in gaze stabilization, can drive OKR behavior in these blind mutants. This is a surprising find considering the large body of evidence showing that the overall architecture of visual processing center is heavily influenced by visual experience and its deprivation – causing visual behavior defects that can last a lifetime. However, these experiments were mainly performed in mammals. Experiments in zebrafish have shown the opposite: that visual experience is dispensable for the development of zebrafish behavior. Zebrafish reared in dark can perform OKR and OMR indistinguishably from their light-reared siblings[108,111]. Additionally, optogenetic activation of a prey-capture circuit results in blind larvae performing prey-capture[150]. Therefore it is maybe less surprising that zebrafish *lakritz* mutants can perform visual behavior under the right conditions.

Do RGCs influence neurogenesis?

Teleosts brains facilitate continuous neurogenesis. Compared to mammals, fish can grow throughout their lives and so can their brain and the peripheral nervous system. Neurogenesis, however, does not occur throughout the entire brain, but rather is confined to specific areas. It is the spatial organization of these areas that allowed researchers to understand the effects of extrinsic and intrinsic cues on the different stages of neurogenesis: proliferation, cell-cycle dynamics, and differentiation[151]. For example: growth hormone (GH) acts broadly to maintain neurogenesis levels via insulin-like growth factor-I (IGF-I). Cells in the proliferative zone of the retina, the ciliary marginal zone (CMZ), express receptors for both GH and IGF-I. And, artificially increasing the concentration of IGF-I increases the rate at which retinal progenitors proliferate[152,153].

Additional extracellular factors and pathways that affect proliferation and differentiation have been identified. The Wnt/beta-catenin and Hedgehog pathways both control the rate of proliferation of CMZ progenitors. In frogs, activation of the Wnt pathway causes increased proliferation of the CMZ. Conversely, blocking the Wnt pathway, inhibits CMZ proliferation[154,155]. Hedgehog activation increases proliferation while also accelerating cell-cycle exit[156,157]. In zebrafish, acute expression of sonic hedgehog (*shh*) during early retinal development is required for RGC differentiation[158,159].

In my thesis, I observed that we can detect the transitions from post-mitotic uncommitted progenitors to committed neurons. This is likely because zebrafish, as other teleosts, undergo continuous neurogenesis. Additionally, the large number of cells I transcriptionally profiled allowed to resolve these rare states. The cluster of precursor cells (undergoing neuronal commitment and terminal differentiation) is different from other single-cell clusters in that its UMAP structure is informative. The cluster clearly marks the path from progenitors to neurons and then to either an excitatory or inhibitory neuronal fate. The markers of this cluster are also different from other clusters as they show peak expression levels along critical fate-decision junctions (compared to terminal markers which show a more homogenous expression in a given cluster).

In *lakritz* blind mutants this process is largely unaltered; the identical trajectories can be inferred from the WT as well as the mutant datasets. This means that lack of retinal input neither disturbs the terminal-selection of cell types, nor the fate decisions a progenitor undertakes to get there. Furthermore, the ratio of excitatory to inhibitory neurons is perfectly conserved between samples – one inhibitory neuron for every three excitatory neurons. While RGCs contribute both activity patterns and molecular factors to innervated downstream brain areas the two are unlikely to contribute equally to the regulation of neurogenesis. Based on existing data on how activity and molecular factors may regulate neurogenesis, It is likely that dysregulation arises because RGC-derived molecules are absent in *lakritz*.

CONCLUSION AND OUTLOOK

My thesis explored the development of visual processing centers in absence of any retinal input, functional or molecular. I resolved for the first time the complete diversity of cell type in a vertebrate retinorecipient area. To understand how these cell types develop in complete absence of retinal input, I transcriptionally profiled cells from visual processing centers in a blind mutant lacking all RGCs throughout its development. I found that most, if not all, cell types developed in this mutant. I also found that they develop in the same way in mutants and WT. The most striking effect I could identify between groups was that in absence of RGCs, cell types differentiate slower. In *lakritz* blind mutants I identified an increase in progenitors over neurons and was able to show that this is likely due to a slower exit from the cell cycle - the first step in the transition from uncommitted, mitotic progenitors to committed, post-mitotic neurons. Lastly, I was able to show that the forebrain, devoid of retinal input throughout its development, is still capable of assembling into circuits capable of generating visual behavior.

The zebrafish nervous system, including its visual system, grows throughout its life and shows continuous plasticity[160]. However, it appears that the sensory surface plays a minor, secondary, role in shaping central targets. Two recent publications have recently explored a similar question in the mouse visual cortex (V1), finding evidence largely agreeing with my findings[119,161]. While these studies have explored V1 in absence of visual experience, my thesis goes a step further by exploring cell composition in retinorecipient sites in complete absence of retinal input. Future experiments taking advantage of recent advances in spatial transcriptomics may further explore this paradigm to understand changes throughout development and into adulthood. Combining electron-microscopy connectomics with spatial transcriptomics may unravel how genetic programs underlying cell fate determination also drive the assembly of neural circuits.

Future studies looking to expand on any topics covered in this thesis could focus on these unanswered questions and future directions:

What would a full characterization of all visual cell types look like? In this thesis I focused on a small number of areas that are part of the “visual brain”, and a small part of the brain altogether. The tectum (the largest retinorecipient brain area) and the hypothalamus were both largely missing from my analysis – a result of using the HGn12C line. While it is possible that cell types in these areas are very similar to cell types uncovered in my work, it is more likely that they harbor novel cell types unique to these area. Future studies could focus on resolving these cell types, or even extend this endeavor to resolving all cell types in the larval zebrafish brain. The tectum specifically would be of great interest to explore possible unidentified differneces between WT and *lakritz*. The layered structure of the tectum would be of specific interest as it mediates both visual function and development. If this structure extends to the organization of

specific tectal cell types, it would be of particular interest to see whether these structures are conserved in the complete absence of retinal input.

