Exploring Neuronal Heterogeneity in the Drosophila Nervous System with Novel Neurotechnologies

by

Ye Li

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cell and Developmental Biology)
in the University of Michigan
2019

Doctoral Committee:

Associate Professor Bing Ye, Chair
Assistant Professor Dawen Cai
Associate Professor Cheng-Yu Lee
Associate Professor Maureen A. Sartor

Ye Li

cnliye@umich.edu

ORCID iD: 0000-0002-8647-384X

Dedication

This thesis is dedicated to my parents, Jianjun Li and Ruilian Xing, and my wife, Meiling Zhao, who give me love and strength every step of the way.

Acknowledgements

I would like to thank my mentor, Dr. Dawen Cai for his continuous support and guidance. This has been a long journey, and it would be impossible for me to carry out all components in my thesis study without his encouragements, constructive criticisms, and most importantly, trust. This is a mutual trust that we developed together through all ups and downs in this small scientific endeavor, something that I treasured greatly. From so many things that I have earned from Dawen, the most important one was "Trust yourself, don't give up." I made more mistakes than success, encountered more setbacks than progress, but every time we figured out a way to get up and approach our goals a bit closer. We share thoughts, scrutinize experimental designs, crunch numbers on the go and inspire each other whenever we can. I am deeply thankful for the training by Dawen.

Dawen created an inclusive lab environment with high scientific standards where everybody in the lab supports each other with their best capacity. I have been benefited so much from numerous sincere suggestions, arguments, and debates from the best band of lab people I could ever expect. Yimeng, Marya, Tiffany, Maggie, HJ, Erica, Fred, Nigel, Logan, Jingqun, Danny, Douglas – thank you all for your help.

I would like to thank my thesis committee members, Dr. Bing Ye, Dr. Cheng-Yu Lee, and Dr. Maureen Sartor. Every meeting has been a growth point for me, and I am sincerely grateful to all of your advices! Your patience and support helped me greatly. I would especially thank Bing, who recruited me to Michigan six years ago, taught me one of the most memorable Developmental Biology classes I ever had (fascinating patterning of neuron arbors!) five years

ago, suggested me to look at serotonergic neurons about four years ago, and supported me through the entire process. Thank you for helping me!

I would also like to thank the Department of Cell and Developmental Biology who has given me the best support as I can imagine from a department. We have the best staff who are always kind and helpful, a tight-knit community in which everyone is supporting each other, a strict but helping graduate program lead by Dr. Roman Giger and Dr. Scott Barolo over the years, and a strongly trainee-focused leadership chaired by Dr. Pierre Coulombe.

Finally, I would thank my family. My parents have been unconditionally supporting me in every stage of my life with their best abilities, without which I could never achieve what I have reached today. My wife has always stood by my side, and we have shared so many valuable moments together. I am very grateful to our extended families on both sides, who have been caring about us for as long as we could remember. Also, my belated thanks to grandma and grandpa, who must be delightful seeing me now writing the finishing pages of this chapter of my career.

Table of Contents

List of Tables

List of Figures

Abstract

The complex nervous system is built upon a vastly heterogeneous population of neurons. In order to decipher how the nervous system operates, it is critical to understand all aspects of neuronal properties such as morphology, lineage, electrophysiology and molecular identity, etc. Canonical neuronal subtype classification is often based on features reflected in one of these attributes. However, canonically defined neuronal subtypes normally compose of individuals that are heterogeneous in other attributes. It is therefore important to study the same neuron based on a collective cohort of properties. In this thesis study, I studied the lineage composition, morphology patterning and molecular heterogeneity of the serotonergic neurons in the fruit fly *Drosophila melanogaster*. I developed a series of novel transgenic tools, collectively called Bitbow, which are capable of generating up to tens of thousands of unique fluorescent barcodes to unambiguously label hundreds of lineages or individual neurons in the same brain. My results indicated that most of the serotonergic neurons arises from distinct lineages. Combining with Expansion Microscopy and multispectral neuronal tracing, a morphological map of serotonergic neurons in the ventral nerve cord was reconstructed from a single Bitbow fly. Using scRNAseq techniques, I found profound molecular heterogeneity of serotonergic neurons, characterized by their differentially expressed genes of GPCRs, ion channels, neurotransmitters, transcription factors and so on. My thesis has provided new methodologies to better define neuronal subtypes with multiple modalities, and accumulated knowledge to allow more precise investigations and manipulations of *Drosophila* serotonergic neurons in future studies.

Chapter 1. Introduction

1.1 Neuronal subtypes: heterogeneous functional units that build up the nervous system

Nervous systems, simple or complex across the span of the evolutionary tree, rely on precise development and wiring of vastly diverse types of neurons. As fundamental functional units in the system, neurons contain common structural features, such as axons, dendrites, somas and synapsis, and common functional features, such as action potentials, gated ion flux, and excitatory and inhibitory connections. Together these basic features ensure proper processing and transmitting of information flows across the nervous system. However, neurons are vastly different in the extent and combinations of how they adopt these features to perform desired functions, resulting in numerous neuronal subtypes. In order to obtain knowledge regarding how the nervous system is operating as an integral entity, it is critical to understand the diversity in neuronal subtypes.

Neurons can be defined through categorical observations in their morphology, lineage, molecular identity, connectivity, electrophysiological properties and so on (Fig 1.1, Kepecs & Fishell, 2014). Since the time when neuron doctrine was established by the studies of Cajal, Golgi and many others in various species (Golgi, 1885; y Cajal, 1888), a solid body of work has been created on the basic morphological description of neurons. These morphological descriptions are direct and indictive. How do major neurites branch from the soma? How dense are the axonal and dendritic projections? Which specific brain regions do the projection arbors cover? Does the neuron have long-distance projections or local projections? These questions form a set of critical categorical survey that builds the foundation to determine neuronal

subtypes. It is also important to understand from which group of neural stem cells that certain neurons are born, i.e. neurons' lineage identities. In mice, excitatory and inhibitory interneurons have been found to be derived from distinct parts of the developing neocortex (Tan & Shi, 2013), and disruptions in progenitor cells lead to changes in cell type composition (Sultan & Shi, 2018), damaging normal functions in the nervous system. Molecular identity also played a strong role in defining neuron types. With proper probing at both the transcription and translation levels (Bhattacherjee et al., 2019; Greig, Woodworth, Galazo, Padmanabhan, & Macklis, 2013; Molyneaux, Arlotta, Menezes, & Macklis, 2007), neurons can be consistently categorized based on their molecular profiles, which also pave the way of tracking and manipulating specific subtypes transgenically (Jenett et al., 2012; Taniguchi et al., 2011). Finally, electrophysiological properties of neurons brings in critical knowledge of how each type of neuron is potentially exerting its role to participate in a function circuitry (Fuzik et al., 2016; Knoblich, Huang, Zeng, & Li, 2019).

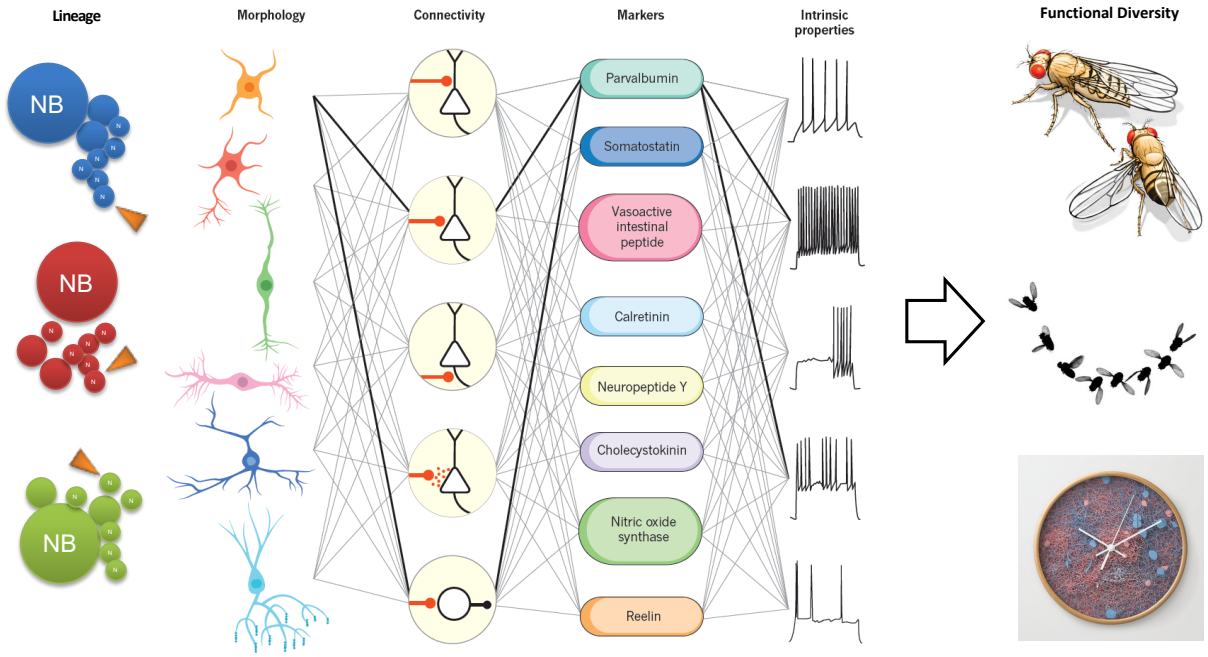


Fig 1.1 Neuronal subtypes can be defined by categorical observations of neuron properties
Neuronal diversities in lineage, morphology, connectivity, molecular identity and electrophysiological properties contribute to the heterogeneity in the nervous system, leading to profound functional diversity. (Modified from Kepecs & Fishell, 2014)

Currently, the foundation of our knowledge regarding the heterogeneity in neuronal subtypes is built upon parallel information collected independently in each of these categories. However, it is becoming more and more evident that functional/behavioral responses from single-categorically defined neurons in an experiment can be very heterogeneous (Cembrowski & Spruston, 2019). This suggests that there are more differences among the cells in each defined group of neurons. To more accurately address these differences, one approach is to find new markers within each category to gain finer definition. This approach is greatly limited by the profiling depth and throughput of available methodologies. Alternatively, if features across multiple categories can be associated together, i.e. generating cross-modality definitions, the subtypes could be more precisely defined utilizing all information collected by existing and novel neurotechnologies. With more accurate subtype definitions, new knowledge and methodology can be obtained to deepen our understanding about how these diverse types of neurons contribute to a uniform and complex entity.

1.2 The *Drosophila* nervous system consists of precisely developed and organized neuron subtypes

The central nervous system (CNS) of fruit fly *Drosophila melanogaster* is an excellent model for understanding neuronal diversity and the potential regulation mechanisms underlying the neuronal diversity. Consisting of 10,000 - 12,000 neurons in larval stages, and ~150,000 neurons in adult (Ohyama et al., 2015), the fly CNS contains vastly different types of neurons to build complex circuits and perform various neurological functions. All neurons are precisely generated through two waves of neurogenesis. The first wave happens in the embryonic stage when a relatively small set of neurons are produced and organized to support activities in the larva stage. After a short period of stem cell quiescence, the second wave of neurogenesis

persists throughout the larval stage to produce ~90% neurons responsible for the adult CNS (Homem & Knoblich, 2012). A diverse population of neurons are generated during these two waves from a rather small pool of neural stem cells, the neuroblasts (NB). There is estimated to be ~200 NBs in the central brain (K. Ito & Hotta, 1992), ~1000 in the two optic lobes (Hofbauer & Campos-Ortega, 1990) and ~800 in the ventral nerve cord (Birkholz, Rickert, Berger, Urbach, & Technau, 2013). NBs in the fly brain are defined in two distinct types based on their proliferation schemes. Type I NBs uses a relatively simple scheme, where the stem cell goes through a limited number of asymmetric divisions to each time generate a progeny named ganglion mother cell (GMC) and self-renew to maintain its stem cell property. Each GMC divides symmetrically once further to generate two neurons (or one neuron and one glia, or two glia). Because most of the NBs in the nervous system are Type I, this scheme produces the majority of neurons in the brain. Type II NBs uses an extra step in the process, where in each asymmetric division, instead of generating a GMC, the NB produces an intermediate neural progenitor (INP) and self-renew; each INP can produce multiple GMCs in their lifetime, and GMCs are still going through one symmetric division to produce two daughter cells. Although there are only 16 Type II NBs in the CNS, this developmental strategy produces a lot more neurons within the same lineage. Together with the Type I NBs, the two strategies ensure the generation of a large population of progeny and lay the foundation of neuronal heterogeneity in the CNS.

Although NBs undertake stereotypic proliferation strategies, they are all uniquely regulated and specified to produce diverse progenies. This is achieved through combinatorial molecular mechanisms with spatial, temporal and hemi-lineage specifications (Fig 1.2).

As early as 5 hours after-egg-lay (AEL), NBs in the embryos start to be patterned with spatially defined molecule combinations, mostly transcription factors, and these spatial factors play vital roles in defining NB identity. NBs in both central brain and VNC are influenced by spatial factors. In VNC, segment polarity genes such as engrailed (en) and gooseberry (gsb) determine anterior and posterior compartments, respectively, within each neuromere (Torsten Bossing, Udolph, Doe, & Technau, 1996). Dorsal/ventral patterning genes such as Msh, Ind, and Vnd exert their function through "stripes" of expression domains running across neuromeres from anterior to posterior (Urbach & Technau, 2003; Urbach, Volland, Seibert, & Technau, 2006). These two sets of orthogonal molecular signals create a meshed spatial determination pattern. Additional "patches" of spatial factors, such as huckebein (hkb) and Dachshund (dac), are expressed in unique patterns across various segments. Together these spatial factor combinations make it possible to stereotypically determine NBs located at specific positions in the developing brain.

However, spatial determination alone is not sufficient to heterogeneously determine hundreds of thousands of neurons in the *Drosophila* CNS. Temporal and hemi-lineage specifications are critical to further diversify the population within each lineage. Hemi-lineage determination is a straightforward and effective strategy to assign different identities for neurons of the same lineage. In the last step of neurogenesis, GMCs produce two daughter cells with similar morphological traits but contrasting molecular determinants. Most notable is the Notch signaling components, where one of the two daughter cells contains numb, the inhibitor of Notch, hence being "Notch-off", and the other does not hence being "Notch-on" (Frise, Knoblich, Younger-Shepherd, Jan, & Jan, 1996; Skeath & Doe, 1998; Spana, Kopczynski, Goodman, & Doe, 1995). This binary mechanism effectively groups all progeny in the same

lineage into two sets, the Notch-off hemi-lineage and the Notch-on hemi-lineage (Bardin, Le Borgne, & Schweisguth, 2004). Neurons of the same hemi-lineage exert similar developmental and terminal properties, such as sharing the same neurite tract, forming similar projection domains, etc., and the two hemi-lineages from the same NB clearly have differentially projecting neurons although all the cell bodies are tightly clustered with each other (Harris, Pfeiffer, Rubin, & Truman, 2015; Truman, Moats, Altman, Marin, & Williams, 2010).

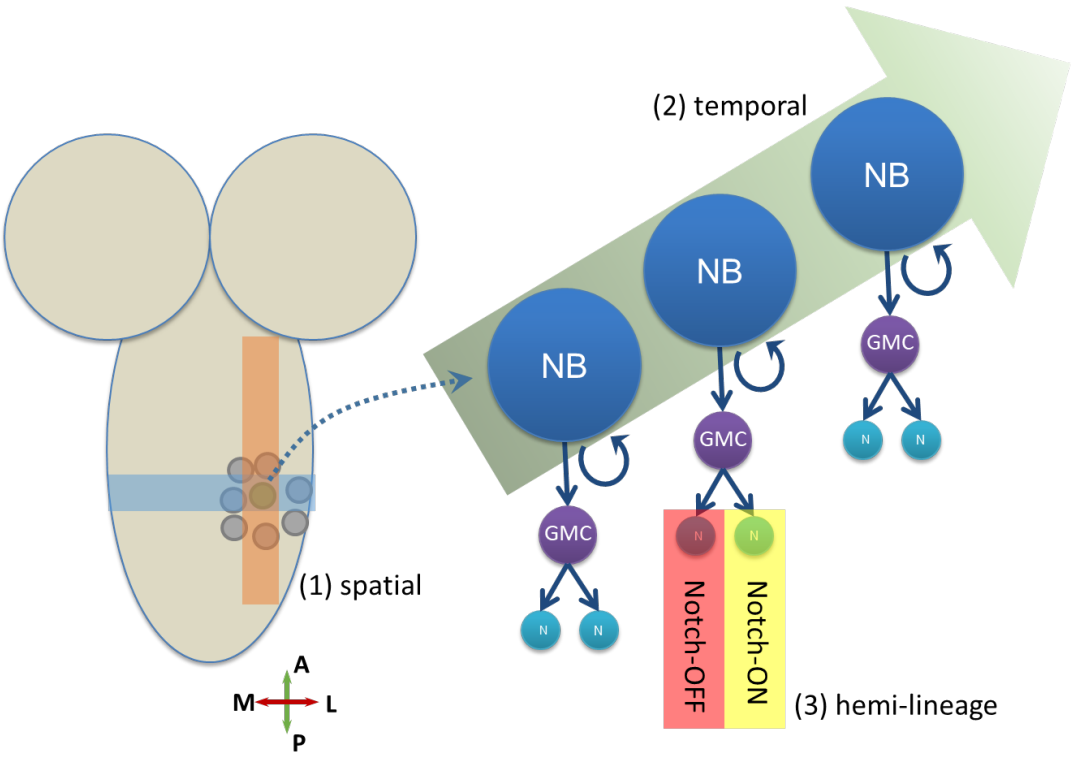


Fig 1.2 Combinatorial molecular specifications permit diverse neuron subtype generation in developing *Drosophila* Ventral Nerve Cord (VNC)
Neurogenesis in *Drosophila* VNC is precisely regulated by specific combinatorial molecular factors patterned in spatial-, temporal- and hemi-lineage-dependent manners. Shaded boxes in different colors represent different molecule combinations. NB, neuroblast. GMC, ganglion mother cell. N, neuron. A-P, Anterior-Posterior axis. M-L, Medial-Lateral axis.

As the third and most dynamic patterning influence, temporal factors have been found to play vital roles in generating diverse neuronal types across the CNS, such as in VNC lineages, central brain type II lineages as well as optic lobe lineages (Bayraktar & Doe, 2013; Isshiki, Pearson, Holbrook, & Doe, 2001; X. Li et al., 2013; Pearson & Doe, 2003; Suzuki, Kaido,

Takayama, & Sato, 2013). Cascades of transcription factors, such as Hb > Kr > Pdm > Caster in VNC NBs (Isshiki et al., 2001), D > Grh > Ey in Type II NBs (Bayraktar & Doe, 2013) and Hth > Klu > Ey > Slp > D > Tll in optic lobe NBs (X. Li et al., 2013), have made it possible to add an additional layer of specification to diversify the possible progeny outcomes from a single lineage. Transitions and overlapping between adjacent temporal factors further strengthen the power of temporal patterning to allow more developmental states of NBs, hence more neuronal types. Disruptions or delay of normal temporal transition in NBs lead to substantial changes of neuron subtype composition within a lineage, resulting in missing certain types and/or overpopulating of others (Brody & Odenwald, 2000; Isshiki et al., 2001; Pearson & Doe, 2003). Together, spatial, temporal and hemi-lineage determinations build up a multiplex and stereotypic intrinsic developmental program to lay the foundation of the neuron diversity in the *Drosophila* CNS.

Equally important but less understood are the extrinsic regulations in neuron shaping and wiring, which gives proper morphology and connectivity to neurons to exert proper functions in a circuit. Understand how different neurons look and wire together is critical to reveal the complexity of neuronal subtypes.

Drosophila CNS neurons have common morphological features with diversely flexible neurite organization. Most of CNS neurons are unipolar; somas (cell bodies) of these neurons contain single primary neurites. Branching off from the primary neurite, elaborative arbors form compartments of afferent and efferent natures. This basic morphological organization leads to clustering of somas at the peripheral of CNS which forms the cortex, and the clustering of arbors which forms the neuropil compartments (Larsen et al., 2009). Because of this morphological compartmentalization of somas and arbors, it is difficult to precisely describe neuron types solely

based on their soma location or arbor location. Somas of different types of neurons might be clustered tightly together (Lai, Awasaki, Ito, & Lee, 2008), and arbors at the same neuropil might be originated from various types of neurons (Fischbach, -F. Fischbach, & Dittrich, 1989; Hanesch, Fischbach, & Heisenberg, 1989). Accurate and complete reconstructions of neuronal morphology are essential to define neuronal subtypes.

Once neuron morphologies can be accurately described, well-guided hypothesis can be derived regarding their functional relevance. For one aspect, because the neuropil compartments are well characterized in the *Drosophila* brain (Kei Ito et al., 2014), and some of them are found to associate with specific neurological functions, e.g. mushroom body to memory (Pascual & Préat, 2001), antenna lobe to sensory decoding (Fishilevich & Vosshall, 2005), central complex to locomotion (Strauss, 2002), etc., morphological information regarding where the studied neuron targets its arbor onto will be predictive of its functional relevance. In another aspect, even if neurons are targeting the same neuropil, their functional differences can be proposed based on their morphological differences. For example in the adult optic lobe (Fischbach et al., 1989), some medulla neurons extend their neurites perpendicularly through multiple medulla layers (such as intrinsic columnar medulla neurons, Mi neurons), while some other ones whose neurite stay mainly in one layer but tangentially cover the entire span of medulla (such as tangential medulla neurons, Mt neurons). Both Mi and Mt neurons are innervating the Medulla only, but because of their morphological differences, they have drastically different functional predictions, where Mi neurons are thought to be responsible for relaying and processing visual information passing down in each visual column, while Mt neurons are hypothesized to be responsible for integration of information across columns. In summary, morphological descriptions of neuronal subtypes are important and highly indicative of neuronal functions.