How are cell types organized in visual brain centers? In this work I began answering this question by amassing a large number of hybridization stains and registering them into a common atlas – generating a spatio-molecular atlas of visual brain areas. This, however, falls short of truly uncovering the organization of cell types in these areas. In rare cases, a single gene could be associated with only one cell type. Hence, in these rare cases, visualizing a single gene is sufficient to understand the organization of a cell type in the brain. To understand how the majority of cell types are organized in the brain, however, one could take advantage from recent advances in spatial transcriptomics. This could be further combined with single-cell or single-nucleus RNA sequencing to assign true cell to putative location in a 3D brain model. Creating such a resource will in no doubt be incredible valuable to any neuroscientist looking to understand the brain of the zebrafish larva.

How do molecular cell types relate to functional cell types? Understanding how neurons establish their functional properties is a central goal of modern neuroscience. This is both due to conceptual as well as technical difficulties. In my thesis, I resolved molecular cell types by transcriptionally profiling a large number of cells from visual processing centers. This, however, left open the question of how do these molecular types relate to functional cell types. In the pretectum are direction-selective cells and ones responding to prey. In the thalamus are cells that respond specifically to conspecifics. Their molecular signature, however, remains unknown. It is likely that the specific set of molecules that can distinguish these cells from others lies in the dataset I collected in my thesis. More work can and should be done to link the two in future studies. Establishing driver lines that differentially label these functional cell types will be crucial to furthering our understanding of how the zebrafish brain drives behavior.

How do RGCs regulate neurogenesis? In my work I observed that *lakritz* mutants lacking RGCs harbor a larger progenitor population compared to WTs. It is of yet unclear how RGCs are involved in this process. Future studies could focus on looking into the different molecules RGCs secrete to understand how they regulate neurogenesis. Using new genetic manipulation tools it is possible to alter secretion of specific peptides only in RGCs. First steps could shed light on which of the major biochemical pathways is involved in this process and how their downstream effectors mediate the progenitor-to-neuron transition. It is also possible that while RGCs do not facilitate neurogenesis per se, they do influence the survival of cells along the neurogenic axis. Higher frequency of cell-death could similarly change the ratio of neurons to progenitors. Future studies exploring this avenue could label specifically the

progenitor and precursor populations to collect more data from these rare populations. Expanding on this will contribute tremendously to our understanding of how the sensory system regulates brain growth and development.

BIBILIOGRAPHY

1. Barlow, H. B. Visual experience and cortical development. *Nat. 1975 2585532* **258**, 199–204 (1975).
2. Daw, N. W. & Wyatt, H. J. Kittens reared in a unidirectional environment: evidence for a critical period. *J. Physiol.* **257**, 155–170 (1976).
3. Grubb, M. S. & Thompson, I. D. The influence of early experience on the development of sensory systems. *Curr. Opin. Neurobiol.* **14**, 503–512 (2004).
4. Katz, L. C. & Shatz, C. J. Synaptic Activity and the Construction of Cortical Circuits. *Science (80-.).* **274**, 1133–1138 (1996).
5. L, B., C, L., P, F., P, M. & LC, K. Odorant receptors instruct functional circuitry in the mouse olfactory bulb. *Nature* **419**, 296–300 (2002).
6. Espinosa, J. S. & Stryker, M. P. Development and Plasticity of the Primary Visual Cortex. *Neuron* **75**, 230 (2012).
7. D'Souza, S. & Lang, R. A. Retinal ganglion cell interactions shape the developing mammalian visual system. *Development* **147**, (2020).
8. Wallace, V. A. & Raff, M. C. A role for Sonic hedgehog in axon-to-astrocyte signalling in the rodent optic nerve. *Development* **126**, 2901–2909 (1999).
9. Golding, B. *et al.* Retinal Input Directs the Recruitment of Inhibitory Interneurons into Thalamic Visual Circuits. *Neuron* **81**, 1057–1069 (2014).
10. Rivlin-Etzion, M., Wei, W. & Feller, M. B. Visual Stimulation Reverses the Directional Preference of Direction-Selective Retinal Ganglion Cells. *Neuron* **76**, 518–525 (2012).
11. Engert, F., Tao, H. W., Zhang, L. I. & Poo, M. ming. Moving visual stimuli rapidly induce direction sensitivity of developing tectal neurons. *Nature* **419**, 470–475 (2002).
12. Ruthazer, E. S. & Aizenman, C. D. Learning to see: patterned visual activity and the development of visual function. *Trends Neurosci.* **33**, 183–192 (2010).
13. White, L. E. & Fitzpatrick, D. Vision and Cortical Map Development. *Neuron* **56**, 327–338 (2007).
14. D, S. & CJ, S. An instructive role for retinal waves in the development of retinogeniculate connectivity. *Neuron* **33**, 357–367 (2002).
15. A, A., P, G. & A, R. Activity dependent mechanisms of visual map formation--from retinal waves to molecular regulators. *Semin. Cell Dev. Biol.* **35**, 136–146 (2014).
16. Ge, X. *et al.* Retinal waves prime visual motion detection by simulating future optic flow. *Science (80-.).* **373**, eabd0830 (2021).

17. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202 (2015).

18. Trivedi, V., Choi, H. M. T., Fraser, S. E. & Pierce, N. A. Multidimensional quantitative analysis of mRNA expression within intact vertebrate embryos. *Development* **145**, (2018).

19. Kay, J. N., Finger-Baier, K. C., Roeser, T., Staub, W. & Baier, H. Retinal Ganglion Cell Genesis Requires lakritz, a Zebrafish atonal Homolog. *Neuron* **30**, 725–736 (2001).

20. Kubo, F. *et al.* Functional architecture of an optic flow-responsive area that drives horizontal eye movements in zebrafish. *Neuron* **81**, 1344–1359 (2014).

21. Wu, Y., dal Maschio, M., Kubo, F. & Baier, H. An Optical Illusion Pinpoints an Essential Circuit Node for Global Motion Processing. *Neuron* **108**, 722-734.e5 (2020).

22. Llinás, R. R. The contribution of Santiago Ramón y Cajal to functional neuroscience. *Nat. Rev. Neurosci.* **4**, 77–80 (2003).

23. Sotelo, C. Viewing the brain through the master hand of Ramon y Cajal. *Nat. Rev. Neurosci. 2003 41* **4**, 71–77 (2003).

24. Cajal, S. R. y & Ramón y Cajal, S. *Histologie du système nerveux de l'homme & des vertébrés. Histologie du système nerveux de l'homme & des vertébrés.* (Maloine, 1909). doi:10.5962/bhl.title.48637

25. Fishell, G. & Heintz, N. The neuron identity problem: form meets function. *Neuron* **80**, 602–612 (2013).

26. Zeng, H. & Sanes, J. R. Neuronal cell-type classification: challenges, opportunities and the path forward. (2017). doi:10.1038/nrn.2017.85

27. Stoeckius, M. *et al.* Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 1–12 (2018).

28. Setliff, I. *et al.* High-Throughput Mapping of B Cell Receptor Sequences to Antigen Specificity. *Cell* **179**, 1636-1646.e15 (2019).

29. Kubota, Y., Sohn, J. & Kawaguchi, Y. Large volume electron microscopy and neural microcircuit analysis. *Front. Neural Circuits* **12**, 98 (2018).

30. Eze, U. C., Bhaduri, A., Haeussler, M., Nowakowski, T. J. & Kriegstein, A. R. Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. *Nat. Neurosci. 2021 244* **24**, 584–594 (2021).

31. JH Lui, D. H. A. K. Development and evolution of the human neocortex. *Cell* **146**, 18–36 (2011).

32. Davie, K. *et al.* A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. *Cell* **174**, 982-998.e20 (2018).

33. Tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).

34. Luo, L., Callaway, E. M. & Svoboda, K. Genetic Dissection of Neural Circuits: A Decade of Progress. *Neuron* **98**, 256–281 (2018).

35. Luo, L., Callaway, E. M. & Svoboda, K. Genetic Dissection of Neural Circuits. *Neuron* **57**, 634 (2008).

36. Tsukahara, T. *et al.* A transcriptional rheostat couples past activity to future sensory responses. *Cell* **184**, 6326-6343.e32 (2021).

37. Regev, A. The Human Cell Atlas. *Elife* **6**, e27041 (2017).

38. Osumi-Sutherland, D. *et al.* Cell type ontologies of the Human Cell Atlas. *Nat. Cell Biol. 2021 2311* **23**, 1129–1135 (2021).

39. Li, H. *et al.* Fly Cell Atlas: a single-cell transcriptomic atlas of the adult fruit fly. *bioRxiv* 2021.07.04.451050 (2021). doi:10.1101/2021.07.04.451050

40. Packer, J. S. *et al.* A lineage-resolved molecular atlas of C. Elegans embryogenesis at single-cell resolution. *Science (80-.).* **365**, (2019).

41. Arendt, D. *et al.* The origin and evolution of cell types. *Nature Reviews Genetics* **17**, 744–757 (2016).

42. Lipovsek, M. *et al.* Patch-seq: Past, Present, and Future. *J. Neurosci.* **41**, 937–946 (2021).

43. Li, H. *et al.* Single-Cell Transcriptomes Reveal Diverse Regulatory Strategies for Olfactory Receptor Expression and Axon Targeting. *Curr. Biol.* **30**, 1189-1198.e5 (2020).

44. Bagnoli, J. W. *et al.* Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nat. Commun. 2018 91* **9**, 1–8 (2018).

45. Zhang, X. *et al.* Comparative Analysis of Droplet-Based Ultra-High-Throughput Single-Cell RNA-Seq Systems. *Mol. Cell* (2018). doi:10.1016/j.molcel.2018.10.020

46. Schaum, N. *et al.* Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nat. 2018 5627727* **562**, 367–372 (2018).

47. Siebert, S. *et al.* Stem cell differentiation trajectories in Hydra resolved at single-cell resolution. *Science (80-.).* **365**, (2019).

48. Taylor, S. R. *et al.* Molecular topography of an entire nervous system. *Cell* **184**, 4329-4347.e23 (2021).

49. Shafer, M. E. R., Sawh, A. N. & Schier, A. F. Gene family evolution underlies cell-type diversification in the hypothalamus of teleosts. *Nat. Ecol. Evol. 2021 61* **6**, 63–76 (2021).

50. Lei, Y. *et al.* Applications of single-cell sequencing in cancer research: progress

and perspectives. *J. Hematol. Oncol. 2021 141* **14**, 1–26 (2021).

51. Farrell, J. A. *et al.* Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science (80-.).* **360**, eaar3131 (2018).

52. Scott, E. K. & Baier, H. The cellular architecture of the larval zebrafish tectum, as revealed by Gal4 enhancer trap lines. *Front. Neural Circuits* **3**, 13 (2009).

53. Scott, E. K. The Gal4/UAS toolbox in zebrafish: new approaches for defining behavioral circuits. *J. Neurochem.* **110**, 441–456 (2009).

54. Halpern, M. E. *et al.* Gal4/UAS Transgenic Tools and Their Application to Zebrafish. *Zebrafish* **5**, 97 (2008).

55. Kawakami, K. Tol2: A versatile gene transfer vector in vertebrates. *Genome Biol.* **8**, 1–10 (2007).

56. Liu, J. *et al.* CRISPR/Cas9 in zebrafish: an efficient combination for human genetic diseases modeling. *Hum. Genet.* **136**, (2017).

57. Luo, J.-J. *et al.* CRISPR/Cas9-based genome engineering of zebrafish using a seamless integration strategy. *FASEB J.* **32**, 5132–5142 (2018).

58. Wang, Y. *et al.* A robust and flexible CRISPR/Cas9-based system for neutrophil-specific gene inactivation in zebrafish. *J. Cell Sci.* **134**, (2021).

59. Kölsch, Y. *et al.* Molecular classification of zebrafish retinal ganglion cells links genes to cell types to behavior. *Neuron* **109**, 645-662.e9 (2021).

60. Helmbrecht, T. O., Dal Maschio, M., Donovan, J. C. & Koutsouli, S. Topography of a Visuomotor Transformation. (2018). doi:10.1016/j.neuron.2018.10.021

61. Kunst, M. *et al.* A Cellular-Resolution Atlas of the Larval Zebrafish Brain. *Neuron* (2019). doi:10.1016/j.neuron.2019.04.034

62. Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M. & Keller, P. J. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat. methods* | **10**, 413 (2013).

63. Pisharath, H. & Parsons, M. J. Nitroreductase-mediated cell ablation in transgenic zebrafish embryos. *Methods Mol. Biol.* **546**, 133–143 (2009).

64. dal Maschio, M., Donovan, J. C., Helmbrecht, T. O. & Baier, H. Linking Neurons to Network Function and Behavior by Two-Photon Holographic Optogenetics and Volumetric Imaging. *Neuron* **94**, 774-789.e5 (2017).

65. Förster, D., Kramer, A., Baier, H. & Kubo, F. Optogenetic precision toolkit to reveal form, function and connectivity of single neurons. *Methods* **150**, 42–48 (2018).

66. Kalueff, A. V. *et al.* Towards a Comprehensive Catalog of Zebrafish Behavior 1.0 and Beyond. *Zebrafish* **10**, 70 (2013).

67. Young, A. P., Jackson, D. J. & Wyeth, R. C. A technical review and guide to RNA

fluorescence in situ hybridization. *PeerJ* **8**, (2020).

68. Baker, M. RNA imaging in situ. *Nat. Methods* **9**, 787–790 (2012).

69. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science (80-.).* **353**, 78–82 (2016).

70. Moffitt, J. R. *et al.* High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proc. Natl. Acad. Sci.* **113**, 11046–11051 (2016).

71. Shah, S., Lubeck, E., Zhou, W. & Cai, L. In Situ Transcription Profiling of Single Cells Reveals Spatial Organization of Cells in the Mouse Hippocampus. *Neuron* **92**, 342–357 (2016).

72. Lee, J. H. *et al.* Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat. Protoc. 2015 103* **10**, 442–458 (2015).

73. Dowling, J. E. The retina : an approachable part of the brain. 355 (2012).

74. Burrill, J. D. & Easter, S. S. Development of the retinofugal projections in the embryonic and larval zebrafish (Brachydanio rerio). *J. Comp. Neurol.* **346**, 583–600 (1994).

75. Baier, H. Synaptic Laminae in the Visual System: Molecular Mechanisms Forming Layers of Perception. doi:10.1146/annurev-cellbio-101011-155748

76. Baier, H. & Wullimann, M. F. Anatomy and function of retinorecipient arborization fields in zebrafish. *J. Comp. Neurol.* **529**, 3454–3476 (2021).

77. Lawrence, P. M. & Studholme, K. M. Retinofugal projections in the mouse. *J. Comp. Neurol.* **522**, 3733–3753 (2014).

78. EM, M. *et al.* Diverse Central Projection Patterns of Retinal Ganglion Cells. *Cell Rep.* **18**, 2058–2072 (2017).

79. Gaillard, F., Karten, H. J. & Sauvé, Y. Retinorecipient areas in the diurnal murine rodent Arvicanthis niloticus: A disproportionally large superior colliculus. *J. Comp. Neurol.* **521**, 1699–1726 (2013).

80. Scalia, F., Rasweiler, J. J. & Danias, J. Retinal projections in the short-tailed fruit bat, Carollia perspicillata, as studied using the axonal transport of cholera toxin B subunit: Comparison with mouse. *J. Comp. Neurol.* **523**, 1756–1791 (2015).

81. Gioanni, H., Rey, J., Villalobos, J., Richard, D. & Dalbera, A. Optokinetic nystagmus in the pigeon (Columba livia) II. Role of the pretectal nucleus of the accessory optic system (AOS). *Exp. Brain Res.* **50**, 237–247 (1983).