1.3 Genetic approaches that enable identification of neuronal subtypes in *Drosophila*

Besides its amenable size of neuron population and circuitry-behavior complexity, *Drosophila* is an attractive research model also due to its large arsenal of available genetic tools. This makes thorough investigations approachable. This is an ongoing effort by the contributions from many generations of scientists, trailblazing from simple, direct prototypes to more and more advanced, complex systems that have been combinatorically built upon previous successes.

To obtain systematic knowledge of neuron types in the *Drosophila* nervous system, whether through a morphological, developmental or circuitry perspective, or the combinations of them, it is crucial to be able to consistently identify and target neuron types of interest. Before the prevalent use of fluorescent proteins, chemical staining (Blest, 1961; Holmes, 1943; y Cajal, 1910) and dye injections (T. Bossing & Technau, 1994) were the major methodologies to differentially label and track neurons. These methods provided high quality detail of neurons but were greatly limited by their technical complexity and throughput. More importantly, it is non-trivial to correlate morphologically determined neurons with their molecular features.

Fluorescent proteins, such as GFP, which are genetically encodable and optically observable, have revolutionized the paradigm of all branches of neuroscience. In *Drosophila*, the strength of fluorescent-protein-based investigation further stands out with the armory of powerful and delicate genetic designs.

Borrowing from a yeast transcription system, the Gal4-UAS transgenics permits specific labeling of molecularly defined population with great flexibility. This two-component system uses a driver-effector scheme, where the transcription activator Gal4 is under the control of a specific enhancer element to be the driver line, and effectors such as a fluorescent protein is

downstream of the UAS to be the reporter line. Because of the specific interactions between Gal4 and UAS, the driver-reporter pair will ensure faithful labeling of only the neurons targeted by the enhancer element in the driver. This allows great flexibility in studying neuronal types, especially with the tremendous resource of Gal4 transgenic library (Jenett et al., 2012) and reporter lines (Pfeiffer et al., 2010).

In many cases, further specification of a group of neurons within a determined population is necessary, especially when the group of interest cannot be simply defined by known enhancer elements, or the neuron population is too large to be optically resolved reliably. Genetic tools with stochastic characteristics, such as Flp-out (Evans et al., 2009; K. Ito, Awano, Suzuki, Hiromi, & Yamamoto, 1997) and MARCM (Lee & Luo, 1999) are able to fulfill these needs. As the core molecular machinery, recombinase flippase (Flp) can recognize specific DNA sequences named flippase recognition target (FRT), and either invert or remove sequences flanked by a pair of FRTs depending on the relative orientation of the two FRTs. Because the amount of flippase in cells can be externally controlled with transgenic lines such as heat-shock-promoter-driven flippase (hsFlp), the level of recombination can be flexibly regulated based on experimental needs. Flip-out and MARCM make it possible to label a subset of neurons within a defined group, and in some extreme cases, a single neuron (Lee & Luo, 1999).

As knowledge and tools advance, there is greater and greater need to not only follow on single or sparse neuronal targets, but to investigate neurons within their whole native circuitry, in a dense but resolvable manner. This is one of the initial rationales of multicolor labeling systems, such as Brainbow (Hadjieconomou et al., 2011; Hampel et al., 2011; Livet et al., 2007) and Multi-color Flip-out (MCFO) (Nern, Pfeiffer, & Rubin, 2015). Through careful genetic designs combining Gal4-UAS, Flp-FRT (or other recombinase systems such as Cre-Lox) systems and a

set of spectrally, antigenically distinguishable fluorescent proteins, multicolor labeling systems have achieved differential labeling of large groups of neurons that are previously unresolvable. Neurons of interest are still selectively targeted with specific genetic control such as enhancer-Gal4 drivers, and multicolor labeling combinations are randomly assigned to individual neurons within the targeted population, which make them distinguishable from each other. This has enabled critical advantages in both morphological and developmental studies. Detailed morphological information regarding cell body position, projection pattern, neurite density and so on can be learned efficiently in a population manner, plus knowledge about potential interactions among the members within the population. When tracing down the developmental processes by assigning multicolor labeling in a lineage-related manner, progeny of the same lineage can inherit the same label, which enables lineage tracing of these progeny to reveal their developmental dynamics and final contribution to the nervous system.

The multicolor genetic systems are powerful in resolving morphological and developmental characteristics of neuronal subtype with high throughput, however, existing tools are still limited by the pool size of available labels, balance of labeling assignment and flexibility in tuning labeling density and time point. These imperfections can be fixed with better genetic designs and inclusion of new recombination tools. In Chapter 2 of this thesis study, I present the design and evaluation of a new generation of multicolor genetic system, which aims to deliver up to tens of thousands of unique labels in the *Drosophila* nervous system with a concise and effective strategy. I also demonstrate its flexible application and tunable labeling density with a carefully designed initiation method. This novel genetic tool will allow more reliable and systematic analyses on lineage and morphological identities of neuronal subtypes.

1.4 Single-cell RNA sequencing permits molecular investigation of neuronal subtypes

In addition to morphological and developmental features, neuronal subtypes can be defined by their molecular identities. This is a critical categorical dimension as it provides important insights of what certain groups of neurons might do to fulfill specific neuronal functions and how they might accomplish it. These pieces of information have become more readily accessible with recent advances in single cell RNA sequencing (scRNAseq).

In succession to bulk RNA sequencing methods (Cloonan et al., 2008; Wang, Gerstein, & Snyder, 2009) which reveal transcriptomic information in a mixed population of cells, scRNAseq aims to precisely attribute these transcriptomes to each individual cells in the population (Klein et al., 2015; Macosko et al., 2015; Pollen et al., 2014; Tang et al., 2009). This makes scRNAseq perfectly suitable in resolving intrinsic heterogeneity at the RNA level among a group of neurons of interest. This permits effective validation of existing knowledge of known neuron subtypes, and allows findings of novel subtypes in a known population or across the whole nervous system. In the nervous system of human (Darmanis, Sloan, & Zhang, 2015), mouse (Lacar et al., 2016; Zeisel et al., 2015), zebrafish (Pandey, Shekhar, Regev, & Schier, 2018; Raj et al., 2018), and fruit fly (Croset, Treiber, & Waddell, 2018; Davie et al., 2018), scRNAseq has propelled findings in defining, correlating neuronal subtypes, and observing changes in them when development progresses or physiological/pathological alteration occurs in the systems. New knowledge acquired with scRNAseq has brought in exciting opportunities in understanding biological and pathological dynamics with unprecedented detail in high-throughput. However, because positional identity is lost as the cells are dissociated from the system to go through scRNAseq, the molecular identity assignment of each subtype needs thorough post-validation on a different animal, or by selected capture of well predefined cell populations before passing them

into the scRNAseq pipeline. In Chapter 4 of this thesis study, scRNAseq is applied on a selected group of *Drosophila* CNS neurons. The depth and throughput of transcriptome profiling on this neuronal group has revealed intriguing subtype findings, and demonstrated a powerful and reliable paradigm to thoroughly reveal molecular identities for accurate neuronal subtype definitions.

1.5 Serotonergic neurons in the *Drosophila* nervous system

Serotonin (or 5-hydroxytryptamine, 5-HT) is an important monoamine neurotransmitter participating in various important neurological functions in nervous systems across multiple species (Gaspar, Cases, & Maroteaux, 2003; Lillesaar, Tannhäuser, Stigloher, Kremmer, & Bally-Cuif, 2007; Sodhi & Sanders-Bush, 2004). In the *Drosophila* nervous system, there is a relatively small subset of neurons that produce serotonin (Lundell & Hirsh, 1994; Vallés & White, 1988), and they have been observed to play critical roles in maintaining and regulating important nervous system functions, such as feeding, courtship, aggression, learning and memory (Becnel, Johnson, Luo, Nässel, & Nichols, 2011; Gasque, Conway, Huang, Rao, & Vosshall, 2013; Johnson, Becnel, & Nichols, 2011; Liu et al., 2011; Sitaraman et al., 2008; Yuan, Joiner, & Sehgal, 2006).

Serotonergic neurons go through neurogenesis in early embryonic stages, and commit to their fate by 16-18 hours after-egg-lay (AEL) to be able to produce 5-HT (Vallés & White, 1988). Most if not all of these neurons persist through developmental stages from embryo, larva, pupa and eventually into adult. There are around 80 serotonergic neurons in the larva CNS, and 100 in the adult CNS (Vallés & White, 1988)(Huser et al., 2012). These neurons are stereotypically clustered to form distinctive neuronal patterns (Fig 1.3). In larva brains, they are symmetrically distributed as clusters located throughout the ventral nerve cord (VNC) which

includes 4 subesophageal (SE0~SE3), 3 thoracic (T1~T3), and 8 abdominal (A1~A8) segments, usually two to three neurons in each hemi-segment, and located in 4 groups in the central brain - SP1, SP2, IP and LP, often three to four neurons per cluster. In adult brains, the serotonergic neuronal patterns reorganized drastically after going through metamorphosis with addition and loss of neurons in the population. Dispersed more broadly in the central brain, adult serotonergic neurons formed 6 anterior clusters (ADMP, ALP, AMP, LP, SEL and SEM), and 4 posterior clusters (PMPd, PMPm, PMPv, PLP). Fewer cells populate the adult VNC which shrinks greatly from its significant larval form to three thoracic segments (pro- (PR), meso- (MS) and meta- (MT) thoracic) and one condensed abdominal segment (AB), and around one to two cells per thoracic hemisegment and ten cells per abdominal segment can be observed. (Niens et al., 2017)

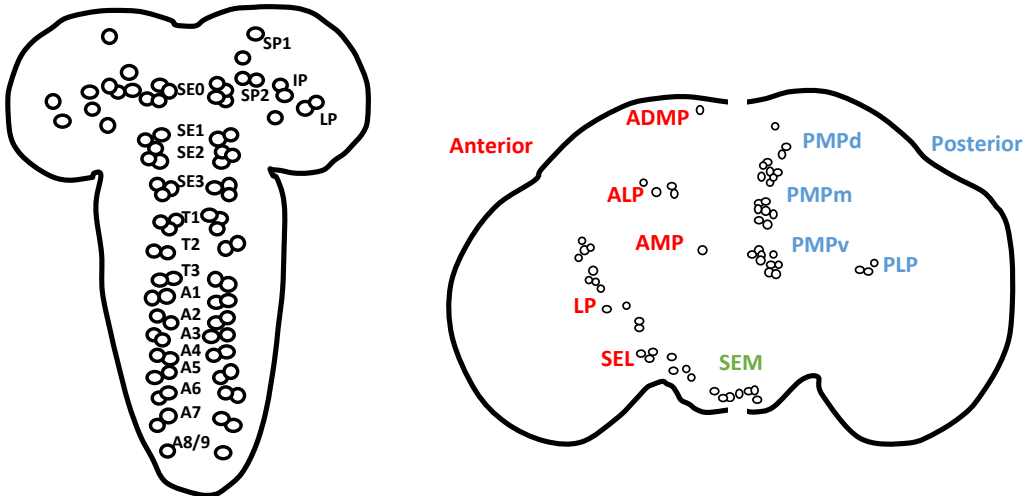


Fig 1.3 Distribution and nomenclature of serotonergic neurons in *Drosophila* CNS
Larval (left panel) and adult (right panel) serotonergic neuron cell body positions are illustrated. See text for more details about neuron nomenclatures.

Several critical questions about the *Drosophila* serotonergic system remains unresolved. Previous studies have set a strong base of morphological knowledge of larva and adult serotonergic neurons (Huser et al., 2012), however because of the density and complexity, detailed morphological properties of these neurons are only partially revealed so far (Chen & Condron, 2008). Further, although serotonergic neurons appear to be closely clustered in each

neuromeres from anterior to posterior, and a recent study suggests that neurons using the same fast-acting neurotransmitter are most likely to be developed from the same lineage (Lacin et al., 2019), it is still unknown whether similar lineage mechanism also applies to the serotonergic neurons. Finally, because of the multifaceted neurofunctional performance of the serotonergic neurons, it is unclear if all the serotonergic neurons are molecularly identical but simply participating in different circuitries, or the serotonergic neurons are substantially diverse however still sharing the same neuromodulation machinery by using serotonin. These unanswered questions are corresponding to the information in three essential subtyping categories, i.e. morphology, lineage and molecular identity. Careful investigations with advanced neurotechnologies are needed to answer these questions and obtain a more complete and systematic view of serotonergic neurons in *Drosophila*. In Chapter 3 of this thesis study, the newly developed multicolor genetic system is applied on the serotonergic neurons. Systematic morphological and lineage analyses have provided new insights of the subtype heterogeneity among the *Drosophila* serotonergic neurons.

Chapter 2. Bitbow: a Digital Format Brainbow Enables Spectral-spatial Barcoding for Lineage and Anatomical Tracing

2.1 Introduction

Central to the fundamental tasks in neuroscience, understanding how neurons look and where they come from, i.e. their morphological and lineage properties, has profound importance in revealing neuron's functional relevance in the nervous system. Morphology reflects the circuit structure by defining physical boundaries of receptive and projective fields, and lineage explains determination and specification of neurogenesis outcomes. To effectively elucidate the heterogeneity of neurons characterized by these two dimensions, proper tools are necessary to enable unambiguous targeting and analysis on any given type of neurons.

Tremendous efforts have been made in the past century to fulfil the fundamental request in tool development, evolving from methodologies that can cope with one or a few neurons at a time, such as sparse application of silver stain (Golgi's method) (Ramon y Cajal, 1911; y Cajal, 1910) or other chemical dyes (Honig & Hume, 1986, 1989) and mosaic genetic labeling (Dymecki & Kim, 2007; Lee & Luo, 1999; Luan, Peabody, Vinson, & White, 2006; Muzumdar, Luo, & Zong, 2007; Pfeiffer et al., 2010), to multispectral labeling technologies (Brainbow) that can differentiate large populations of neurons in the same tissue (Livet et al., 2007). Brainbow and Brainbow-like tools developed in recent years are capable to label neurons in distinct colors by expressing random ratios of different fluorophores upon genome recombination (Richier & Salecker, 2015). The recombinase-based "molecular switches", such as Cre/Lox (Livet et al., 2007), Flp/FRT (Cai, Cohen, Luo, Lichtman, & Sanes, 2013), and PhiC31/att (Kanca, Caussinus,

Denes, Percival-Smith, & Affolter, 2014) are used to control the specificity of labeled cell types and timing of color generation. Reagents, including mice, fruit flies, zebrafish, and viruses are now broadly available for studying the neurogenesis process and tracing the neuronal morphology. While enlightening the possibility of discerning individual cells by color in a complex system, current genetic designs are often limited to generating up to tens of reliably distinguishable colors. This makes it particularly challenging to interpret morphology or lineage tracing results because there is a high probability of labeling multiple cells or lineages with the same color.

Here in this chapter I present the design and evaluation of Bitbow, a set of transgenic tools that can generate up to tens of thousands of unique barcodes through a concise and powerful design. This improvement will allow unambiguous labeling of a much larger number of neurons in the nervous system. Additional genetic strategies in Bitbow also permit broad and tunable coverage of targeted cells, providing flexibility in studying complex neuron populations.

2.2 Results

Bitbow: the digital genetic design

We reasoned that permitting independent recombination of each FP would allow fully occupying the whole color space, therefore, generating the most dynamic colors from a single transgenic cassette. To do so, we utilized a pair of inverted FRT sites flanking a reversely positioned FP (Fig 2.1). This default “Off” state guaranteed a non-fluorescent expression without Flp activity. Flp recombination could flip the FP between the inverted FRT sites, resulting in either an “On” or “Off” state of the FP, similar to the digital “1” or “0” information that is stored in one bit of computer memory, respectively. Such combination scheme exponentially increases color coding capacity when increasing the number of bits (FPs) in the same transgenic animal,

e.g. 3 FPs generate 2^3=8 Bitbow barcodes, and 5 FPs generate 2^5=32 Bitbow barcodes. We decided to utilize known incompatible FRT sites FRT^{F3}, FRT^{5T2}, FRT^{545} (Cai et al., 2013), and three spectrally separable FPs, mTFP1, mCitrine and tdTomato to ensure independent events for each Bitbow bit. Further, each Bitbow bit is downstream of a UAS sequence, so the expression could be controlled by established enhancer-Gal4 drivers (Fig 2.1).

We first generated the 3FP-Bitbow transgene with validation in the cell culture, and a transgenic fly confirmed by its performance with a driver fly line containing flippase regulated by a heat-shock promoter (hsFlp) and a pan-neuronal Gal4 driver (elav-Gal4) (Fig 2.2). Bitbow labeling was successfully generated in the whole nervous system of the progeny flies from the cross between 3FP-Bitbow flies and driver flies. 7 Bitbow barcodes (8 total barcodes minus the case when all three FPs are "off") were observed, and the three bits appeared to be balanced in all observed Bitbow barcodes (Fig 2.2). Success of the 3FP-Bitbow proved the practicality of our digital genetic design in Bitbow.

Bitbow: utilize proper FPs and new FRTs to enable five bits

We moved to further improve Bitbow's labeling capacity to 5 bits, aiming for generating 2^5-1=31 unique Bitbow barcodes. In order to achieve this goal, we first selected five FPs for the design. We chose five FPs to be used in Bitbow: mAmetrine, mTFP1, mNeonGreen, mKusabira-Orange2, tdKatushaka2 (Ai, Hazelwood, Davidson, & Campbell, 2008; Ai, Henderson, Remington, & Campbell, 2006; Sakaue-Sawano et al., 2008; Shaner et al., 2013; Shcherbo et al., 2009). These FPs were chosen with the consideration of their brightness, photo-stability, antigenicity, and spectral separation (Fig 2.3). Next, in order to ensure independent recombination for each bit, we screened more FRTs that are incompatible with the three incompatible FRTs available. We found that FRT^{F13}, FRT^{F14} and FRT^{F15} showed satisfying

incompatibility (Fig 2.4), and we determined to use FRT^{F13} and FRT^{F14} together with FRT^{F3}, FRT^{5T2}, FRT^{545} in our final design, since FRTF 15 has lower recombination efficiency than the others (data not shown). Finally, in consideration of using Bitbow in the nervous system where it is crucial to reveal neuronal structure details for morphological studies, we decided to target Bitbow FPs onto cell membrane with Myristoylation signal from dSrc64B (1-10aa, (Struhl & Adachi, 1998)). We named this transgenic design mBitbow1.0, in which the "m" indicates the membrane targeting signal (Fig 2.5).

By crossing mBibow1.0 to the hsFlp;;elav-Gal4 driver, we successfully confirmed its performance in both larval and adult CNS (Fig 2.6). All 31 Bitbow barcodes was observed, and the five bits had overall balanced appearance except slightly higher observation of the mNeonGreen bit with FRT^{545} pair. This could be due to intrinsic reactivity differences among the FRT sites. Bitbow was able to resolve densely wired neurons, such as the ones in the optic lobe, with differentiating Bitbow barcodes and reveal clear details such as patches formed by the neurite terminus (Fig 2.6). Additionally, by controlling the time point of flippase activation through heat shock, Bitbow labeling could be initiated in neurons or in neuroblasts. In the latter case, because the Bitbow barcode could be inherited by all progeny generated by the same neuroblast, the entire lineage would be uniquely marked out by the same Bitbow barcode. In larval CNS where the progeny stayed close to the neuroblast, it was clear that each lineage was labeled by the same Bitbow barcode (Fig 2.6), highlighting the fidelity and practicality of Bitbow when applying to lineage tracing.

Bitbow: subcellular localizations to further expand barcoding capacity

Although the 31 unique Bitbow barcodes generated by mBitbow1.0 would be sufficient for many applications, especially in revealing neuronal morphological details, it was not

adequate for studies such as whole *Drosophila* central brain lineage tracing, where the number of targeting entities (such as >200 lineages in the central brain) outnumbered the labeling capacity. In those cases, repeated Bitbow barcodes (i.e. "barcoding collisions") would occur, and if adjacent lineages were labeled by the repeated Bitbow barcodes, they would be erroneously regarded as one lineage. Therefore, more Bitbow barcodes were needed to resolve this challenge.

How many more Bitbow bits would be needed? We set to estimate it with statistical simulations by randomly accessing barcoding pools of different sizes, such as of 2^5, 2^{10}, 2^{15} unique barcodes, and quantify how frequently repeated barcodes were seen, i.e. the barcode collision rate (Fig 2.7). It was evident that the three levels of labeling capacity produced different speed of collision rate increase as well as different saturation points, which was the point where certain number of entities were labeled while barcode collision happened in 100% of the simulations; the 2^5 pool has the fastest collision rate growth and lowest saturation point, and the 2^{15} pool had the slowest growth and highest saturation point. When considering the case of labeling ~200 neural lineages in the *Drosophila* central brain, the larger barcode pool clearly had the best performance, where the 2^{15} pool exerted <0.2% collision rate, and the 2^5 or 2^{10} pools had collision rates close to 100% or >10% (Fig 2.7).