82. Deichler, A. *et al.* The nucleus pretectalis principalis: A pretectal structure hidden in the mammalian thalamus. *J. Comp. Neurol.* (2018). doi:10.1002/cne.24540

83. Masseck, O. A. & Hoffmann, K. P. Comparative neurobiology of the optokinetic

reflex. in *Annals of the New York Academy of Sciences* **1164**, 430–439 (2009).

84. Collewijn, H. Oculomotor areas in the rabbit's midbrain and pretectum. *J. Neurobiol.* **6**, 3–22 (1975).

85. Kerschensteiner, D. & Guido, W. Organization of the dorsal lateral geniculate nucleus in the mouse. *Vis. Neurosci.* **34**, E008 (2017).

86. Robles, E., Laurell, E. & Baier, H. The retinal projectome reveals brain-area-specific visual representations generated by ganglion cell diversity. *Curr. Biol.* **24**, 2085–2096 (2014).

87. Stoeckli, E. T. Understanding axon guidance: Are we nearly there yet? *Dev.* **145**, (2018).

88. Chédotal, A. & Richards, L. J. Wiring the brain: The biology of neuronal guidance. *Cold Spring Harb. Perspect. Biol.* **2**, (2010).

89. Franze, K. The mechanical control of nervous system development. *Development* **140**, 3069–77 (2013).

90. Matsuzaki, F. & Shitamukai, A. Cell Division Modes and Cleavage Planes of Neural Progenitors during Mammalian Cortical Development. *Cold Spring Harb. Perspect. Biol.* **7**, (2015).

91. Beattie, R. & Hippenmeyer, S. Mechanisms of radial glia progenitor cell lineage progression. *FEBS Lett.* **591**, 3993–4008 (2017).

92. Hobert, O. Terminal Selectors of Neuronal Identity. *Curr. Top. Dev. Biol.* **116**, 455–475 (2016).

93. Lois, C., García-Verdugo, J. M. & Alvarez-Buylla, A. Chain Migration of Neuronal Precursors. *Science (80-.).* **271**, 978–981 (1996).

94. Sawada, M. & Sawamoto, K. Neuronal migration in the postnatal brain. *Cell. Migr. Form. Axons Dendrites* 465–478 (2020). doi:10.1016/B978-0-12-814407-7.00021-3

95. Velasco, S. *et al.* A Multi-step Transcriptional and Chromatin State Cascade Underlies Motor Neuron Programming from Embryonic Stem Cells. *Cell Stem Cell* **20**, 205-217.e8 (2017).

96. Aberle, H. Axon guidance and collective cell migration by substrate-derived attractants. *Front. Mol. Neurosci.* **12**, 148 (2019).

97. Goodhill, G. J. Can Molecular Gradients Wire the Brain? *Trends Neurosci.* **39**, 202–211 (2016).

98. Van Horck, F. P. G., Weinl, C. & Holt, C. E. Retinal axon guidance: novel mechanisms for steering. *Curr. Opin. Neurobiol.* **14**, 61 (2004).

99. Kim, B. Evolutionarily conserved and divergent functions for cell adhesion molecules in neural circuit assembly. *J. Comp. Neurol.* **527**, 2061–2068 (2019).

100. Sanes, J. R. & Yamagata, M. Many Paths to Synaptic Specificity. *http://dx.doi.org/10.1146/annurev.cellbio.24.110707.175402* **25**, 161–195 (2009).

101. Cameron, S. & McAllister, A. K. Immunoglobulin-like receptors and their impact on wiring of brain synapses. *Annu. Rev. Genet.* **52**, 567–590 (2018).

102. Luhmann, H. J. *et al.* Spontaneous neuronal activity in developing neocortical networks: From single cells to large-scale interactions. *Front. Neural Circuits* **10**, 40 (2016).

103. Feller, M. B. Retinal waves are likely to instruct the formation of eye-specific retinogeniculate projections. *Neural Development* **4**, 24 (2009).

104. Torborg, C. L. & Feller, M. B. Spontaneous patterned retinal activity and the refinement of retinal projections. *Progress in Neurobiology* **76**, 213–235 (2005).

105. Zhang, R. W., Li, X. Q., Kawakami, K. & Du, J. L. Stereotyped initiation of retinal waves by bipolar cells via presynaptic NMDA autoreceptors. *Nat. Commun.* **7**, 12650 (2016).

106. Le Vay, S., Wiesel, T. N. & Hubel, D. H. The development of ocular dominance columns in normal and visually deprived monkeys. *J. Comp. Neurol.* **191**, 1–51 (1980).

107. Issa, N. P., Trachtenberg, J. T., Chapman, B., Zahs, K. R. & Stryker, M. P. The critical period for ocular dominance plasticity in the Ferret's visual cortex. *J. Neurosci.* **19**, 6965–6978 (1999).

108. Easter, S. S. & Nicola, G. N. The Development of Eye Movements in the Zebrafish (Danio rerio). *Dev. Psychobiol.* **31**, 267–276 (1997).

109. Bilotta, J. Effects of abnormal lighting on the development of zebrafish visual behavior. *Behav. Brain Res.* **116**, 81–7 (2000).

110. Portugues, R. & Engert, F. The neural basis of visual behaviors in the larval zebrafish. *Current Opinion in Neurobiology* **19**, 644–647 (2009).