How could we improve Bitbow so it would be possible to generate up to 2^{15} Bitbow barcodes? A straightforward strategy would be to expand the number of bits in Bitbow, i.e. adding in more FPs in each transgenic line, and this increase would be exponential. However, due to difficulties in imaging multiple FPs in the limited visible spectrum, it would be challenging to fit more than 5FPs together. An alternative strategy would be adding in more Bitbow bits with the same 5FP set but with different subcellular localizations (Fig 2.8). With sufficient separation between the localizations, each bit could still work independently and

faithfully, even in the extreme cases where the same FPs are expressed in different subcellular localizations.

We chose to target three subcellular locations: cell membrane, nucleus and the Golgi apparatus, for their best spatial separation and structural integrity inside cells. In addition to the myristoylation signal peptide from *Drosophila* Src64B, full length human Histone H2B (Shaner et al., 2013), and N-terminus of mouse Mannosidase II alpha 1 (Ye et al., 2007) were used to fuse with FPs and target them to cell membrane, nucleus or Golgi, respectively (Fig 2.8). Making each FP at each subcellular localization as a single "bit", we would be able to get a labeling system with 15-bit capacity (Fig 2.8). Three localization tags were tested to be effective in cultured *Drosophila* cells (data not shown), and individual Bitbow transgenic flies were established accordingly. We named the two new transgenics as nBitbow1.0 and gBitbow1.0, representing their nucleus- and Golgi-targeting designs. We also generated fly lines that combine multiple Bitbow transgenes together, either through meiotic recombination on the same chromosome, or combination of 2nd and 3rd chromosome transgenes. We named these lines reflecting their transgene content accordingly, such as mnBitbow1.0, mngBitbow1.0.

Multi-subcellular-localization Bitbow worked effectively. Using the driver line hsFlp;;elav-Gal4 and conducting heat shock at 1st instar larva stage (Fig 2.9), we were able to label neuroblasts that were ready to enter the second wave of neurogenesis. Neuroblasts with recombined Bitbow labeling passed the barcode onto their progeny, and all cells in the same lineage were labeled by the same Bitbow barcode. Bitbow labeled lineages were readily evident in the 3rd instar brains, with one of the three subcellular localizations (Fig 2.9), or the combination of them (Fig 2.10). In mnBitbow1.0 and mngBitbow1.0 labeled brains, we had observed diverse Bitbow barcodes consisting of one to three localizations, and unambiguous

barcoding calling was practical due to the sufficient spatial separation of three subcellular structures as well as spectral separation of the five FPs. We observed 80 barcodes from mnBitbow1.0 labeled brains (4 brains, 286 clusters quantified), and 153 barcodes from mngBitbow1.0 labeled brains (2 brains, 247 clusters quantified) (Fig 2.11). Among these barcodes, 38 of the 80 mnBitbow1.0 barcodes and 111 of the 153 mngBitbow1.0 barcodes were observed to be unique across all quantified brain samples (Fig 2.11). In each brain, we found on average 30% of the clusters were labeled by a unique mnBitbow1.0 barcode, and 57% by a unique mngBitbow1.0 barcode (Fig 2.11). Using observed statistics of the barcodes in mngBitbow labeled third instar brains, we derived the possibility of occurrence of all 32767 mngBitbow barcodes, and composed lists of barcodes including top 32000 or 32700 ones with the least possibility of occurrence. With this underlying discrete probability distribution function, we applied Monte Carlo simulation to aim to label up to 10^4 lineages with top 32000 or top 32700 or all 32767 mngBitbow barcodes, as well as with two existing genetic tools, MARCM (Lee & Luo, 1999) and Brainbow 1.1 (Livet et al., 2007) (Fig 2.12). We discovered the need of animal numbers along the increasing targeting lineages grew much slower in Bitbow groups, comparing to MARCM or Brainbow 1.1. In typical cases where ~100 lineages needed to be labeled, less than 10 animals were needed for Bitbow (32767 or 32700 barcodes) to ensure full coverage of all lineages, whereas hundreds of animals were needed when using MARCM or Brainbow 1.1 (Fig 2.12). In summary, Bitbow transgenics with multi-subcellular-localization designs were capable to generate large number of unique barcodes to allow effective unambiguous lineage analyses.

Bitbow: novel genetic design to allow self-initiation and broader coverage

Bitbow 1.0 transgenes achieved desirable barcoding capacity for neuronal tracing and lineage tracing with flexible control of initiation time by hsFlp. However, the amount of flippase activity provided by this exogenous method was limited. Multiple heat shocks were often needed to ensure sufficient Bitbow coverage among all cells (Fig 2.13). This was compatible with experiments where the timing of heat shocks was not critical, but unfavorable when precise labeling initiation is needed. Alternatively, if a higher level of flippase activity was provided by longer heat shocks (such as 60-min rather than 30-min heat shocks), the animals were challenged by stronger stress, and the survival rate was lower (data not shown). Further, the requirement of hsFlp based driver lines for Bitbow1.0 brought in a relatively complicated experimental scheme, where a hsFlp + enhancer-Gal4 line needed to be generated, animal collection as well as heat shock needed to be properly conducted and the flies needed to be reared at lower temperature to avoid undesired labeling triggered by ambient temperature. We reasoned that we could overcome these difficulties by extending the Bitbow1.0 design with sufficient amount of endogenously supplied flippase.

Bitbow 2.0 transgenes were created following this principle of design. Specifically, a self-excising flippase module was put together with the mBitbow1.0, consisting of the flippase cDNA flanked by a pair of FRT sites in the same direction, driven by the promoter of *Drosophila* neuronal Synaptobrevin (nSyb). This design permitted a strong burst of neuronal-specific expression of flippase which would react with Bitbow modules to generate barcodes randomly and eventually excise out the flippase cDNA to prevent excessive activities. To ensure sufficient amount of flippase could be produced before the coding sequence was excised out, we chose FRT^{F13} or FRT^{F15} as the flanking FRTs which have weaker reactivity among the incompatible

FRTs screened previously in our study, and we named the corresponding transgenes as mBitbow2.0 and mBitbow2.1, respectively (Fig 2.14).

Labeling coverage was greatly improved by the Bitbow2.0 design. Comparing to mBitbow1.0 when labeling the larval serotonergic neurons, it is evident that mBitbow2.0 achieved much better labeling coverage (Fig 2.15), with a largely simplified experimental setup where only a one-step Gal4 cross was needed. We further improved the labeling coverage by supplying more endogenous flippase. This was achieved either through the use of two copies of mBitbow2.0, or the use of mBitbow2.1 which has the weaker FRT pair (FRT^{F15}), or two copies of mBitbow2.1. Better coverage was observed when more flippase was included in the system where as high as a 93.83% coverage can be achieved, evaluated by the proportion of serotonergic was labeled by using different Bitbow reporters or an UAS-TagBFP reporter that labeled all the cells (Fig 2.15). Aside from the high labeling coverage, Bitbow 2.0 also inherited the high coding capacity from the Bitbow1.0 series, evident by the diverse membrane color combinations (Fig 2.16). Also, Bitbow2.0 transgenes were flexibly applicable to label different neuron subsets with ease (Fig 2.16).

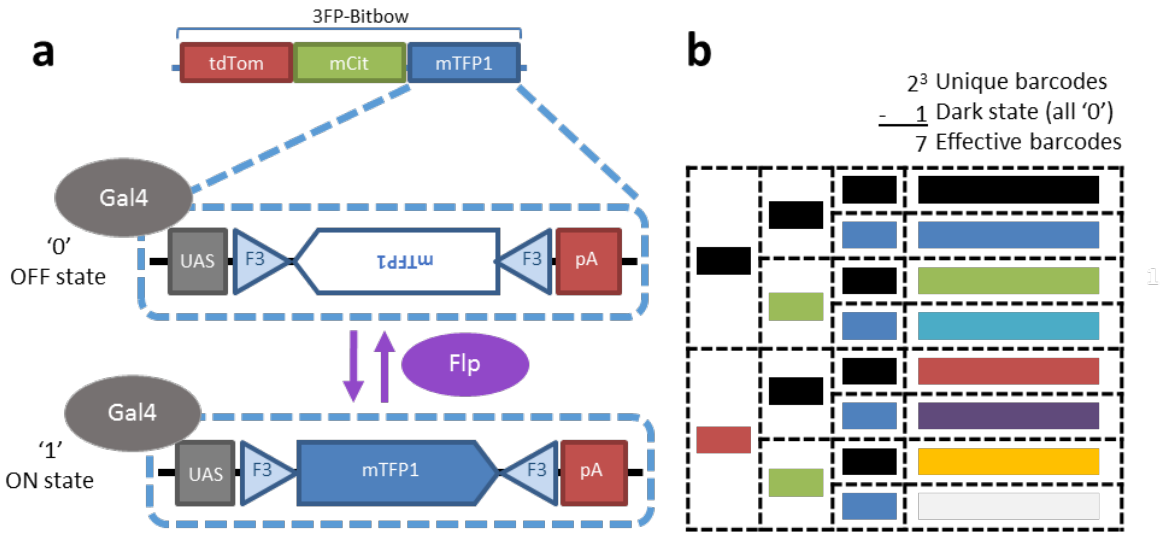


Fig 2.1 Genetic design of the 3FP-Bitbow

(a) Design of the 3FP-Bitbow. Each bit is organized as UAS-FRT-inv(xFP)-FRT-pA from 5' to 3'. Upon flippase recombination, the inverted FP is capable being switched between ON or OFF state. All bits are under regulation of Gal4-UAS sysem.

(b) All seven possible combinations of Bitbow barcodes generated by the 3FP-Bitbow.

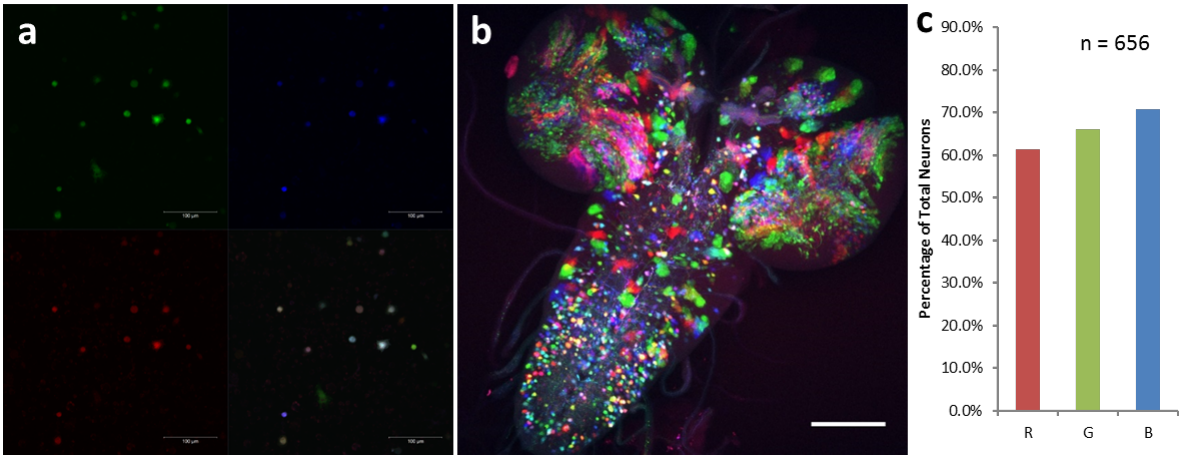


Fig 2.2 3FP-Bitbow performance confirmed in *Drosophila* S2 cell culture and third instar larva brain.

(a) clockwise: mCitrine, mTFP1, tdTomato.

(b) third instar larva brain labeled by hsFlp;;elav-Gal4 crossed to 3FP-Bitbow, with a heat shock during the first instar larva stage. Scasle bar, 100um.

(c) Quantification of the percentage of cells containing tdTomato (R), mCitrine (G) or mTFP1 (B) expression.

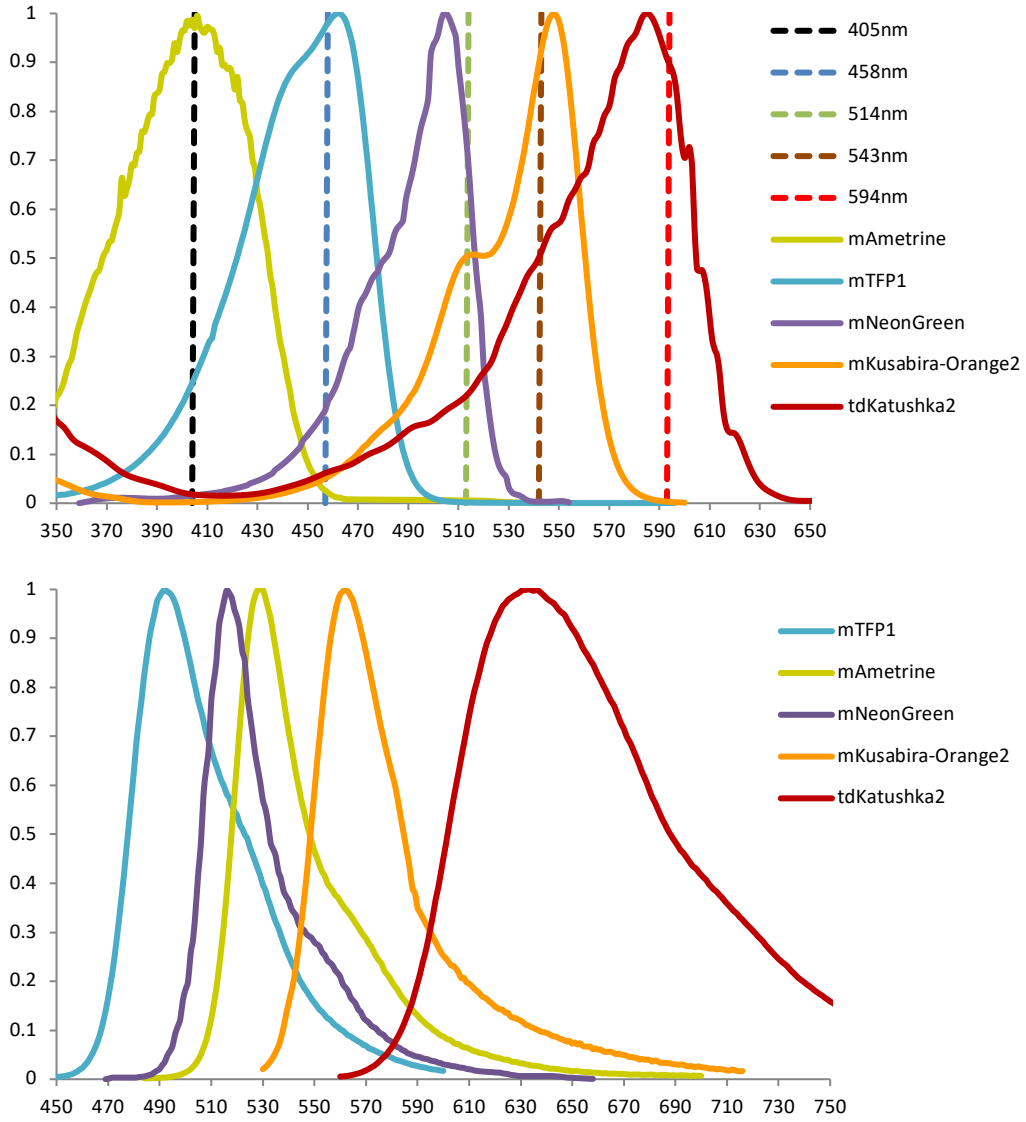


Fig 2.3 Spectral profiles of five fluorescent proteins in Bitbow.
(top) Excitation spectra of the five FPs, with wavelength indicator of five common laser lines.
(bottom) Emission spectra of the five FPs.

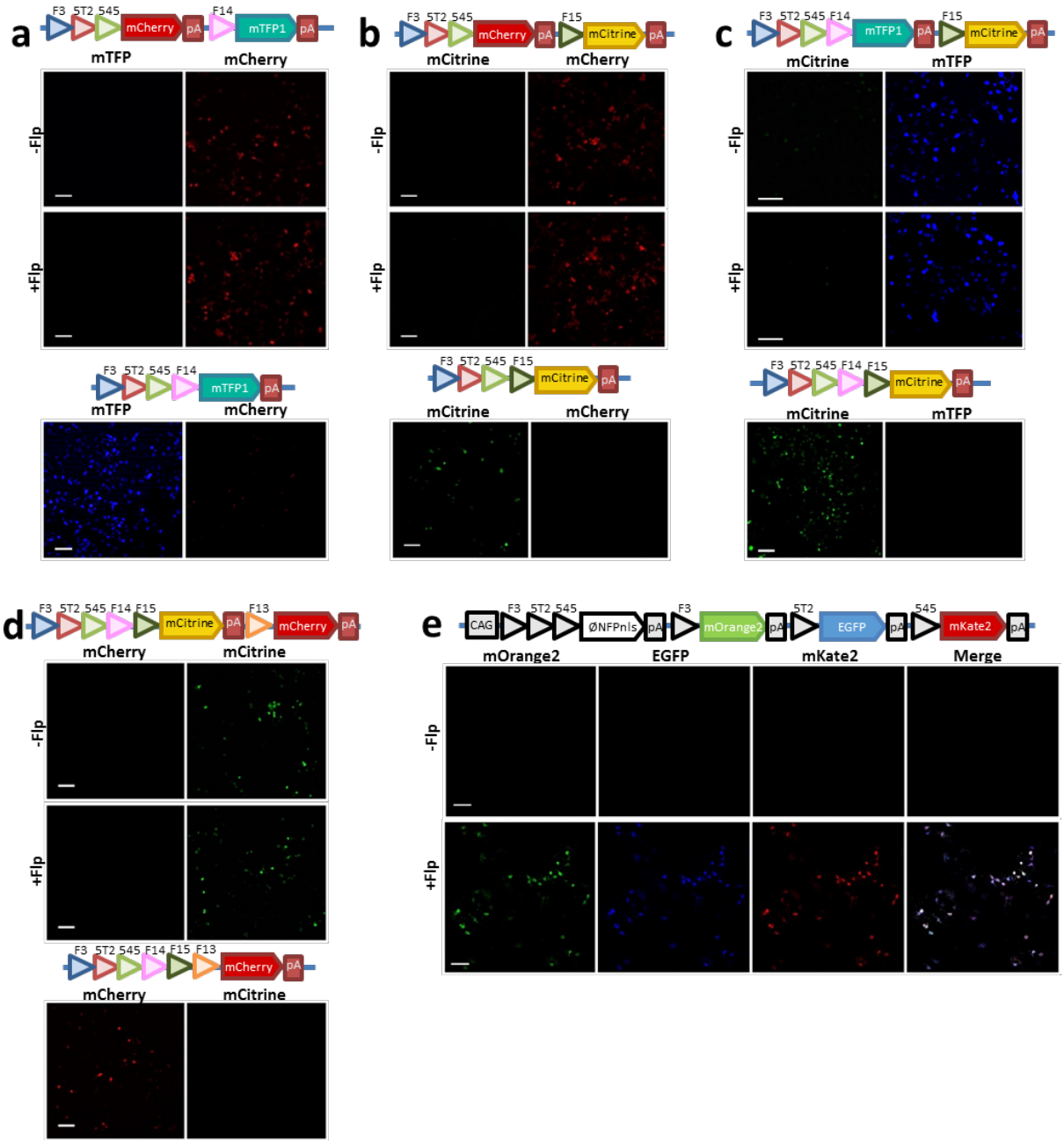


Fig 2.4 Additional FRT sites are selected to ensure no cross-bit recombination would occur between the known and new FRT sites

(a) FRT^{F14} is screened with the three known incompatible FRT sites. A dual FP reporter construct is co-transfected into HEK293T cells with or without a pCMV-Flp construct; no mTFP expression was detected in either cases, indicating FRT^{F14} being incompatible with the other FRT sites. The transfection of a control plasmid in which the mCherry is removed has shown clear mTFP expressions, indicating that mTFP signal would come up, should the F14 be compatible with the other three FRT sites.

(b-d) FRTF15 and FRTF13 are screened in similar genetic setups, and have shown to be incompatible with all known incompatible FRT sites, and between them.

(e) A control transfection of Flpbow (Cai, et al., 2013) with or without pCMV-Flp construct in the same batch of experiment, indicating a functional Flp construct.

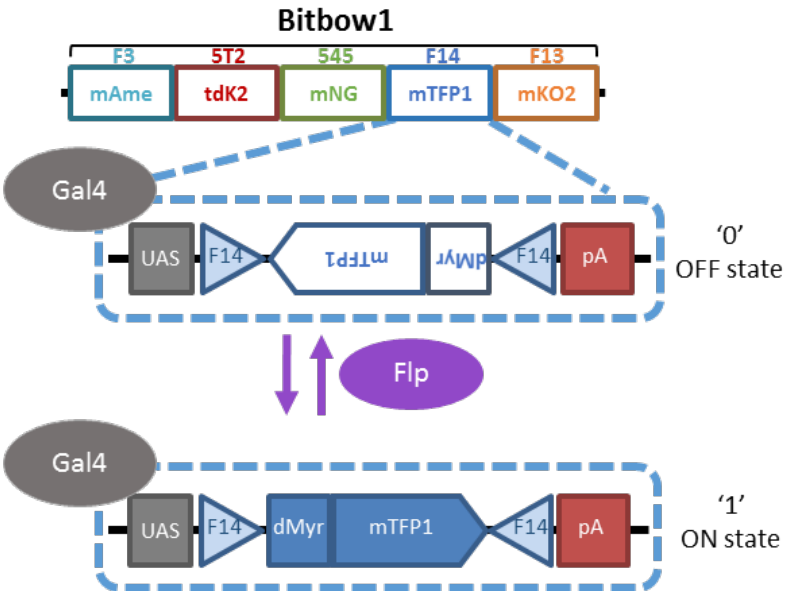


Fig 2.5 Genetic design of mBitbow 1.0

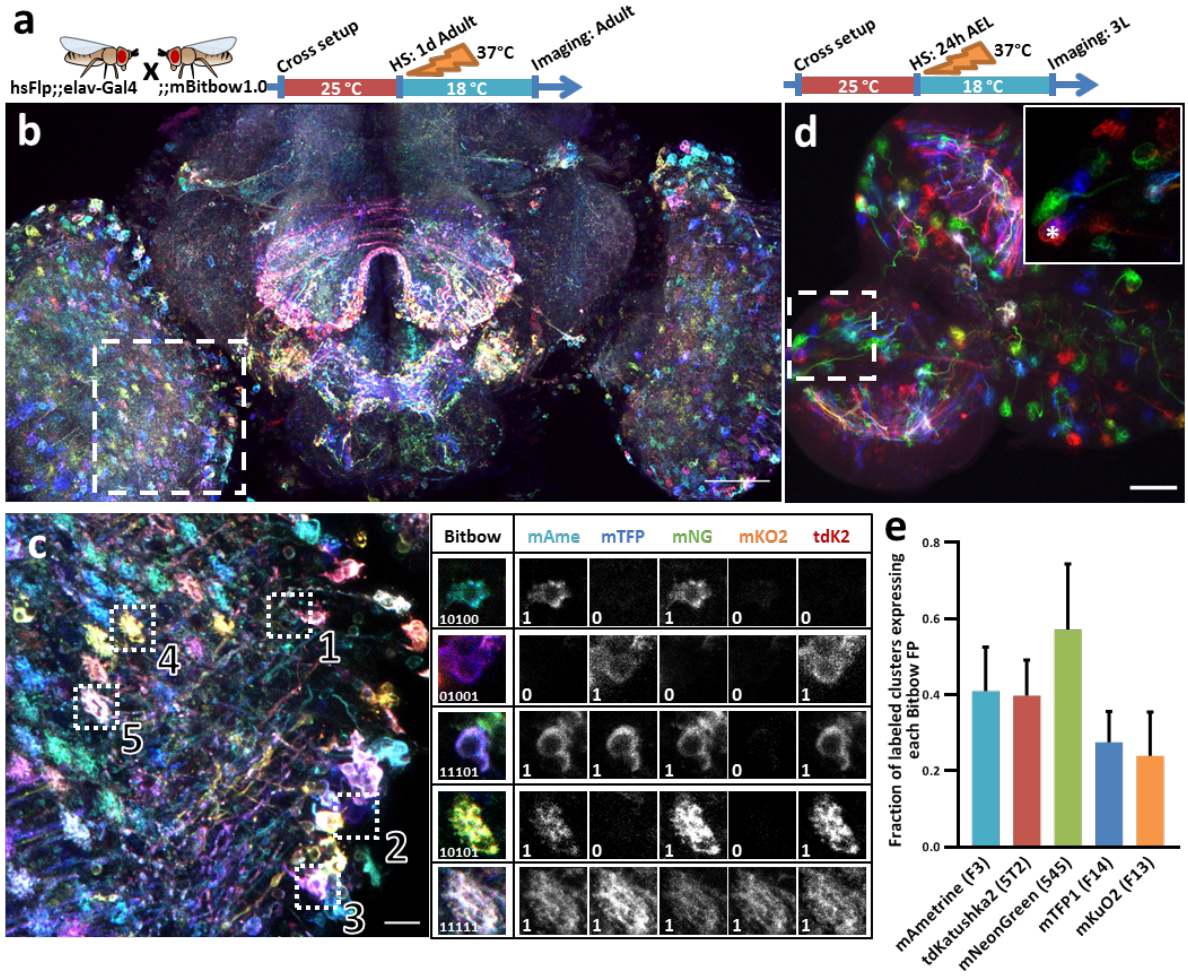


Fig 2.6 mBitbow 1.0 performance confirmed in adult and larval brains

(a) Experimental setups to generate heat shock induced Bitbow labeling
(b) Adult brain labeled by hsFlp;;elav-Gal4 and mBitbow1.0. Dashed white box, region of (c). Scale bar, 50um.
(c) Maximum projection of dashed squre region in (b). Bitbow barcodes were breakdown of five selected somas and terminals. Scale bar, 10um
(d) Larval brain labeled by hsFlp;;elav-Gal4 x mBitbow1.0. Insert, enlarged image of the dash box region. Asterisk indicates a neuroblast. Scale bar, 50um.
(e) Quantification of activation coverage by each Bitbow module. N=787 clusters, 6 brains. Error bars are SD.

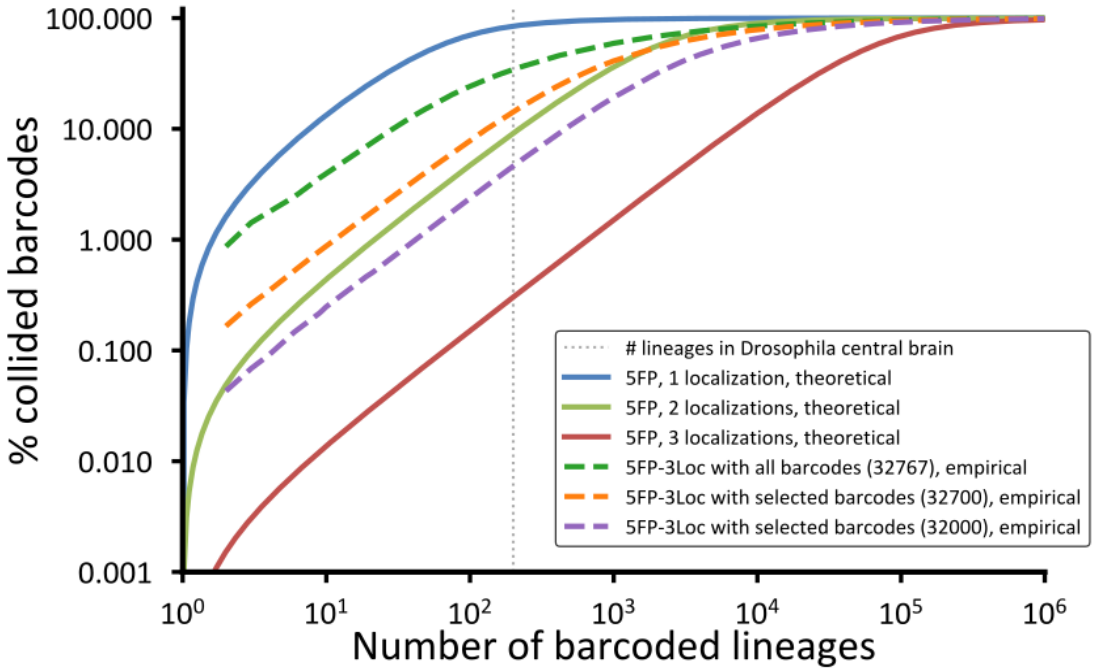
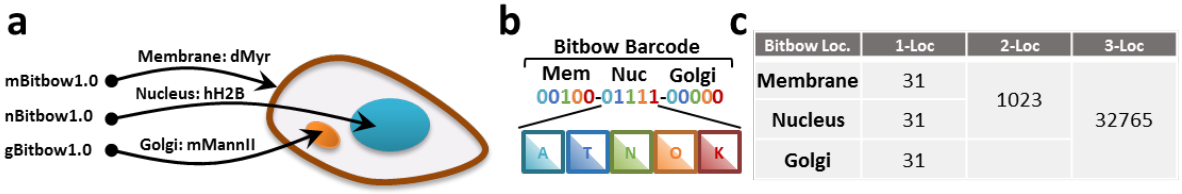


Fig 2.7 Randomized simulation to estimate barcode collision rates when labeling various amount of lineages with different amount of Bitbow barcodes

Solid plots: theoretical barcode collisions of 2^5 (1-localization), 2^{10} (2-localization), and 2^{15} (3-localization) barcodes were calculated with a closed-form solution. Dashed plots: Monte-Carlo simulations based on real barcode occurrence frequency observed in mngBitbow labeled 3rd instar brains, with all (32767), less frequent 32700 or less frequent 32000 barcodes included in the simulations. Dotted vertical line: 200 lineages, which is the amount of lineages estimated in the *Drosophila* central brain.



Bitbow Loc.	1-Loc	2-Loc	3-Loc
Membrane	31	1023	32765
Nucleus	31		
Golgi	31		

Fig 2.8 Bitbow targeting to multiple sub-cellular spatial localizations generates up to 15-bit labeling capacity

(a) Bitbow1.0 transgenes are designed to targeting to membrane, nucleus or Golgi-apparatus.

(b) Bitbow barcode is organized in sequential 15-bit format, with each bit representing a FP at a specific subcellular location.

(c) Amount of unique barcodes that could be generated by 1-, 2- or 3- localization Bitbow.

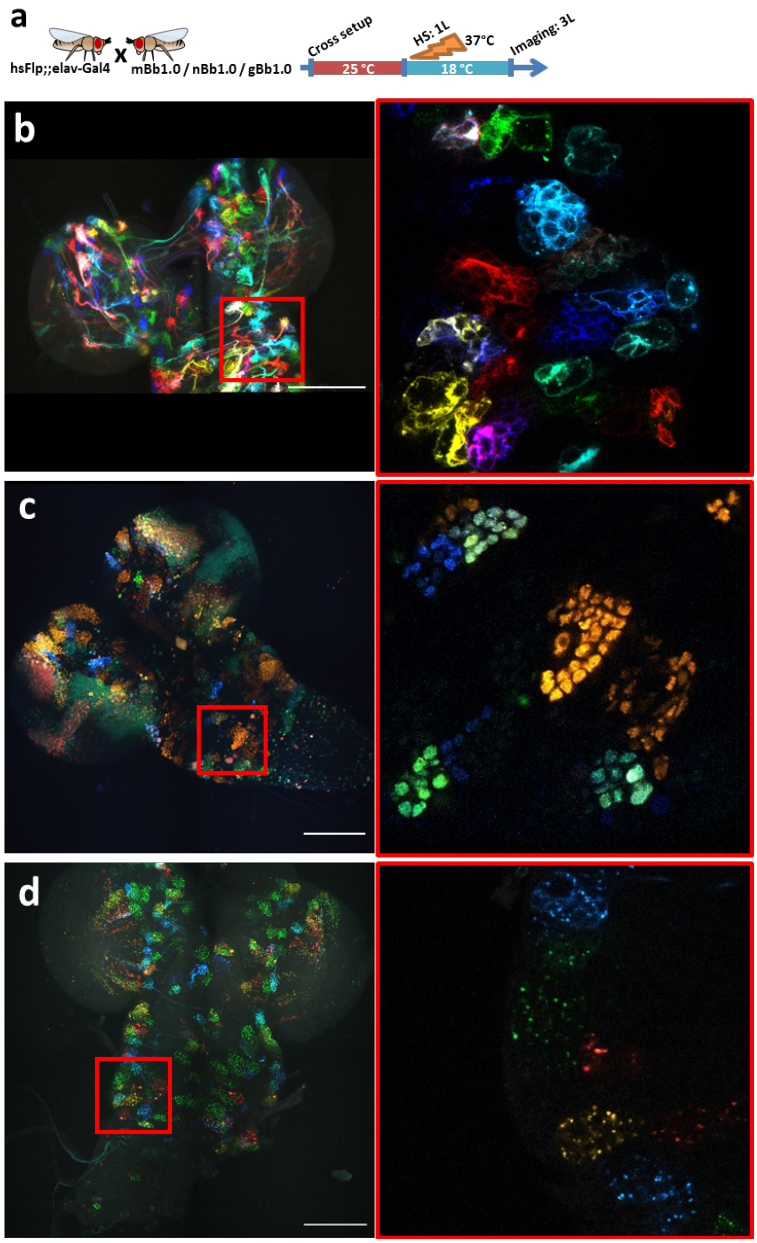


Fig 2.9 Performance of Bitbow targeting to the three sub-cellular spatial localizations were confirmed in third instar larva brains

(a) Experimental setup of the crosses and heat shock plan.
(b) mBitbow1.0 expression in the 3rd instar larval brain. Red box, an enlarged region displaying multiple labeled lineages.
(c) nBitbow1.0 expression in the 3rd instar larval brain.
(d) gBitbow1.0 expression in the 3rd instar larval brain.

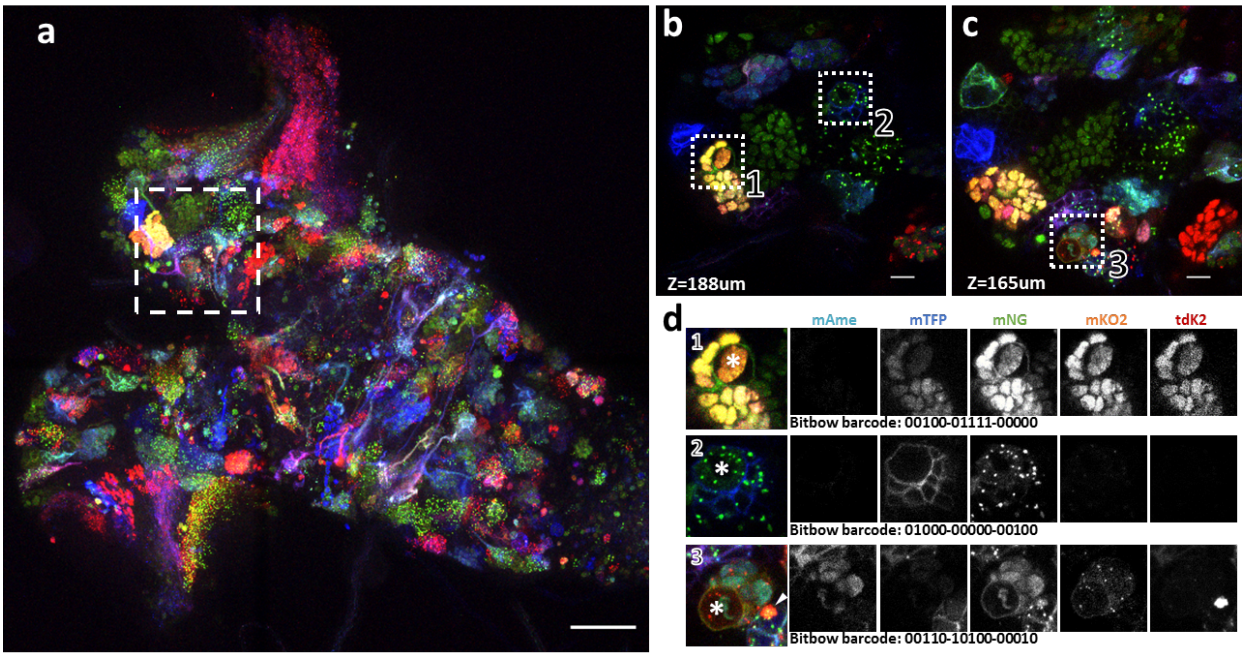


Fig 2.10 Three-localization containing mngBitbow1.0 was capable of labeling lineages in the third instar brain with diverse Bitbow barcodes

(a) A 3rd instar brain labeld by mngBitbow1.0 and hsFlp;;elav-Gal4. Scale bar, 50um.
(b) Two z-slices corresponding to the dashed-box area marked in (b). Scale bar, 10um.
(c) 3 clusters marked in (c) are analyzed by 3-localization Bitbow barcodes. Asterisks indicate neuroblasts of each cluster. The arrowhead indicates a cell labeled by a distinct Bitbow barcode located in between two adjacent clusters.

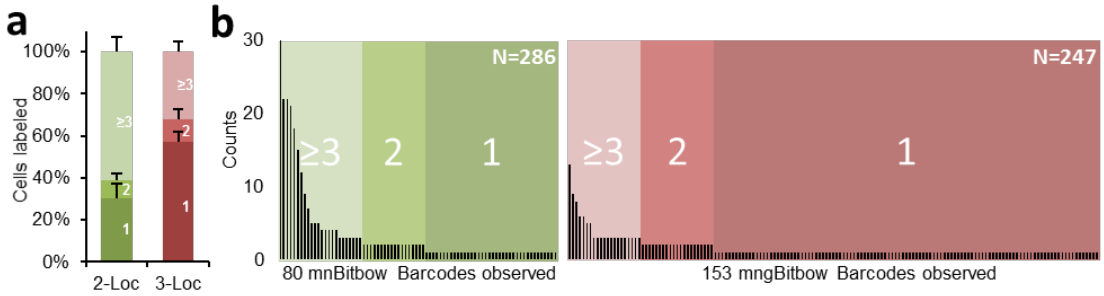


Fig 2.11 Quantification of mnBitbow (2-Loc) and mngBitbow (3-Loc) barcodes observed in third instar larva brains.

(a) Proportion of cells labeled by unique (1), duplicated (2), or multirepeat (≥3) barcodes (2-Localization or 3-Localization) in each brain sample. Mean and SD are shown.

(b) Histogram of mnBitbow (2-Localization) and mngBitbow1.0 (3-Localization) barcodes that are in labeled neuronal clusters 2 brains, respectively. Horizontal axis, individual Bitbow barcodes. Vertical axis, observation counts of each particular barcode in all brain samples. Shaded regions indicate barcodes that been observed 1, 2, or ≥3 times.

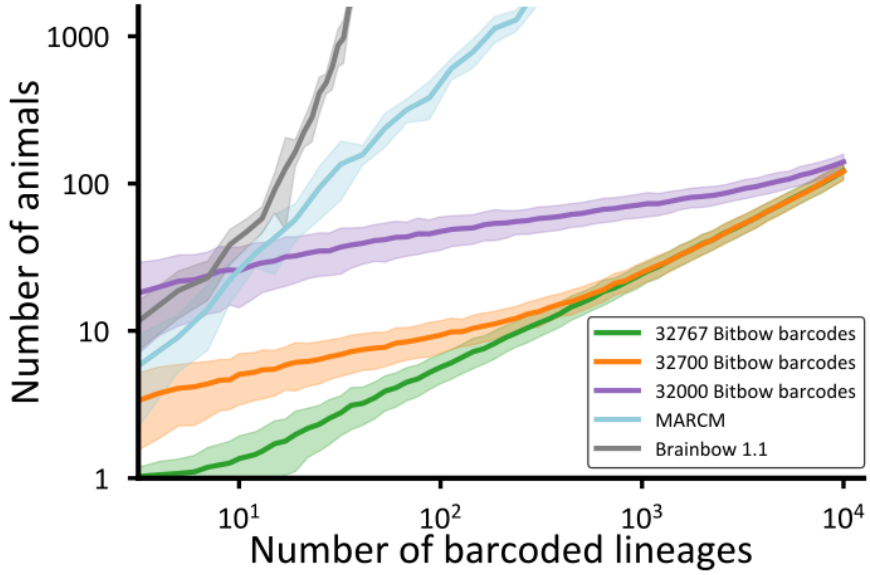


Fig 2.12 Monte-Carlo simulation of number of experimental animals needed to unambiguously resolve various number of barcoded lineages with different genetic approaches

Number of animals needed (y-axis) to achieve full coverage of given number of lineage barcoding (x-axis) with different barcode pool are estimated by Monte Carlo simulations. Solid lines, mean of simulations; shaded area, standard deviation of simulations.

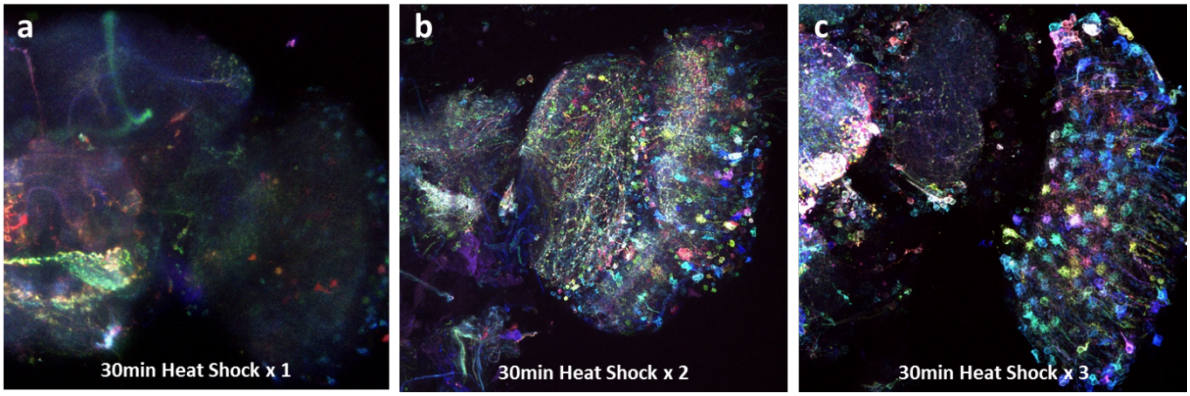


Fig 2.13 Multiple heat shocks were needed to increase labeling coverage with mBitbow1.0
Adult brains of the mBitbow1.0 and hsFlp;;elav-Gal4 cross, with 1, 2 or 3 heat shocks before they are dissected. Half of the brain and one of the optic lobes are displayed.