111. Easter, S. S. & Nicola, G. N. The development of vision in the zebrafish (Danio rerio). *Dev. Biol.* **180**, 646–663 (1996).

112. Mercer, T. R. *et al.* The Human Mitochondrial Transcriptome. *Cell* **146**, 645–658 (2011).

113. Phillips, J. W. *et al.* A repeated molecular architecture across thalamic pathways. *Nat. Neurosci.* **22**, 1925–1935 (2019).

114. Anneser, L. *et al.* The neuropeptide Pth2 dynamically senses others via mechanosensation. *Nat. 2020 5887839* **588**, 653–657 (2020).

115. Kay, J. N., Finger-Baier, K. C., Roeser, T., Staub, W. & Baier, H. Retinal ganglion cell genesis requires lakritz, a zebrafish atonal homolog. *Neuron* **30**, 725–736 (2001).

116. Freytag, S., Tian, L., Lönnstedt, I., Ng, M. & Bahlo, M. Comparison of clustering tools in R for medium-sized 10x Genomics single-cell RNA-sequencing data. *F1000Research* **7**, 1297 (2018).

117. Zhang, L. & Zhang, S. Comparison of computational methods for imputing single-cell RNA-sequencing data. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 1–1 (2018). doi:10.1109/TCBB.2018.2848633

118. Schiebinger, G., Shu, J., Jaenisch, R., Regev, A. & Lander, E. S. Optimal-Transport Analysis of Single-Cell Gene Expression Identifies Developmental Trajectories in Reprogramming. *Cell* **176**, 928-943.e22 (2019).

119. Cheng, S. *et al.* Vision is Required for Cell Type Specification in the Visual Cortex. *bioRxiv* 2021.08.10.455824 (2021). doi:10.1101/2021.08.10.455824

120. Vasconcelos, F. F. & Castro, D. S. Transcriptional control of vertebrate neurogenesis by the proneural factor ascll. *Front. Cell. Neurosci.* **8**, 412 (2014).

121. Velkey, M. M. & O'Shea, K. S. Expression of Neurogenin 1 in mouse embryonic stem cells directs the differentiation of neuronal precursors and identifies unique patterns of down-stream gene expression. *Dev. Dyn.* **242**, 230–253 (2013).

122. Mueller, T. & Wullimann, M. F. BrdU-, neuroD (nrd)- and Hu-studies reveal unusual non-ventricular neurogenesis in the postembryonic zebrafish forebrain. *Mech. Dev.* **117**, 123–135 (2002).

123. Mueller, T. & Wullimann, M. F. Anatomy of neurogenesis in the early zebrafish brain. *Dev. Brain Res.* **140**, 137–155 (2003).

124. D'Souza, S. & Lang, R. A. Retinal ganglion cell interactions shape the developing mammalian visual system. *Dev.* **147**, (2020).

125. Antinucci, P., Folgueira, M. & Bianco, I. H. Pretectal neurons control hunting behaviour. *Elife* **8**, (2019).

126. Jones, A. R., Overly, C. C. & Sunkin, S. M. The Allen Brain Atlas: 5 years and beyond. *Nat. Rev. Neurosci. 2009 1011* **10**, 821–828 (2009).

127. Marquart, G. D. *et al.* High-precision registration between zebrafish brain atlases using symmetric diffeomorphic normalization. *Gigascience* **6**, 1 (2017).

128. Fast release clones: a high throughput expression analysis – ScienceOpen. Available at: https://www.scienceopen.com/document?vid=86459e8b-890d-46a2-b5b8-f84bfce42a1c. (Accessed: 19th January 2022)

129. Li, Y., Fitzpatrick, D. & White, L. E. The development of direction selectivity in ferret visual cortex requires early visual experience. *Nat. Neurosci. 2006 95* **9**, 676–681 (2006).

130. Fagiolini, M., Pizzorusso, T., Berardi, N., Domenici, L. & Maffei, L. Functional postnatal development of the rat primary visual cortex and the role of visual experience: Dark rearing and monocular deprivation. *Vision Res.* **34**, 709–720

(1994).

131. McLaughlin, T., Torborg, C. L., Feller, M. B. & O'Leary, D. D. M. Retinotopic Map Refinement Requires Spontaneous Retinal Waves during a Brief Critical Period of Development. *Neuron* **40**, 1147–1160 (2003).

132. Niell, C. M. & Smith, S. J. Functional imaging reveals rapid development of visual response properties in the zebrafish tectum. *Neuron* **45**, 941–951 (2005).

133. Stout, R. P. & Graziadei, P. P. C. Influence of the olfactory placode on the development of the brain in Xenopus laevis (Daudin): I. Axonal growth and connections of the transplanted olfactory placode. *Neuroscience* **5**, 2175–2186 (1980).

134. Code, R. A., Durham, D. & Rubel, E. W. Effects of cochlea removal on GABAergic terminals in nucleus magnocellularis of the chicken. *J. Comp. Neurol.* **301**, 643–654 (1990).

135. WIESEL, T. N. & HUBEL, D. H. EFFECTS OF VISUAL DEPRIVATION ON MORPHOLOGY AND PHYSIOLOGY OF CELLS IN THE CAT'S LATERAL GENICULATE BODY. *https://doi.org/10.1152/jn.1963.26.6.978* **26**, 978–993 (1963).

136. WIESEL, T. N. & HUBEL, D. H. SINGLE-CELL RESPONSES IN STRIATE CORTEX OF KITTENS DEPRIVED OF VISION IN ONE EYE. *https://doi.org/10.1152/jn.1963.26.6.1003* **26**, 1003–1017 (1963).