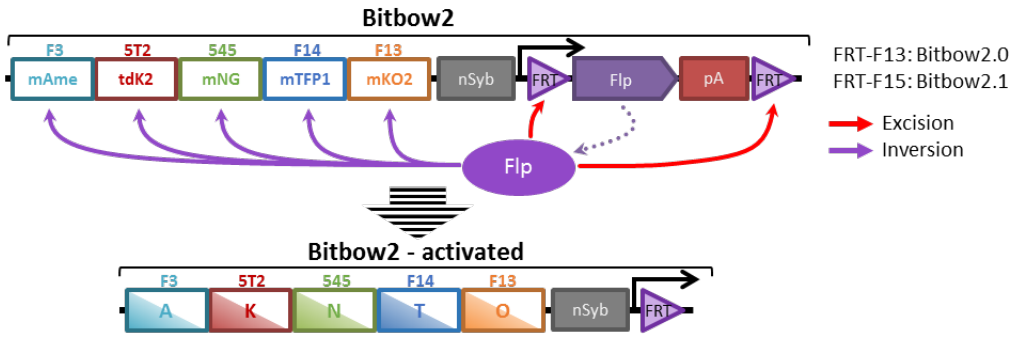


Fig 2.14 Genetic design of Bitbow2

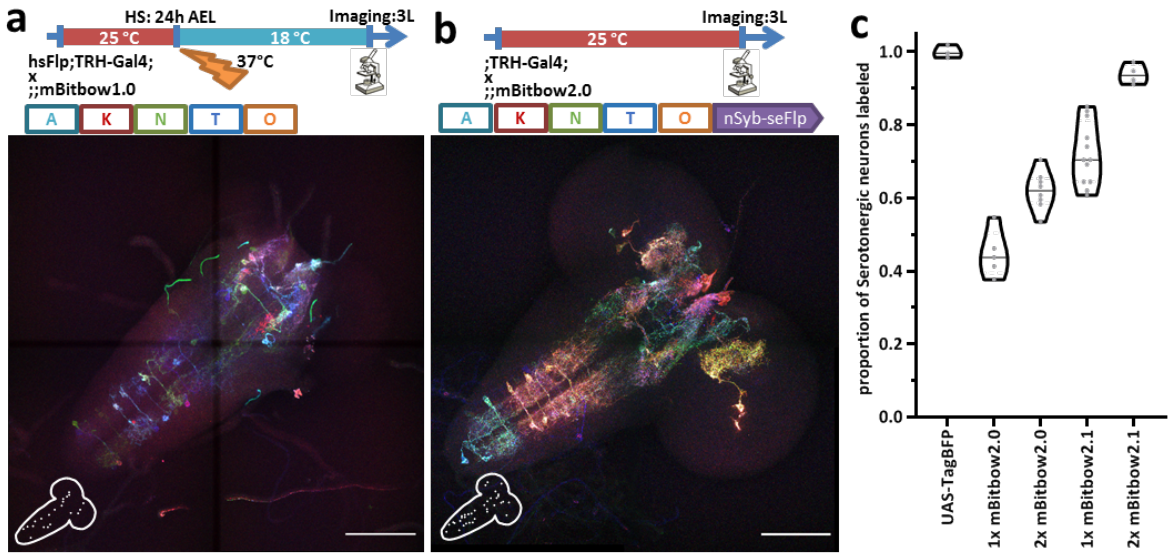


Fig 2.15 Bitbow2 designs achieved improved labeling coverage
(a-b) Larval serotonergic neurons are labeled by mBitbow1.0 and mBitbow2.0. Experimental setups are shown above the images, respectively. Cell bodies of labeled serotonergic neurons are marked in the abstract illustration in the bottom left corners. Scale bars, 100um.
(c) Quantification of proportion of neurons being labeled by Bitbow2 constructs. Horizontal lines inside the violin plots indicate means of the groups.

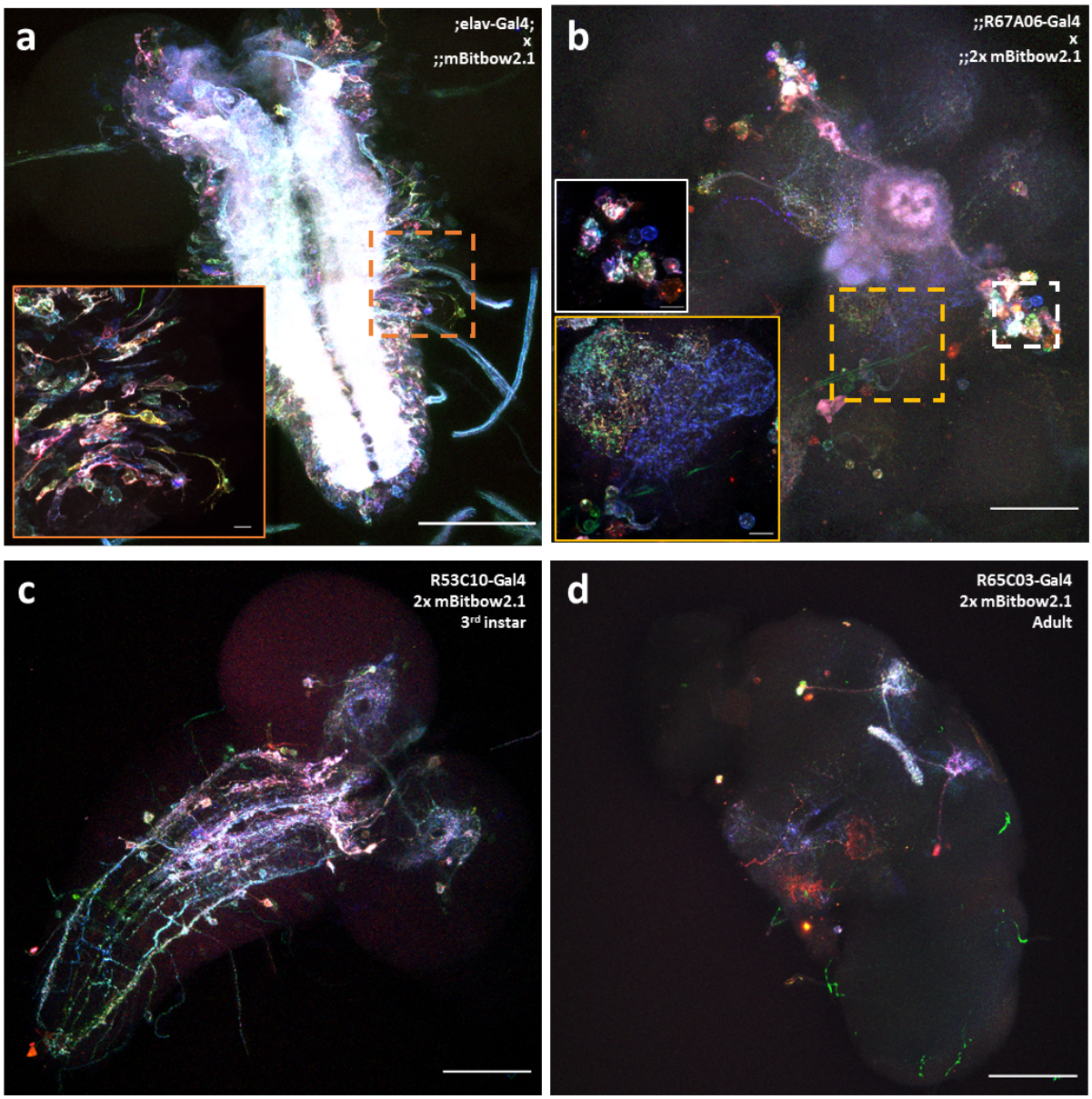


Fig 2.16 Bitbow2 was able to flexibility label groups of neurons with diverse Bitbow colors in high coverage

(a) Larva neurons are labeled by elav-Gal4 and mBitbow2.1. Scale bar, 100um. Insert scale bar, 10um.
(b) Adult neurons are labeled by R67A06-Gal4 and 2x mBitbow2.1. Scale bar, 50um. Insert scale bars, 10um.
(c) Larva neurons are labeled by R53C10-Gal4 and 2x mBitbow2.1. Scale bar, 100um
(d) Adult neurons are labeled by R65C03-Gal4 and 2x mBitbow2.1. Scale bar, 100um

Table 2.1 List of Bitbow transgenic flies created in this study

Bitbow Version	Subcellular Loc.	Insertion site (Chr.)	Contain self-activating Flp?	Labeling density
3FP-Bitbow	Whole cell	attP40 (2L)	No	Variable (with hsFlp)
mBitbow 1.0	Cell membrane	attP40(2L), attP2(3L)	No	Variable (with hsFlp)
nBitbow 1.0	Nucleus	attP40(2L), VK27(3R)	No	Variable (with hsFlp)
gBitbow 1.0	Golgi	attP40(2L), attP2(3L)	No	Variable (with hsFlp)
mnBitbow 1.0	Cell membrane & Nucleus	attP2(3L), VK27(3R)	No	Variable (with hsFlp)
mngBitbow 1.0	Cell membrane, Nucleus & Golgi	attP40(2L), attP2(3L), VK27(3R)	No	Variable (with hsFlp)
mBitbow 2.0	Cell membrane	attP2(3L), VK27(3R)	Yes	Medium
mBitbow 2.1	Cell membrane	attP2(3L), VK27(3R)	Yes	High

2.3 Discussion

Bitbow has demonstrated its capability in delivering unambiguous barcoding for anatomical and lineage studies. Bitbow1 series are a company of transgenes fully utilizing spectral and spatial dimensions to maximize possible barcoding outcomes from a single copy of transgene. With temporal controls that are governed by external manipulations, such as heat shock induced activations, desired density in morphology tracing and desired initiation time in lineage dating can be tuned without interfering with barcoding outcomes. Bitbow2 series has inherited advantages from Bitbow1, and extended it by a self-initiation/self-termination mechanism, which overcomes the need of external input for Bitbow initiation.

There are still several aspects of Bitbow design that can be further improved in the future development. First, Bitbow will be greatly benefited from a more balanced set of FRT sites. Currently, the five FRTs exert different levels of reactivity reflected by their relative occurrence in labeled neurons, as it is evident that FRT^{545} is more active than the others, and FRT^{F14} and FRT^{F13} being less active. This uneven reactivity has caused some barcodes to appear more frequent than the others, and eventually caused higher barcode collision rate in the system than the theoretical estimation, in which case all barcodes are equally probable (Fig 2.7). If more incompatible FRT sites with better reactivity could be found, a more balanced set of FRTs can be selected and the performance of Bitbow could be improved using the new set. Second, more bits and more identifiable levels in each bit are always welcomed. Although concurrent imaging of five FPs in the range of visible spectrum is already getting challenging, it is still worth consideration to expand the number of bits by introducing Far-red fluorophores, or dark fluorophores who can be distinguished antigenically for multiple rounds of detections. Because Bitbow barcoding capacity grows exponentially, each additional bit will double the available

barcodes, which will greatly suppress barcode collision rate and reduce the number of animals needed for the same experimental coverage. If more than 2 levels expression of each bit can be reliably called, this will transform the base of Bitbow barcode capacity from 2 to 3 or more. This approach will allow more information collected from every bit, and significantly improve Bitbow labeling without introducing more bits. Consistent multi-level signal processing could be achieved with the advancement in imaging and post-imaging processing methods. Finally, the Bitbow2 design can extend to specifically label other cell types in the nervous system, or cell types in other systems, by changing the neuronal-specific promoter nSyb to other appropriate promoter/enhancer components. However, it is worth emphasizing that one important criterion for such modifications to be successful is to ensure the desired promoter/enhancer element will not have any significant function in germline cells. Otherwise, the designed transgene will be prematurely recombined in the germline, and no longer functional in desired cell populations.

Even without the aforementioned possible improvements, Bitbow is capable in many other experimental tasks in its current format. Since a set of bright, stable and spectrally resolvable fluorescent proteins are carefully selected in the design, Bitbow can be effective in prolonged multicolor live imaging. We and collaborators have preliminarily demonstrated (Veiling, Li, Veiling, et al, 2019, in revision) the performance of nBitbow1.0 in live tracking embryonic neurons during early developments. Also, Bitbow can empower neuron subtype studies in combination with molecular profiling at transcriptional and translational levels. Since Bitbow can bring in unambiguous neuron morphology labeling in given set of neurons, staining or tracking of endogenous RNA/protein targets in combination with Bitbow will greatly accelerate the discovery of novel neuronal subtypes.

2.4 Materials and Methods

Fly husbandry. Flies were reared at 25c on standard CT food. Following fly lines were acquired from Bloomington Drosophila Stock Center (BDSC): w;;elav-Gal4 (#8760), w;TRH-Gal4; (#38388), w;;R53C10-Gal4 (#38873), w;;R67A06-Gal4 (#39397). Heat-shock promoter driven flippase line $hsFlp^{112}$;Sp/CyO;TM2,Ubx/TM6B,Tb was gifted by Dr. Bing Ye lab (University of Michigan). Driver lines was created by combining the hsFlp transgene with Gal4 transgenes, yielding hsFlp;;elav-Ga4 and hsFlp;TRH-Gal4;.

Dissection and mounting. Drosophila 3rd instar and adult brains were dissected in PBS at room temperature (RT) within 30min before proceeding to fixation. Dissected brains were fixed in 4% PFA (Sigma #158127, diluted in PBS) at RT with gentle nutation for 20min, followed by three quick PBST (PBS+1% Triton X-100) washes, then PBS washes for 15min x 3. Brains then either proceeded to direct mounting (for native fluorescence imaging) or immuno-stainings. Vectashield (Vector Laboratories, H-1000) were used as the mounting medium.

Molecular Cloning and Fly Transgenics. cDNAs encoding the following FPs were used: mAmetrine, mTFP, mNeonGreen, mKusabira-Orange2, and tdKatushka2. Drosophila myristoylation signal peptide of dSrc64B (1-10aa, dMyr), Human histone 2B protein (full length, hH2B) or Mouse Mannosidase II alpha 1 (1-112aa, mManII) was fused in frame to the N-terminus of individual FPs (dMyr-FP, hH2B-FP, mMannII-FP), to achieve targeted labeling at cell membrane, nucleus or Golgi apparatus.

Individual incompatible FRT sequence pairs (FRT^{F3}, FRT^{F14}, FRT^{545}, FRT^{F13}, or FRT^{5T2}) were then placed in the opposing direction on both ends of the dMyr-FP / hH2B-FP / mMannII-FP sequence (Cai et al., 2013; McLeod, Craft, & Broach, 1986; Schlake & Bode, 1994; Turan, Kuehle, Schambach, Baum, & Bode, 2010; Volkert & Broach, 1986). Afterwards, each of the

five FRT flanked FP cassettes were assembled with an upstream activation sequence (UAS) and a p10 poly-adenylation sequence (p10pA) (Pfeiffer, Truman, & Rubin, 2012), upstream and downstream, respectively into a pJFRC-MUH backbone vector (Addgene #26213) by standard cloning methods. The Bitbow1.0 plasmids were finally created by constructing five individual modules together through Gibson assembly (Gibson et al., 2009).

For Bitbow 2 plasmids, the nSyb-promoter-driving self-excisable flippase module is constructed by flanking FlpINT (Cai lab, modified based on (Davis, Morton, Carroll, & Jorgensen, 2008)) with FRT^{F14} pairs or FRT^{F15} pairs in the same direction, and then introduced downstream of a Drosophila n-Synaptobrevin promoter (Addgene #46107). The module is then added to mBitbow1.0 plasmid through Gibson Assembly to generate mBitbow2.0 or mBitbow2.1.

The final Bitbow plasmids were integrated into *Drosophila melanogaster* genome docking site attP40, attP2 or VK00027 through the ΦC31 integrase-mediated transgenesis systems (Bateman, Lee, & -ting Wu, 2006; Bischof, Maeda, Hediger, Karch, & Basler, 2007; Groth, Fish, Nusse, & Calos, 2004; Markstein, Pitsouli, Villalta, Celniker, & Perrimon, 2008; Venken, He, Hoskins, & Bellen, 2006). The targeted genome insertion was carried out by BestGene Inc (Chino Hills, CA). All generated fly lines are listed in Table 2.1.

Confocal Microscopy and Linear Unmixing. Confocal images were acquired with Zeiss LSM780 with 20x 1.0 NA water immersion objective or 40x 1.3 NA oil immersion objective. The 32-channel GaAsP array detector was used to allow multi-track detection of five fluorophores. For detailed setups of spectral ranges see (Fig 2.3).

Spectral Unmixing plug-in (by Joachim Walter) in Fiji was used to perform linear unmixing on Bitbow images. Reference unmixing matrix was measured by imaging cultured

mouse N2A cells expressing mAmetrine, mTFP, mNeonGreen, mKO2 or tdKatushka2, with the exact same multi-track setup intended for Bitbow brains. Customized scripts were used to automate the unmixing process as well as creating composite image stacks from unmixed channels.

Statistics. All descriptive quantifications were reported by mean with standard deviation. Theoretical barcode collision rate for 2^5 (1-localization), 2^{10} (2-localization), and 2^{15} (3-localization) barcodes were calculated with a closed-form solution:

$$N = n - b * (1 - (1 - \frac{1}{b})^n)$$

Where N denotes the number of collisions, n is the number of lineages to be labeled, and b is the number of barcodes available. Monte-Carlo simulation was done to estimate the number of animals needed to ensure full coverage of given number of lineages, with different barcoding capacities.

Chapter 3. Anatomical and Lineage Investigation of Drosophila Larval Serotonergic Neurons by Bitbow

3.1 Introduction

As two of the most fundamental modalities in defining neuronal subtypes, morphology and lineage identities are prevalently used to derive the most basic descriptions of neuron groups. The accuracy of subtype definitions are improving quickly as methodologies in each category become more powerful. Fluorescence based sparse labeling, such as MARCM (Lee & Luo, 1999; Yu, Chen, Shi, Huang, & Lee, 2009), Flp-out based methods (Evans et al., 2009; K. Ito et al., 1997), has made it possible to produce high-resolution morphology reconstructions of individual neurons; permanent tracing systems conducted by transgenic or viral deliveries has provided valuable information on lineage identities of specific neurons. Currently, there is no effective system to ensure anatomical and lineage analysis across many neurons in the same brain, which is crucial in revealing whole-system organization, and enabling analysis on animal-to-animal plasticity or global changes under physiological/pathological conditions in the future.

In the last chapter, Bitbow has shown its power in generating large amount of unique fluorescent barcodes that are suitable for anatomical or lineage analysis. Here we focus on a small but important group of neurons in the *Drosophila* central nervous system - the serotonergic neurons (Kasture, Hummel, Sucic, & Freissmuth, 2018; Vallés & White, 1988), and aim to explore the intrinsic heterogeneities among these neurons in terms of their anatomical and lineage properties, which are indicators and driving forces for their functional diversity.

3.2 Results

Bitbow and expansion microscopy permits high resolution visualization and tracing of serotonergic neurons

With previous success of applying Bitbow2 transgenes on serotonergic neurons, we determined to study morphological properties of these neurons in the same brain. We could clearly observe serotonergic neuron clusters in central brain, and bilaterally distributed serotonergic pairs in the ventral nerve cord in diverse Bitbow colors. However, although mBibow2.1 ensured high labeling coverage and diverse labeling colors, many morphological details, especially those regarding neurite terminals were not obtainable as it was restricted by the diffraction limit of light microscopy. It was critical to use other methods with better optical resolution to realize the full performance of Bitbow.

Protein-retention Expansion Microscopy (ProExM) (Tillberg et al., 2016) is a great tool to achieve super-resolution by physically expand the samples in an isotropic manner while maintaining cellular contents in their relative positions. Since ProExM worked by retaining all proteins in the tissue, we reasoned that the membrane-targeted Bitbow2.0 FPs should be compatible with the ExM protocol. We modified the ExM protocol with suggestions from a previous ExM application in *Drosophila* (Cahoon et al., 2017) to maximize the signal quality (Fig 3.1), and successfully acquired images of expanded brains (Fig 3.2) with conventional confocal microscopy. Drastically improved image resolution permitted clear visualization of previously unresolvable neuronal densities (Fig 3.2). Further, with the high-resolution Bitbow image stack, it became possible to generate faithful neuronal tracing with nTracer (Roossien et al., 2019), a Fiji plug-in developed in our lab that is specialized in user-guided tracing on multi-channel datasets.

Through careful tracing of serotonergic neurons in the abdominal segments of fly CNS, we were able to generate high-resolution neuronal reconstructions of these neurons in the same brain (Fig 3.3). All abdominal serotonergic neurons are unipolar, with a single primary neurite extruding from the cell body and quickly branching into more and thinner neurite branches, covering one to one and half segments with extensive arborizations. Collectively these arborization cover segments along the anterior-posterior axis, and mainly at the ventral half of the neuropile (Fig 3.4), which indicates the function of serotonergic neurons in participating and modulating sensory circuits. Except for the single neuron in the A8/9 segment whose soma is positioned posteriorly and projects anteriorly into the neuropile, all abdominal serotonergic neurons send their primary neurite to the contralateral side through the major horizontal neural tract in each segment.

Bitbow is able to resolve lineage relationships between 5HT neuron pairs

Since the majority of larval neurons are born in the second wave of neurogenesis and very little migration occurs in this process, it is very possible that the two serotonergic neurons in each hemisegment, whose somas are located very close to each other, are likely to be generated from the same lineage. However, since VNC NBs are tightly packed in each hemisegment during the development, it is equally possible that the two serotonergic neurons actually come from two different lineages which happen to be right adjacent to each other. It would be impractical to explore this lineage relationship with previous genetic tools, but with Bitbow, especially mngBibow1.0 which could generate up to 32767 unique barcodes, it would be possible to systematically investigate the lineage identities of serotonergic neurons.

We decided to use the driver line hsFlp;TRH-Gal4; which would drive expression in all serotonergic neurons, together with mngBitbow1.0 to reveal the lineage identities of serotonergic

neurons (Fig 3.5). Heat shock was conducted during the embryonic neurogenesis period (3-4 hours after egg-lay) to ensure Bitbow initiation in the progenitors of serotonergic neurons. Because of the genetic inheritance of Bitbow barcodes from progenitors to neurons and the low collision rate of the mngBitbow1.0 system, we expected to observe the same Bitbow barcodes in later developmental stages only if the serotonergic neurons are coming from the same lineage. Contrast to the expectation that both neurons in each hemi-segment came from the same progenitors, we observed neurons in the same segments across the larval brain (Subesophageal - SE, Thorax - T, Abdominal - A) always being labeled by different Bitbow barcodes (Fig 3.5). In summary, Bitbow lineage tracing indicated that the majority of serotonergic neuron pairs in each hemi-segment of the larval CNS came from different lineage origins.

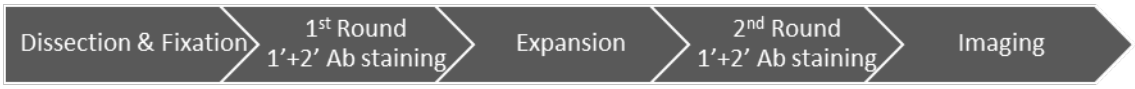


Fig 3.1 Experimental procedures of applying ExM on Bitbow labeled brains

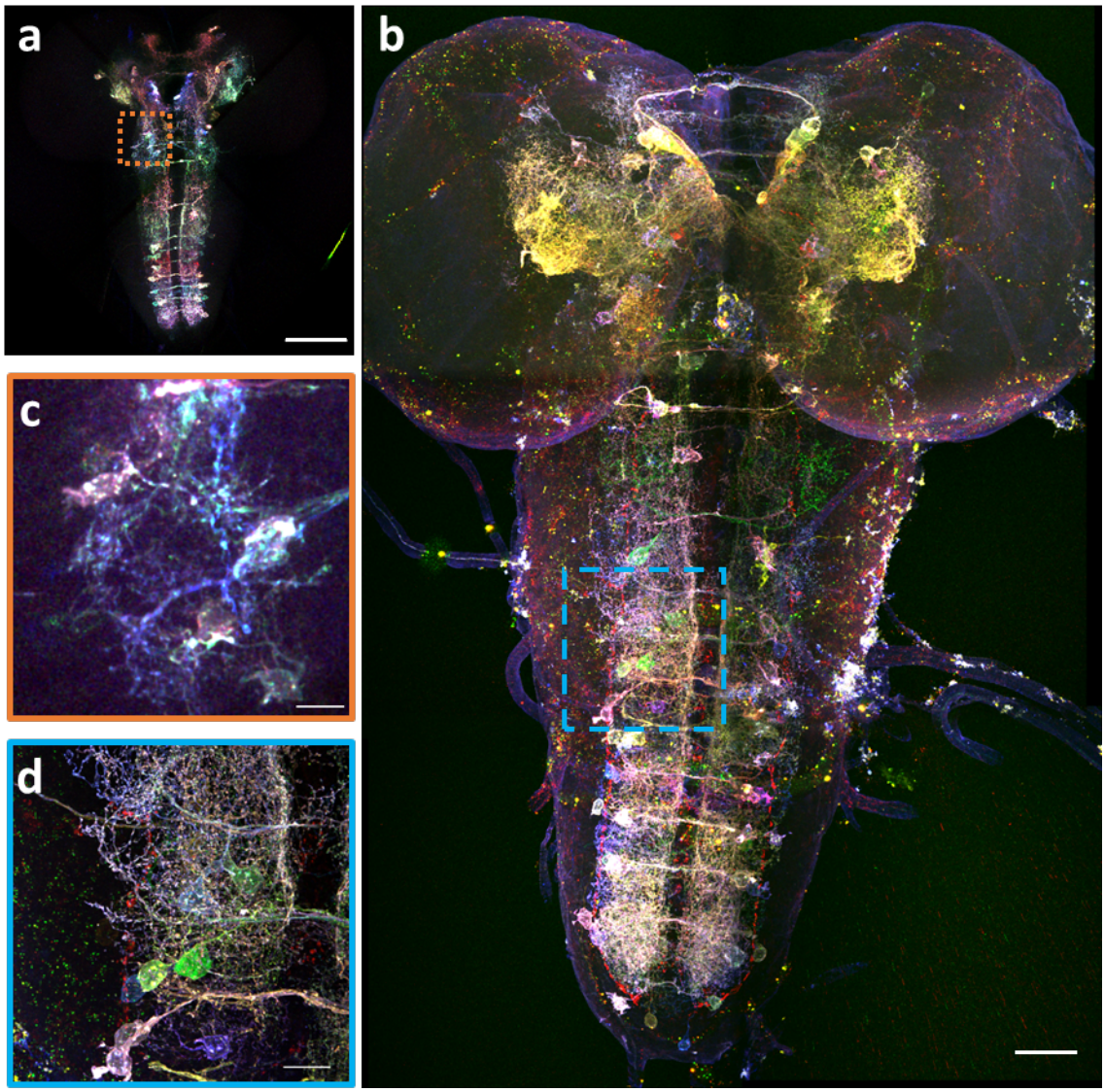


Fig 3.2 A third instar brain with mBitbow2.1 labeled serotonergic neurons was successfully processed with ExM

(a) Serotonergic neurons labeled by TRH-Gal4 driven 2x Bitbow2.1. Scale bar, 100um.
(b) ExM processed brain of TRH-Gal4 x 2x Bitbow2.1. The image is displayed in the same physical scale as in (b). Scale bar, 100um.
(c) Enlarged display of marked region in (b). Scale bar, 10um.
(d) Enlarged display of marked region in (c). Scale bar, 35um.

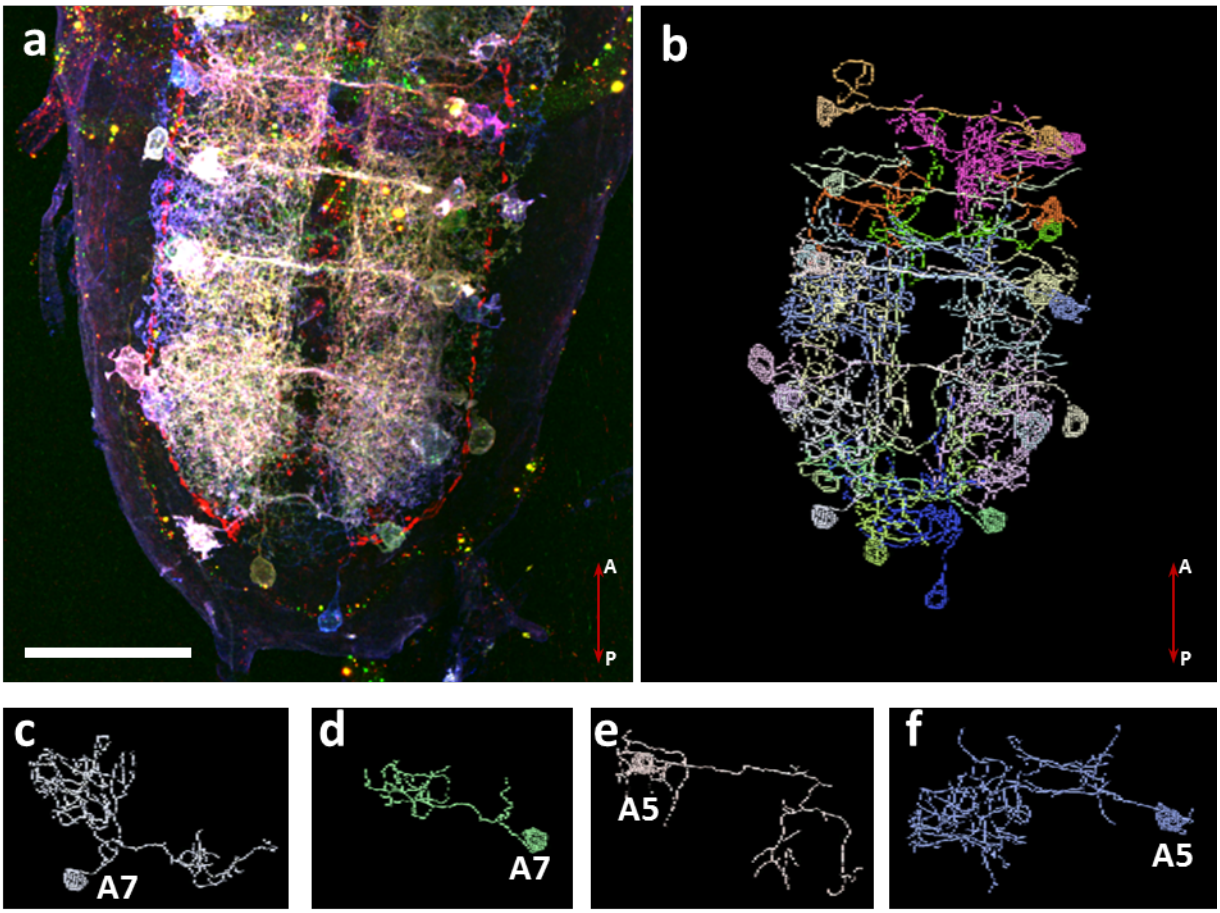


Fig 3.3 Overview of traced and reconstructed abdominal serotonergic neurons

(a) Abdominal VNC portion of the expanded Bitbow serotonergic neurons. Anterior-posterior axis is shown to the bottom right. Scale bar, 100um

(b) Rendered 3D display of traced serotonergic neurons in the abdominal VNC.

(c-f) Four traced neurons, two in segment A7 and two in segment A5 are individually displayed.

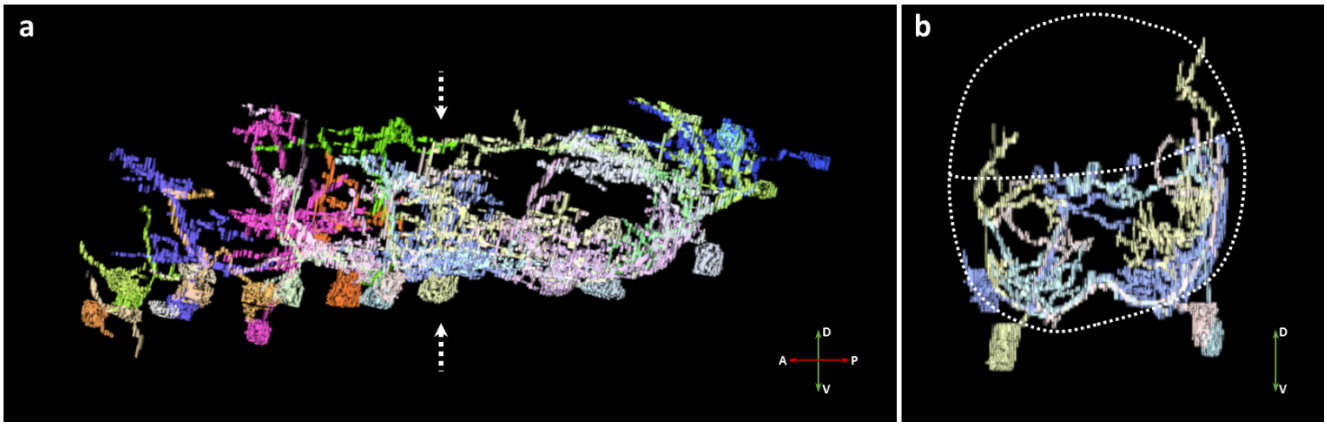


Fig 3.4 Neuropiles of abdominal serotonergic neuron dominantly covered the ventral half of VNC

(a) Side view of traced serotonergic neurons. White dashed arrows indicate the plane which (b) is corresponding to, containing neurons from segment A5. Relative coordination is shown at bottom right, A - anterior, P - Posterior, D - Dorsal, V – Ventral.

(b) Anterior view of four A5 neurons. Relative coordination is shown at bottom right.

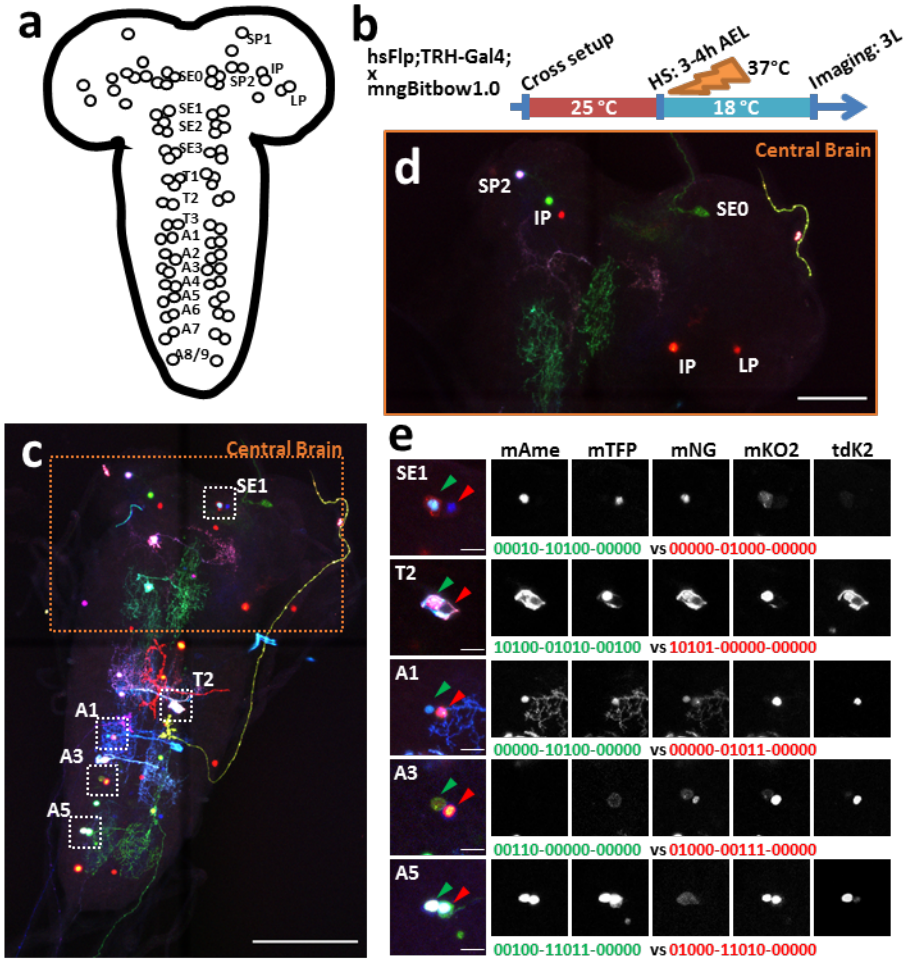


Fig 3.5 lineage tracing of serotonergic neurons with 3-Localization Bitbow1.0

(a) Clusters of serotonergic neurons in the 3rd instar brain.
(b) Schematics of experimental setup.
(c) A 3rd instar brain with serotonergic neurons labeled by mngBitbow. Scale bars, 100um.
(d) Zoom-in view of the central brain area in (c). Cell cluster identities are estimated based on relative cell body location.
(e) Pairs of serotonergic neurons in the same hemi-segments from (c). Scale bars, 10um.

3.3 Discussion

A large body of work has illustrated the importance of serotonergic neurons in various critical neuronal functions in flies and other animals (Gaspar et al., 2003; Lillesaar et al., 2007; Sodhi & Sanders-Bush, 2004). While all commonly producing the slow-acting neuromodulator, serotonin, serotonergic neurons are intrinsically heterogeneous as they participate in many aspects of brain functions in a non-collective manner. It is intriguing if we can understand the cellular and circuitry mechanisms relevant to the diverse functions of these neurons. Understand the exact meaning of diversity of each individual serotonergic neuron would lay the foundation of discovering the correlative and causative relationships to their brain functions.

With more and more systematic anatomical and lineage information in hand, we will be able to precisely categorize these neurons. Individual anatomical features will disclose potential circuits participated by individual neurons from known relationships between given neuropile to given function, and collectively anatomical features of all serotonergic neurons could elucidate their convergence and divergence in shared neural tracts, hence possible collaborative or parallel roles. Lineage identities of this diverse group of neurons could help to explain the determination of final neuronal types, and allow future lineage tracing and comparisons which could establish the mechanism of neuron subtype generation with emphasis on critical patterning steps. Bitbow has made systematic studies on these two aspects effective. By studying more animals to generate a common pattern, we will have a much deeper understanding about this critical group of neurons. It is also foreseeable when combining the knowledge in these two modalities with others, such as molecular properties, the serotonergic neuron subtypes can be thoroughly and accurately defined, to allow further studies of their specialized function in normal and disease conditions, and also translatable in other comparable species' nervous systems.

3.4 Materials and Methods

Fly husbandry. Flies were reared at 25c on standard CT food. Details of the acquired and created fly lines see Chapter 2 Methods.

Dissection and mounting. *Drosophila* 3rd instar and adult brains were dissected in PBS at room temperature (RT) within 30min before proceeding to fixation. Dissected brains were fixed in 4% PFA (Sigma #158127, diluted in PBS) at RT with gentle nutation for 20min, followed by three quick PBST (PBS+1% Triton X-100) washes, then PBS washes for 15min x 3. Brains then either proceeded to direct mounting (for native fluorescence imaging) or immuno-stainings. Vectashield (Vector Laboratories, H-1000) were used as the mounting medium.

Expansion Microscopy (ExM). ExM brain samples were generated following the ProExM protocol (Tillberg et al. 2016) with modifications. Antibody-stained Bitbow samples were treated in NHS-ester (Sigma, A8060) at RT for 1 overnight, followed by PBS washes for 15min x 3. Samples were then incubated in ExM monomer solution (Acrylate, Acrylamide, Bis-acrylamide) at 4c for 1 overnight. Samples were transferred to fresh ExM monomer solution with gel initiators (APS, TEMED, 4-HT) at 4c for 15min, and then quickly mounted on a sample chamber made with 200um adaptors (Sun lab) on a glass slide, sealed with a 22x30 coverslip on top (Fisher, 12-544). The slide was transferred to a humidity box, and incubated at 37c for 2 hours to polymerize the gel. The gels were trimmed carefully with a razor to allow as little of excessive space around the brains as possible. Trimmed gel pieces were transferred to an EP tube and digested with Proteinase K (NEB, P8107) at 37c for 1 hour. Three quick PBST washes and PBS washes for 15min x 3 were done, before the brains were put into the second round of antibody staining, following the same IHC protocol as above. After the second-round staining,

the gels were slowly expanded to the final size by changing the submerging solution from PBS to pure diH2O, and ready for imaging.

Confocal Microscopy. Confocal images were acquired with Zeiss LSM780 with 20x 1.0 NA water immersion objective or 40x 1.3 NA oil immersion objective. The 32-channel GaAsP array detector was used to allow multi-track detection of five fluorophores.

Tracing and reconstruction. Image stacks of mBitbow labeled brains were loaded into Fiji plugin nTracer (Roossien et al., 2019), and serotonergic neuron morphologies were constructed with user-guided semi-automatic tracing. 3D-skeleton image stacks were generated within the built-in nTracer function, and were opened in Vaa3D (Peng, Ruan, Long, Simpson, & Myers, 2010) for 3D reconstruction and display.

Chapter 4. Molecular Subtype Investigation of Drosophila Larval Serotonergic Neurons by Single-cell RNA Sequencing

4.1 Introduction

Single cell RNA sequencing (scRNAseq) brings new opportunity in revealing transcriptomic information in large number of cells with high throughput. Comparing to other candidate-based RNA probing methods, such as quantitative PCR or microarray chip, scRNAseq is advantageous as an open-ended discovery method; comparing to pooled RNA sequencing, scRNAseq makes it possible to reveal molecular identity of individual cells in heterogeneous cell populations. These advantages make scRNAseq perfectly suitable for resolving transcriptomic subtypes in the complex nervous system.

There has been efforts in applying scRNAseq to the nervous system in various species, including flies (Davie et al., 2018; Klein et al., 2015; Li et al., 2017; Macosko et al., 2015). These studies have proven the power of scRNAseq in revealing large scale differences among different subsets of neurons within the nervous system, or global changes as development progresses or physiological/pathological conditions incur. With these demonstrations of the technology, we reasoned that it would also be very suitable for investigations in finding molecular subtypes among the serotonergic neurons, which contribute to many kinds of neurobehavioral functions in flies. We determine to adapt current protocols to make it possible to specifically select this relatively small group of neurons, and process them through the scRNAseq pipeline based on the 10X Chromium platform. We aim to validate our dataset with known knowledge about the serotonergic neurons, make new discoveries of subtypes that could

be defined by specific gene set. This will pave the way for precise characterization and manipulation of these serotonergic neuron subtypes in future studies.