137. Mooney, R., Penn, A. A., Gallego, R. & Shatz, C. J. Thalamic Relay of Spontaneous Retinal Activity Prior to Vision. *Neuron* **17**, 863–874 (1996).

138. Spitzer, N. C. Electrical activity in early neuronal development. *Nat. 2006 4447120* **444**, 707–712 (2006).

139. Yap, E.-L. & Greenberg, M. E. Activity-Regulated Transcription: Bridging the Gap between Neural Activity and Behavior. *Neuron* **100**, 330–348 (2018).

140. Greenberg, M. E., Thompson, M. A. & Sheng, M. Calcium regulation of immediate early gene transcription. *J. Physiol.* **86**, 99–108 (1992).

141. Altar, C. A. *et al.* Electroconvulsive Seizures Regulate Gene Expression of Distinct Neurotrophic Signaling Pathways. *J. Neurosci.* **24**, 2667–2677 (2004).

142. Thoenen, H. Neurotrophins and Neuronal Plasticity. *Science (80-.).* **270**, 593–598 (1995).

143. Selleck, S. B. & Steller, H. The influence of retinal innervation on neurogenesis in the first optic ganglion of drosophila. *Neuron* **6**, 83–99 (1991).

144. Huang, Z. & Kunes, S. Signals transmitted along retinal axons in Drosophila: Hedgehog signal reception and the cell circuitry of lamina cartridge assembly. *Development* **125**, 3753–3764 (1998).

145. Huang, Z., Shilo, B. Z. & Kunes, S. A Retinal Axon Fascicle Uses Spitz, an EGF

Receptor Ligand, to Construct a Synaptic Cartridge in the Brain of Drosophila. *Cell* **95**, 693–703 (1998).

146. Huang, Z. & Kunes, S. Hedgehog, Transmitted along Retinal Axons, Triggers Neurogenesis in the Developing Visual Centers of the Drosophila Brain. *Cell* **86**, 411–422 (1996).

147. Von Bartheld, C. S., Byers, M. R., Williams, R. & Bothwell, M. Anterograde transport of neurotrophins and axodendritic transfer in the developing visual system. *Nat. 1996 3796568* **379**, 830–833 (1996).

148. Menna, E., Cenni, M. C., Naska, S. & Maffei, L. The anterogradely transported BDNF promotes retinal axon remodeling during eye specific segregation within the LGN. *Mol. Cell. Neurosci.* **24**, 972–983 (2003).

149. Liu, H., Lu, M. & Guthrie, K. M. Anterograde trafficking of neurotrophin-3 in the adult olfactory system in vivo. *Exp. Neurol.* **241**, 125 (2013).

150. Antinucci, P., Folgueira, M. & Bianco, I. H. Pretectal neurons control hunting behaviour. *Elife* **8**, (2019).

151. Cerveny, K. L., Varga, M. & Wilson, S. W. Continued growth and circuit building in the anamniote visual system. *Dev. Neurobiol.* **72**, 328–345 (2012).

152. Boucher, S.-E. M. & Hitchcock, P. F. Insulin-Related Growth Factors Stimulate Proliferation of Retinal Progenitors in the Goldfish. *J. Comp. Neurol* **394**, 386–394 (1998).

153. Otteson, D. C., Cirenza, P. F. & Hitchcock, P. F. Persistent neurogenesis in the teleost retina: evidence for regulation by the growth-hormone/insulin-like growth factor-I axis. *Mech. Dev.* **117**, 137–149 (2002).

154. Yamaguchi, M. *et al.* Histone deacetylase 1 regulates retinal neurogenesis in zebrafish by suppressing Wnt and Notch signaling pathways. *Development* **132**, 3027–3043 (2005).

155. Denayer, T. *et al.* Canonical Wnt Signaling Controls Proliferation of Retinal Stem/Progenitor Cells in Postembryonic Xenopus Eyes. *Stem Cells* **26**, 2063–2074 (2008).

156. Locker, M. *et al.* Hedgehog signaling and the retina: insights into the mechanisms controlling the proliferative properties of neural precursors. *Genes Dev.* **20**, 3036–3048 (2006).

157. Stephens, W. Z., Senecal, M., Nguyen, M. & Piotrowski, T. Loss of adenomatous polyposis coli (apc) results in an expanded ciliary marginal zone in the zebrafish eye. *Dev. Dyn.* **239**, 2066–2077 (2010).

158. Neumann, C. J. & Nuesslein-Volhard, C. Patterning of the zebrafish retina by a wave of sonic Hedgehog activity. *Science (80-.).* **289**, 2137–2139 (2000).

159. Masai, I., Yamaguchi, M., Tonou-Fujimori, N., Komori, A. & Okamoto, H. The

hedgehog-PKA pathway regulates two distinct steps of the differentiation of retinal ganglion cells: the cell-cycle exit of retinoblasts and their neuronal maturation. *Development* **132**, 1539–1553 (2005).

160. Cline, H. T. Activity-dependent plasticity in the visual systems of frogs and fish. *Trends Neurosci.* **14**, 104–111 (1991).

161. Tan, L. *et al.* Changes in genome architecture and transcriptional dynamics progress independently of sensory experience during post-natal brain development. *Cell* **184**, 741-758.e17 (2021).