4.2 Results

Targeted collection of larval serotonergic neurons for scRNAseq

Serotonergic neurons were labeled in the progeny larvae from crosses of transgenic lines TRH-Gal4 (Alekseyenko, Lee, & Kravitz, 2010) and UAS-H2B-2xNeonGreen (Cai lab, this study). Specificity of the labeling was confirmed by cell body positions through imaging (Fig 4.1). A total of 53 late 3rd instar larval brains were dissected in 1 hour and quickly proceeded to brain dissociation. The dissociation was effective based on imaging quantification of the cell population (Table 4.1), where 3960, or 93.4% of the NG+ cells were retained (53x80 = 4240 total NG+ cells as 100% in estimation). The cells were then selected by FACS, where 1360 NG+ cells (32.08% of estimated total) were collected. After a quick wash and gentle pellet down, a final 1140 of NG+ cells (26.89% of estimated total) were collected and proceeded to 10X Chromium pipeline to generate the single cell RNA library. The library was sequenced on the NovaSeq 6000 platform. From the Cell Ranger report with preliminary filtering, 308 cells (7.26% of estimated total) were present in the dataset with >1.9 million mean reads per cell and 3660 mean genes detected per cell. Based on an estimation of 80 total serotonergic neurons in the larval nervous system, and an assumption that our cell selection method was unbiased, our data represented a 3.85x coverage of targeted larval serotonergic neurons. This scRNAseq experimental pipeline is the first of the kind to successfully process a rare neuronal population in the *Drosophila* nervous system.

The scRNAseq dataset contained high-quality reads from serotonergic neurons

After demultiplexing and mapping, 299 cells passed through with at least 1000 genes detected per cell. The number of genes (nGene) and number of deduplicated reads (nUMI) presented clear single peak distribution (Fig 4.2), with means of 3991 and 71812, respectively. Mitochondrial read counts appeared to be low among all cells, also with a unimodal distribution and a mean at 3.5% in each cell (Fig 4.2), which is a sign of gentle and proper cell dissociation. When graph-based cell clustering was conducted using top 50 dimensions from Principal Component Analysis (PCA), followed by 2D visualization with Uniform Manifold Approximation and Projection (UMAP), it was clear that there is no nGene or nUMI bias in any cell clusters (Fig 4.3), indicating a fair collection and representation of all serotonergic neuron in the system. Overall, this dataset reflected a batch of sequencing prep with high quality, where the sequencing power was evenly distributed in the population with minimum cell damage.

Collected cells were indeed serotonergic neurons with high purity. Cells in the population universally expressed basic neuronal markers such as n-Synaptobrevin (nSyb) and embryonic lethal abnormal vision (elav), and were clearly lacking markers of neural progenitor cells (dpn) or glia (repo) (Fig 4.4). Genes associated with serotonin production were readily confirmed in this dataset. Canonical genes such as Tryptophan hydroxylase (Trh), Dopa decarboxylase (Ddc), Serotonin transporter (SerT) and Vesicular monoamine transporter (Vmat) was universally expressed across all cells with no significant bias (Fig 4.5).

Serotonergic neurons possess differential adoption of neurotransmitters and neuromodulators

One of the unsettled questions about the serotonergic neurons has been whether these neurons could also produce accompanying fast-acting neurotransmitters to serotonin, which is a slow-acting neurotransmitter and thought to mainly perform as a modulator. We examined key

components of the three fast-acting neurotransmitter production pathways (VGlut, ChAT, Gad1 for glutamate, acetylcholine, GABA, respectively), and found most of the serotonergic neurons did not contain transcripts of these genes (Fig 4.6). However, there were small but distinctive clusters of cells contained exclusively one of the three genes. This line of evidence indicated that although the majority of the serotonergic neurons were unlikely to concurrently use the fast-acting neurotransmitters with serotonin, there were small subset of serotonergic neurons that were capable to generate one of three neurotransmitters.

On the receiving end, we aimed to reveal the potential differences of the serotonergic neurons in responding to the fast-acting neurotransmitters. We found that glutamate-gated chloride channel (GluClalpha) was broadly expressed across the population, and on the contrary, other ionotropic glutamate receptors (GluRIA, GluRIB, GluRIIE, NMDAR1, NMDAR2) were less prevalent, while the metabotropic glutamate receptor (mGluR) only expressed in a small subset of the neurons (Fig 4.7). We also found that subunits of nicotinic acetylcholine receptors (nAChRalpha1-7, nAChRbeta1-2) and muscarinic acetylcholine receptor A (mAChR-A) were broadly expressed among the serotonergic neurons, but mAChR-B had relatively low expression in only a small subset of neurons (Fig 4.8). GABA receptors also had very distinct expression patterns among the serotonergic neurons. Among the three ionotropic GABA receptors, Rdl had the broadest coverage and relatively high expression, Lcch3 had broad but much weaker expression, and Grd expression was only seen in a few cells with much lower expression. The three metabotropic GABA receptors also showed strong differences in expression coverage and level, where GABA-B-R1 being the broadest and strong, GABA-B-R2 being broad and weaker, and GABA-B-R3 being only in a few cells with weak expressions (Fig 4.9). Taken together, serotonergic neurons were equipped with necessary receptors to receive the three fast-acting

neurotransmitter signals, and the differential distribution of these receptors hint potential intrinsic differences among the serotonergic neurons to respond to these input signals, which could serve as strong molecular categorical evidence for accurate subtyping.

We also set to probe whether serotonergic neurons had the potential to receive dopamine or octopamine signal input. By surveying all four dopamine receptors (Dop1R1, Dop1R2, Dop2R, DopEcR) and five of the six octopamine receptors (Oamb, Oct-TyrR, Octbeta1R, Octbeta2R, Octbeta3R) in the dataset, we discovered in general there was broad expression of these receptors, with the exceptions of Dop1R2, Oct-TyrR and Octbeta1R in which case they appeared to be strongly expressed but in a small subset of the neurons (Fig 4.10).

Finally, we also examined the expression of serotonin receptors among the serotonergic neurons, which could implicate their possible strategies in reuptake and feedback regulation. Of the 5 types of serotonergic receptors, we discovered that 5HT1A and 5HT1B had the broadest and highest expression among all serotonergic neurons; 5HT2A and 5HT2B had much lower expression and only in a few cells; 5HT7 had high expression restricted to a small subset of neurons (Fig 4.11). Since 5HT1A and 5HT2B were coupled with G protein Gi/Go associating with inhibitory cellular mechanisms, and 5HT2A, 5HT2B were coupled with Gq, 5HT7 was coupled with Gs associating with stimulative functions, the discovered expression pattern indicated a general feedback inhibitory mechanism in serotonergic neurons to prevent excessive production of serotonin, while in a small subset of neurons there could be feedforward mechanism to allow strong burst of serotonin expressions specifically.

Serotonergic neurons can be well clustered with GINNAT genes

Although graph-based clustering using all genes surveyed in the sequencing could group the serotonergic neurons into six clusters, the separation between each cluster was not optimal.

Clusters revealed on UMAP space were largely mixed with each other with no apparent boundary, indicating a lack of differentiating ability with current gene set (Fig 4.3). We reasoned that a smaller but more biological specific set of genes regarding GPCRs, IonChannels, Neuropeptides, NeurotransmitterSecretion, AxonGuidance, TFs ("GINNAT" set) could be helpful to provide stronger discerning powers for the clustering mechanism. A total of 1244 GINNAT genes was found in our dataset, and they were used to generate PCA dimensions and go throught the same graph-based clustering methods. We found the new clustering yielded much better separated subgroups. (Fig 4.12)

Looking closer at the six clusters proposed by the new clustering pipeline, we have found strong differentiating makers to represent each group. Three of the six clusters, #3, #4, and #5 as marked in the dataset, contained the cleanest separation with a set of marker genes with strong differential expression. Cluster #3 can be well described by transcription factors bi (T-box family), Fer2 (bHLH family), gcm2, zld (zinc finger family) and an undefined CG43689, which was predicted to contain zinc finger transcription factor features. Cluster #4 can be described by transcription factors slp1, slp2 (fork head box family), Vsx1, Vsx2 (Paired-like homeobox family), sim (bHLH family). Cluster #5 can be described by transcription factors disco, disco-r (C2H2 zinc finger family), Ets21C (ETS family), Dr (NK-like homeobox family) and a neuropeptide sNPF (homolog of mammalian neuropeptide Y, NPY). These maker genes all appeared high coverage, high expression among the defined neurons but essentially absent in all the other neurons (Table 4.2). The other three cluster of neurons, #0, #1 and #2 also contained sets of differential genes with good statistical power (all adjusted p values less than $1x10^{-5}$), however these maker genes still possessed non-exclusive expressions outside of the designated

clusters, making their differentiation power limited, and making the three clusters appeared more close to each other on the UMAP display (Fig 4.12).

Finally, to confirm the new clustering produced with the selected genes was consistent with our previous nearest-neighbor estimation based on all genes, we examined the expression patterns of all the marker genes on the full-gene-set UMAP display (Fig 4.13). We found genes for all six clusters of neurons, even the less differentiating genes, could specifically label subsets of neurons that clustered together. This strongly suggested that the intrinsic common characteristics of serotonergic neuron subsets were well maintained with our GINNAT gene sets, and the marker genes revealed above could serve as powerful markers or indicators for future validation and manipulations.

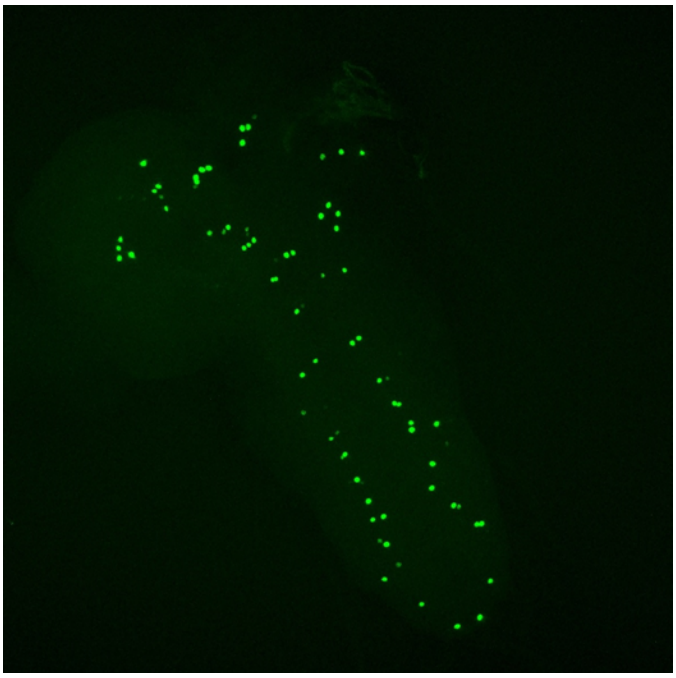


Fig 4.1 Larval serotonergic neurons labeled by nucleus-localized mNeonGreen
3rd instar larva brain from the cross between TRH-Gal4 and UAS-H2B-2xmNeonGreen. Brain is anterior-posterior organized from top-left to bottom-right.

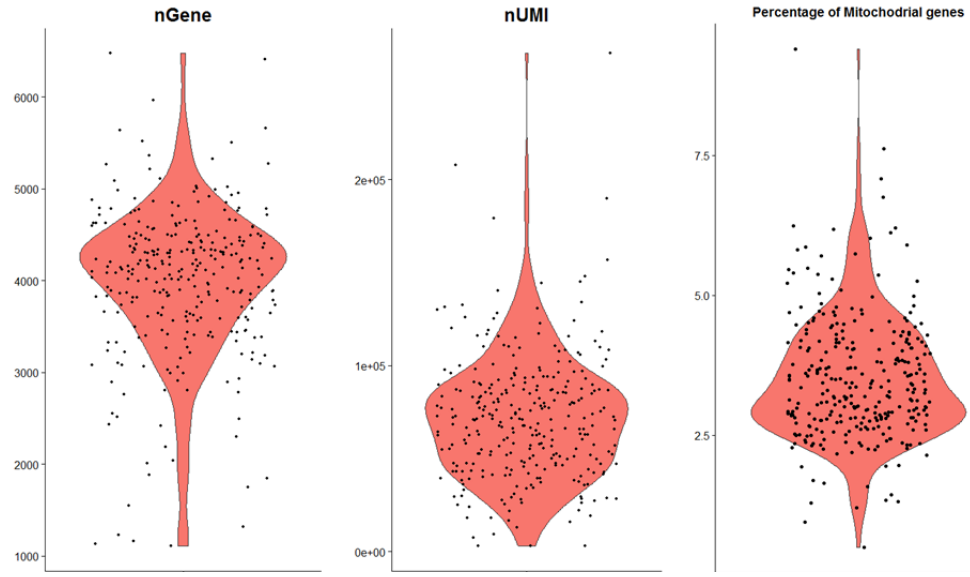


Fig 4.2 Basic statistics of sequenced cells
Violin plots with scattered dot displaying number of genes, number of de-duplicated reads (UMIs) and mitochondrial content of all cells that are passed filtering step.

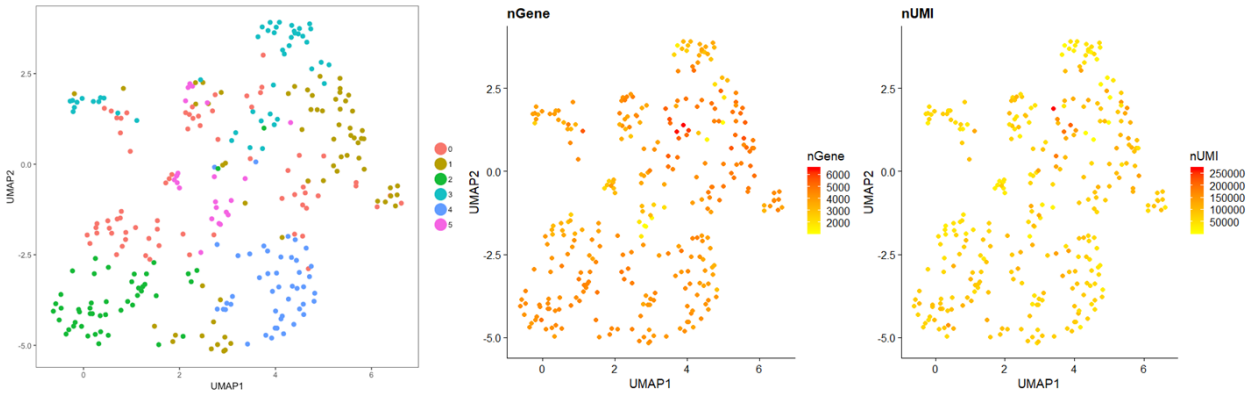


Fig 4.3 Dimensionality reduced visualization shows 6 cell clusters based on their similarity in mRNA levels of all genes, and distribution of nGene & nUMI among the cells

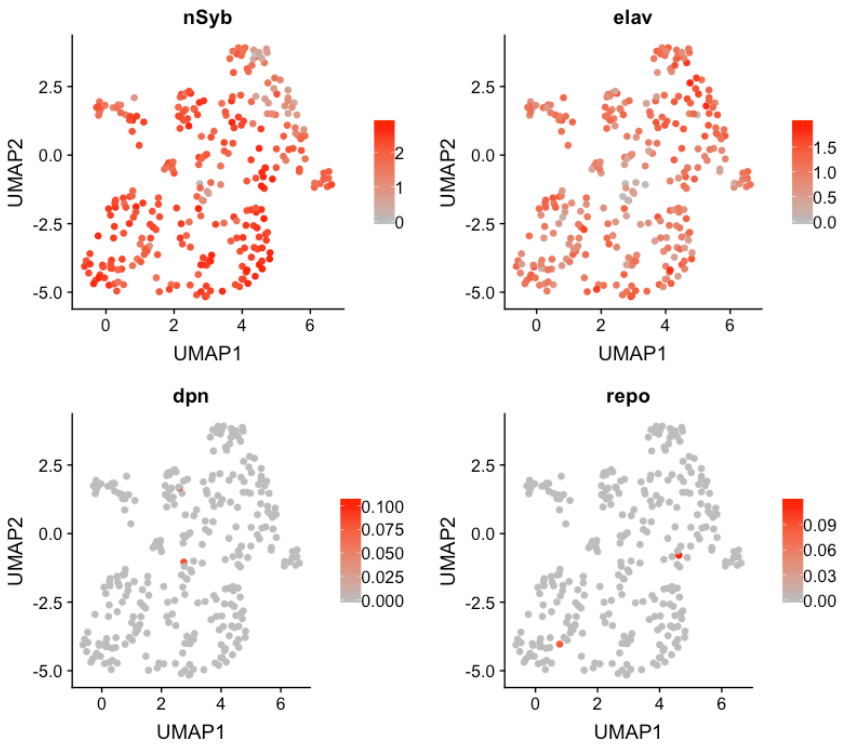


Fig 4.4 Expression of neuron (nSyb and elav), neuroblast (dpn) and glia (repo) marker genes

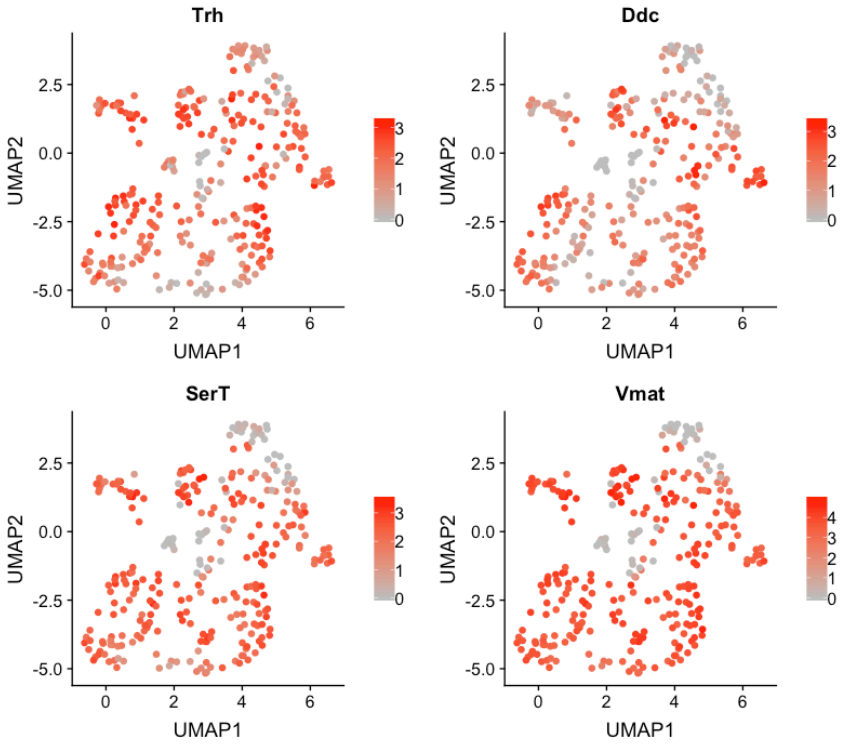


Fig 4.5 Expression of Serotonin synthesis- (Trh and Ddc) and transport-related (SerT and Vmat) genes

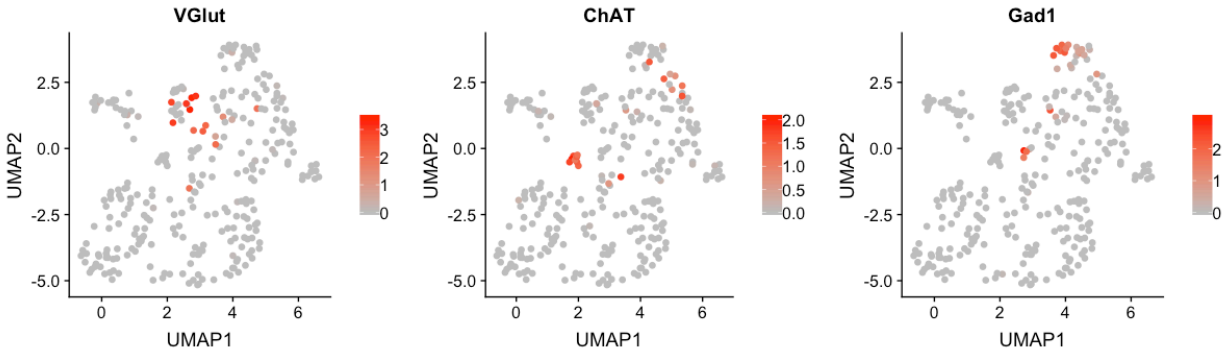


Fig 4.6 Expression of genes related to fast-acting neurotransmitters glutamate (VGlut), acetylcholine (ChAT) and GABA (Gad1) production