APPENDIX

CURRICULUM VITAE

UNIVERSITY EDUCATION

2016 - 2023	Ph.D. at the Department of Genes – Circuits – Behavior Max Planck Institute of Neurobiology Advisor: Prof. Dr. Herwig Baier Thesis: Genetic underpinnings of neural circuit assembly Graduate schools: IMPRS, GSN, LMU
2014 – 2016	M.Sc. at the Department of Life Sciences Ben-Gurion University (Beer-Sheva, Israel) Advisor: Dr. Natalie Elia
2011 – 2014	B.Sc. at the Department of Life Sciences Ben-Gurion University (Beer-Sheva, Israel)

AWARDS

2016	Dean's award for academic achievements and excellence in research

LIST OF PUBLICATIONS

Johannes M. Kappel, Dominique Förster, Katja Slangewal, Inbal Shainer, Fabian Svara, Joseph C. Donovan, **Shachar Sherman**, Michal Januszewski, Herwig Baier, Johannes Larsch. Visual recognition of social signals by a tectothalamic neural circuit. Nature 608, 146–152 (2022). https://doi.org/10.1038/s41586-022-04925-5

Dovrat D*, Dahan D*, **Sherman S***, Tsirkas I, Elia N, Aharoni A. A Live-Cell Imaging Approach for Measuring DNA Replication Rates. Cell Rep. 2018 Jul 3;24(1):252-258. doi: 10.1016/j.celrep.2018.06.018. PMID: 29972785.
*These authors contributed equally

Ott C, Nachmias D, Adar S, Jarnik M, **Sherman S**, Birnbaum RY, Lippincott-Schwartz J, Elia N. VPS4 is a dynamic component of the centrosome that regulates centrosome localization of γ-tubulin, centriolar satellite stability and ciliogenesis. Sci Rep. 2018 Feb 20;8(1):3353. doi: 10.1038/s41598-018-21491-x. PMID: 29463826; PMCID: PMC5820263.

Sherman S, Kirchenbuechler D, Nachmias D, Tamir A, Werner S, Elbaum M, Elia N. Resolving new ultrastructural features of cytokinetic abscission with soft-X-ray cryo-tomography. Sci Rep. 2016 Jun 10;6:27629. doi: 10.1038/srep27629. PMID: 27282220; PMCID: PMC4901327.

Jamasbi J, Megens RT, Bianchini M, Uhland K, Münch G, Ungerer M, **Sherman S**, Faussner A, Brandl R, John C, Buchner J, Weber C, Lorenz R, Elia N, Siess W. Cross-Linking GPVI-Fc by Anti-Fc Antibodies Potentiates Its Inhibition of Atherosclerotic Plaque- and Collagen-Induced Platelet Activation. JACC Basic Transl Sci. 2016 Apr 25;1(3):131-142.. doi: 10.1016/j.jacbts.2016.03.008. PMID: 27766315; PMCID: PMC5063538.

Gershony O, **Sherman S**, Adar S, Segal I, Nachmias D, Goliand I, Elia N. Measuring abscission spatiotemporal dynamics using quantitative high-resolution microscopy. Methods Cell Biol. 2017;137:205-224. doi: 10.1016/bs.mcb.2016.03.032. Epub 2016 Jul 9. PMID: 28065306.

Sherman S, Nachmias D, Elia N. A simple, straightforward correlative live-cell-imaging-structured-illumination-microscopy approach for studying organelle dynamics. Microsc Res Tech. 2015 Sep;78(9):777-83. doi: 10.1002/jemt.22540. Epub 2015 Jul 15. PMID: 26178911.

Jamasbi J, Megens RT, Bianchini M, Münch G, Ungerer M, Faussner A, **Sherman S**, Walker A, Goyal P, Jung S, Brandl R, Weber C, Lorenz R, Farndale R, Elia N, Siess W. Differential Inhibition of Human Atherosclerotic Plaque-Induced Platelet Activation by Dimeric GPVI-Fc and Anti-GPVI Antibodies: Functional and Imaging Studies. J Am Coll Cardiol. 2015 Jun 9;65(22):2404-15. doi: 10.1016/j.jacc.2015.03.573. PMID: 26046734; PMCID: PMC4452546.

Submitted manuscripts (under peer-review)

A single-cell resolution gene expression atlas of the larval zebrafish brain
Inbal Shainer, Enrico Kuehn, Eva Laurell, Mariam Al Kassar, Nouwar Mokayes, **Shachar Sherman**, Johannes Larsch, Michael Kunst, Herwig Baier
bioRxiv 2022.02.11.479024; doi: https://doi.org/10.1101/2022.02.11.479024

Retinal input influences pace of neurogenesis but not cell-type configuration of the visual forebrain
Shachar Sherman, Koichi Kawakami, Herwig Baier
bioRxiv 2021.11.15.468630; doi: https://doi.org/10.1101/2021.11.15.468630

DECLARATION OF AUTHOR CONTRIBUTION

Herwig Baier[1] and Shachar Sherman[1-3] designed this study. Shachar Sherman carried out experiments and analysis.

[1] Max Planck Institute of Neurobiology, Department of Genes – Circuits – Behavior, Am Klopferspitz 18, D-82152, Martinsried, Germany

[2] IMPRS for Molecular Life Sciences, Am Klopferspitz 18, 82152 Martinsried, Germany

[3] Graduate School of Systemic Neurosciences, LMU BioCenter, Großhaderner Str. 2, D-82152 Planegg-Martinsried

Shachar Sherman

Prof. Dr. Herwig Baier