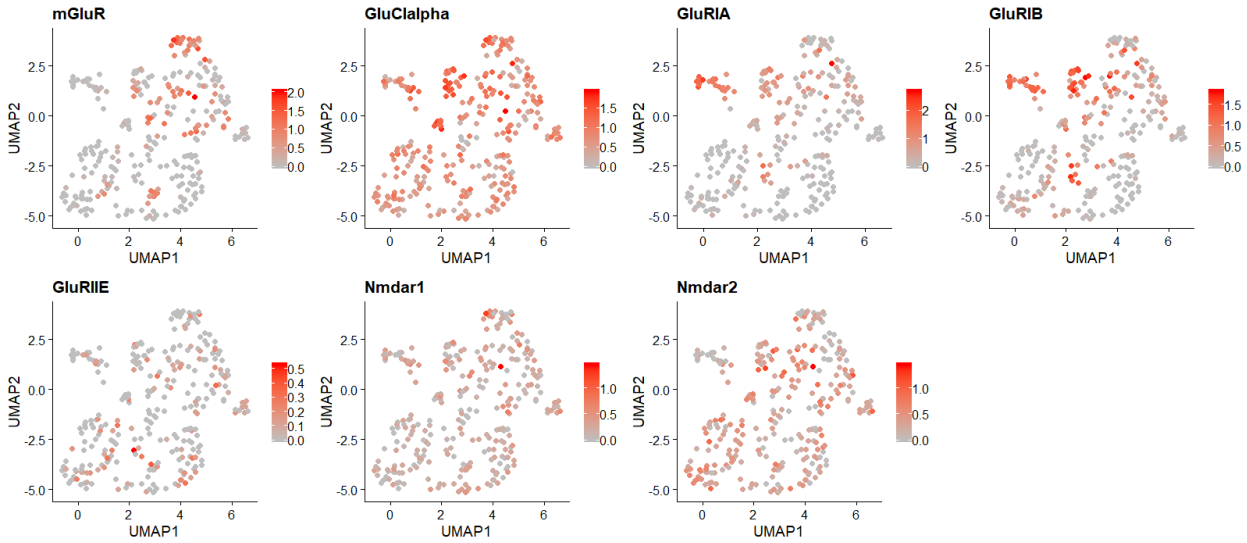


Fig 4.7 Expression of glutamate receptor genes

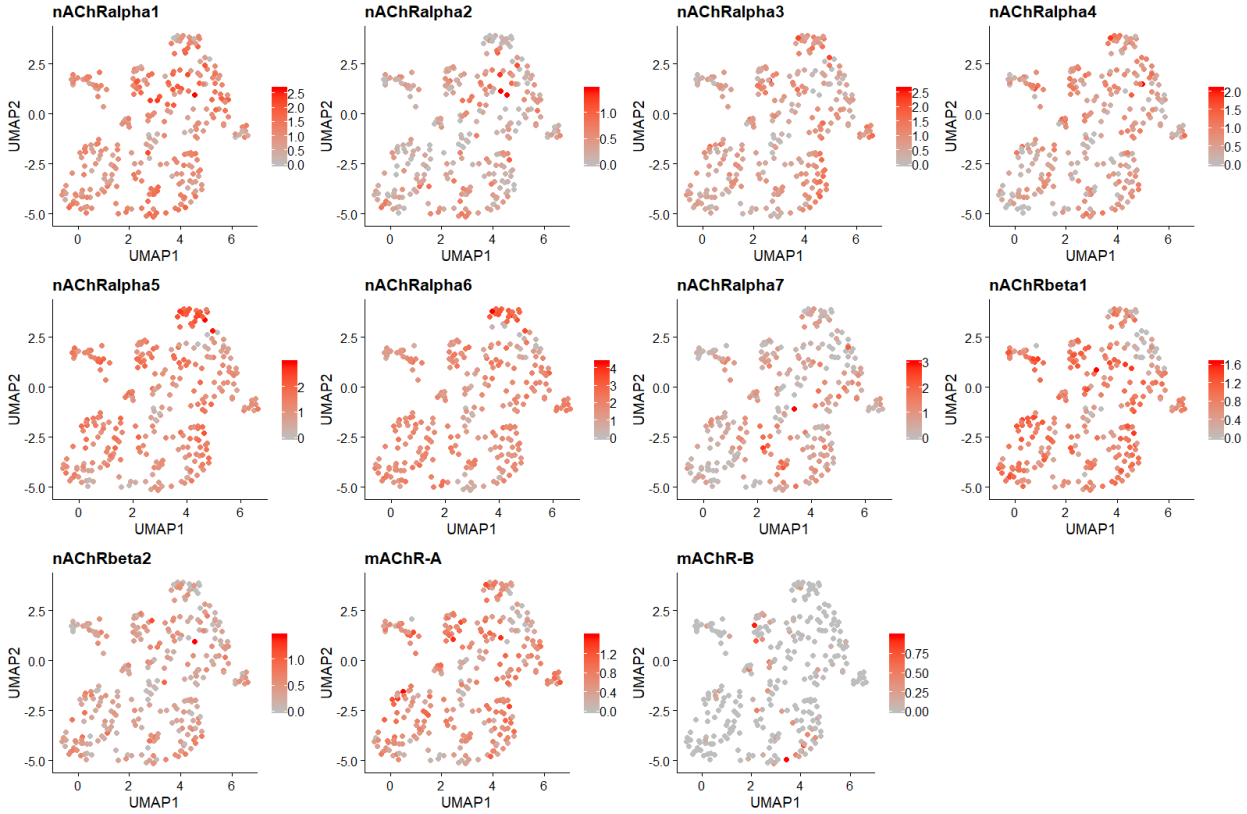


Fig 4.8 Expression of acetylcholine receptor genes

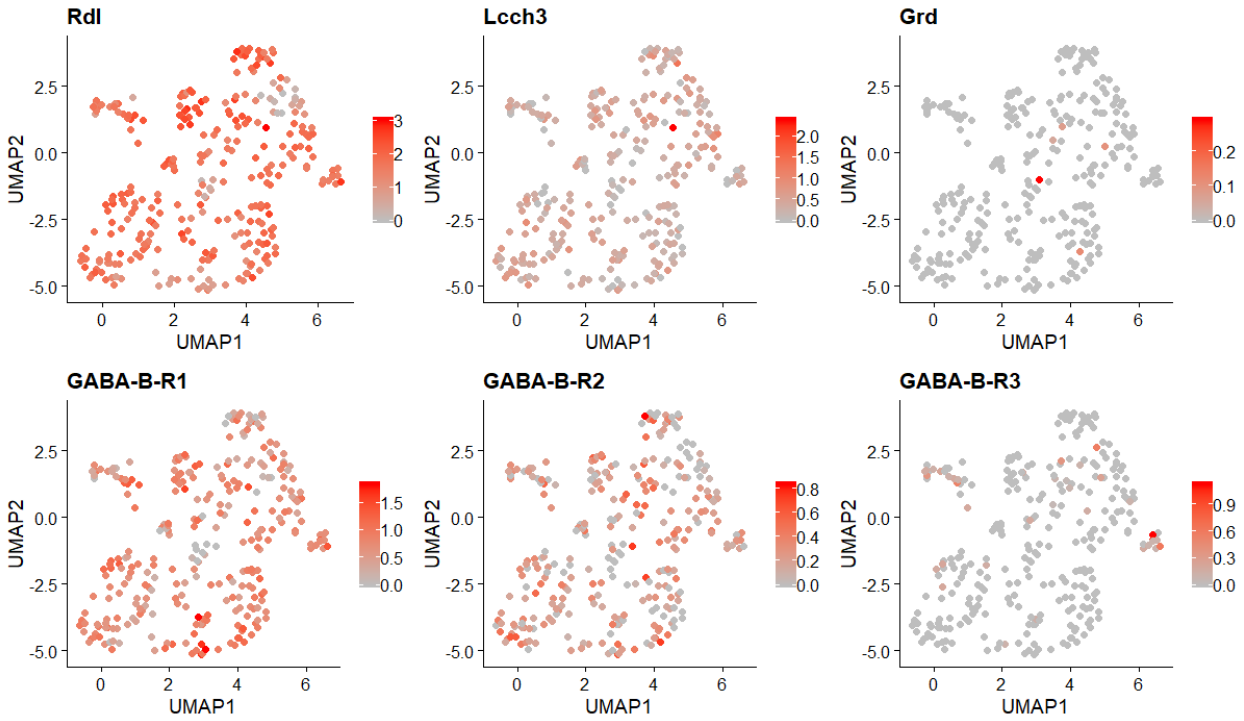


Fig 4.9 Expression of GABA receptor genes

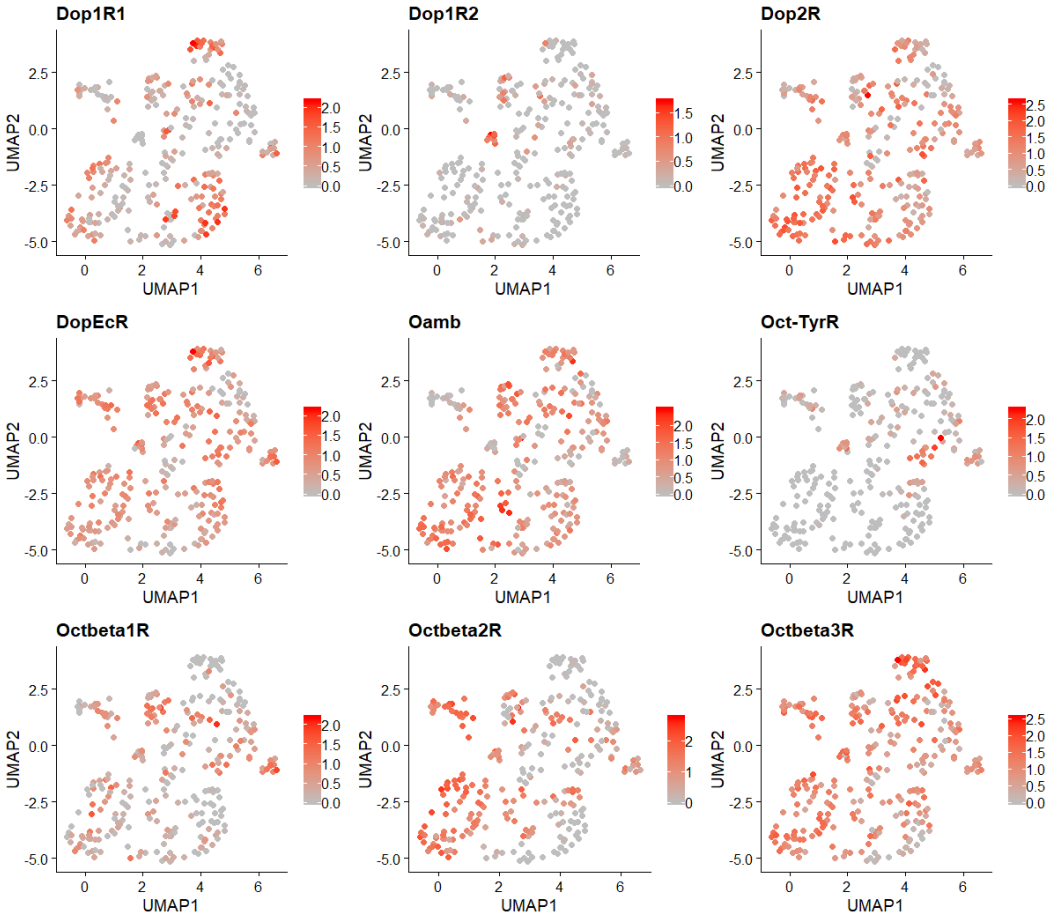


Fig 4.10 Expression of dopamine and octopamine receptor genes

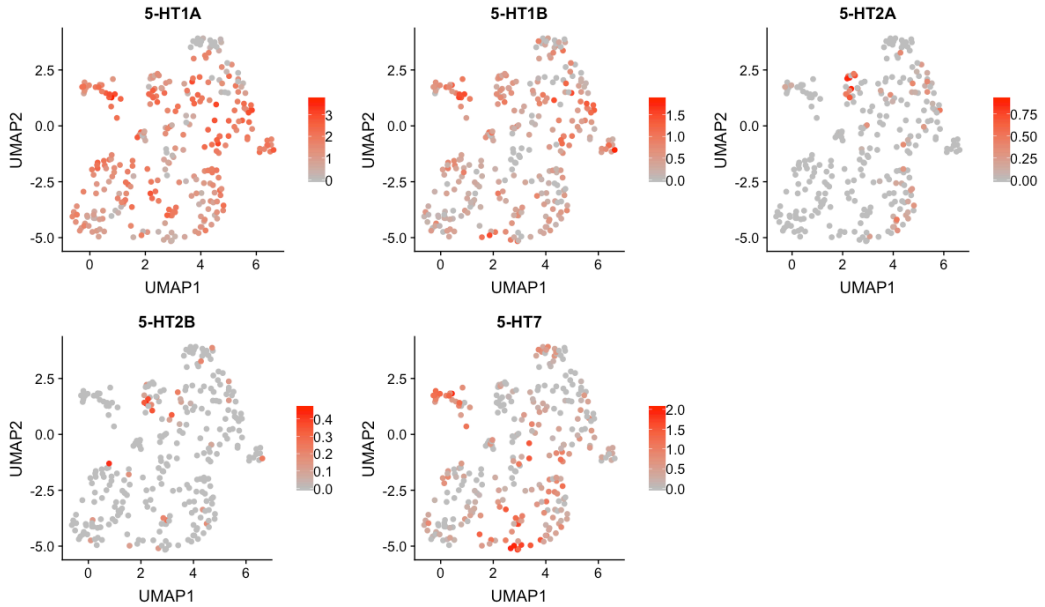


Fig 4.11 Expression of serotonin receptor genes

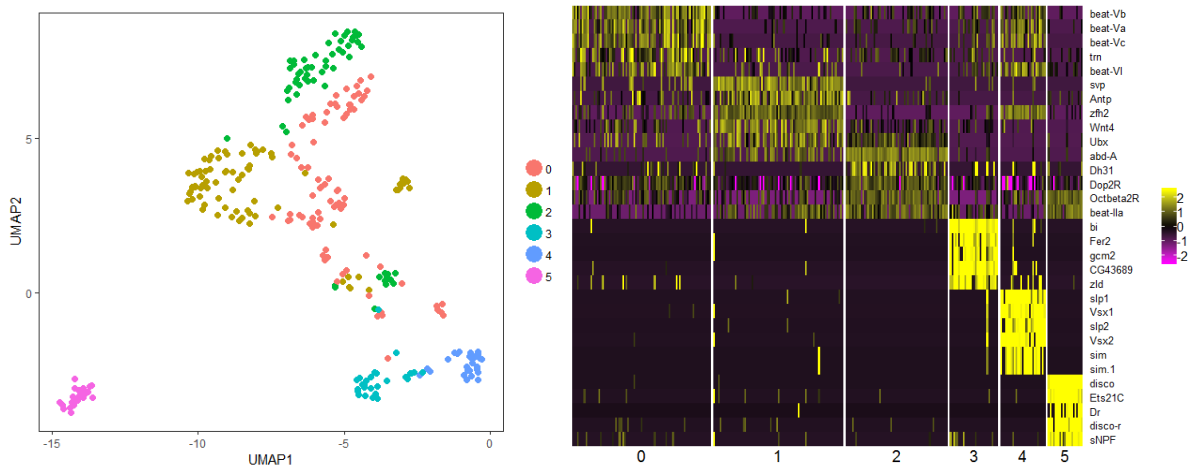


Fig 4.12 Dimensionality reduced visualization of single cell clusters defined by the "GINNAT" gene set
(left) Dimension reduced visualization of all cells with five detected cell clusters.
(right) Heatmap display of top-5 marker genes in each cluster. Horizontal axis, individual cells grouped by clusters; Vertical axis, marker genes. Expression levels are natural-log transformed and color coded.

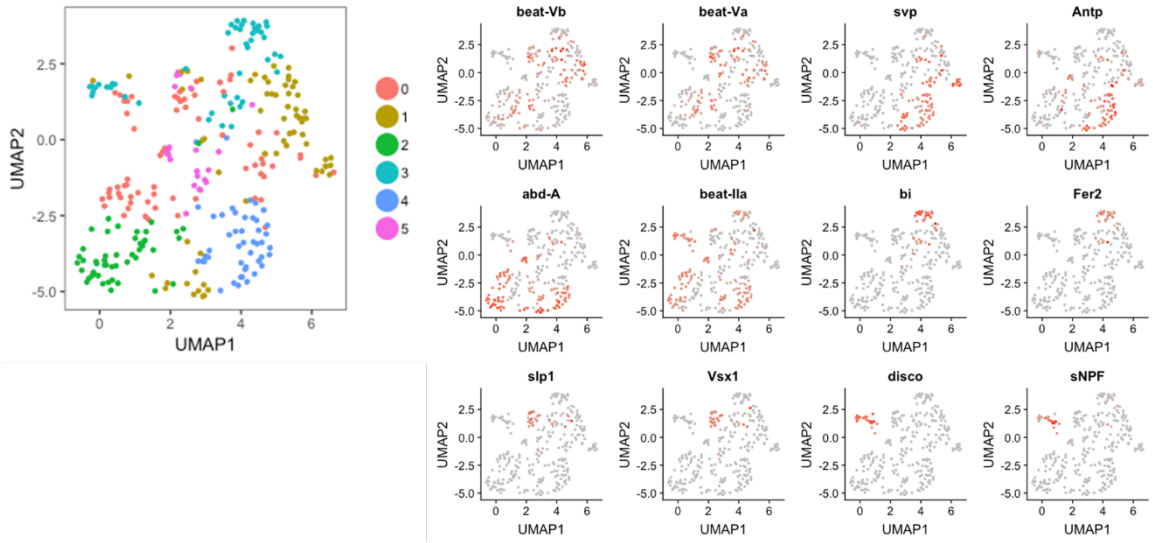


Fig 4.13 Expression of "GINNAT" marker genes mapped on the full-gene-set derived dimension-reduction display

Table 4.1 Cell counts after each step in dissociation and sorting

Step	Cell Count	% of previous step	% of first step
Dissection	4240 (estimate)	-	100%
Dissociation	3960	93.40%	93.40%
FACS Sorting	1812 (collected)	45.76%	42.74%
Count after FACS	1360	75.06%	32.08%
Centrifuge & transfer	1140	83.82%	26.89%

Table 4.2 Marker genes for clusters resolved by dimension reduction based on the GINNAT gene set

cluster	gene	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene	avg_logFC	pct.1	pct.2	p_val_adj
0	beat-Vb	1.4204209	0.904	0.435	1.40E-19	3	bi	3.8612524	1	0.03	7.28E-50
0	beat-Va	1.4669871	0.843	0.301	1.13E-18	3	Fer2	2.2903812	0.897	0.011	1.01E-49
0	beat-Vc	1.4780523	0.795	0.241	2.90E-17	3	gcm2	1.8666414	0.862	0.007	8.20E-49
0	trn	1.2732277	0.747	0.403	7.79E-08	3	CG43689	3.6018055	0.966	0.03	2.47E-47
0	beat-VI	0.9129595	0.663	0.329	7.87E-06	3	zld	2.7564391	0.966	0.067	4.19E-35
1	svp	2.033477	0.91	0.104	2.85E-35	4	slp1	3.3559311	0.963	0.011	2.34E-54
1	Antp	1.5165184	0.872	0.249	1.01E-22	4	Vsx1	3.5442972	0.963	0.015	5.19E-53
1	zfh2	1.1819142	0.987	0.394	1.21E-19	4	slp2	2.9517363	0.926	0.011	3.78E-52
1	Wnt4	2.0792579	0.795	0.385	3.31E-16	4	Vsx2	3.3347155	0.926	0.015	4.05E-50
1	Ubx	1.3366246	0.718	0.376	3.39E-08	4	sim	2.1832845	0.815	0.007	3.64E-46
2	abd-A	1.674062	0.967	0.286	3.94E-26	5	disco	3.7071872	1	0.007	2.59E-58
2	Dh31	0.6653482	0.689	0.164	1.70E-13	5	Ets21C	3.2845108	1	0.043	1.57E-41
2	Dop2R	0.7800274	0.984	0.912	3.44E-11	5	Dr	1.4680166	0.667	0.004	5.90E-38
2	Octbeta2R	0.9497823	0.951	0.622	2.33E-10	5	disco-r	3.3692599	0.952	0.05	1.71E-36
2	beat-IIa	0.8994759	0.918	0.546	5.13E-09	5	sNPF	4.240777	1	0.086	7.89E-31

4.3 Discussion

The fast advances in single cell RNA sequencing technology have made deep transcriptomic profiling of large amount of cells practical. Transcriptomic information provides a valuable exploratory entry point for mechanistic studies in any given type of cells, since this is a direct snapshot of molecular machineries within a population of cells, which is highly predictive of cellular functions at given biological conditions. Present success of scRNAseq application in systems such as peripheral blood mononuclear cells (Zheng et al., 2017), cancer tissues (Guo et al., 2018), early developing embryonic stem cells (Tang et al., 2010) and so on have provided valuable insights of how these dynamic systems might function in great detail. As the cell processing throughput and the sequencing quality advance in scRNAseq, it is time for the field to move onto more complicated applications, such as in highly heterogeneous tissues e.g. the whole nervous system (Croset et al., 2018; Davie et al., 2018), highly dynamic processes e.g. the developmental program in human epiblast and the later stages (Tang et al., 2010, 2009; Yan et al., 2013; Zhou et al., 2019). Through robust sequencing protocols and effective bioinformatics analysis, these challenging goals are becoming reachable.

However, there are still several difficulties that could not be easily surmounted solely by the advancement in sequencing power or analysis algorithms. For one, it is not always trivial to maximize the sequencing power only on selected group of cells, especially when facing more and more heterogeneous systems. ScRNAseq is very effective when the targeted cell populations are easy to obtain, and/or methods of selective separation exists, e.g. well defined surface markers of circulating blood cells. Otherwise, the sequencing pipeline would be troubled by inadequate cells collected for sequencing, or too much sequencing power wasted on unwanted cells. Effective selection method that is compatible with scRNAseq is critical in solving this

challenge. In this chapter, I presented a transgenic solution in the model system *Drosophila* fruit fly, in which a strong nucleus-localized fluorescent marker was used to enable the selection of a rare neuron population in the nervous system (80 serotonergic neurons out of estimated 10,000 total CNS neurons). Considering the size of *Drosophila* neurons, and the fact that these H2B-mNG labeled cells were cleanly separable on a conventional FACS sorter, this work has indicated a practical approach transferable to mammalian systems, whose cells are larger and are able to contain stronger fluorescence.

Also, current scRNAseq approaches are strongly challenged by the loss of positional/morphological information during cell dissociation. Not a critical issue for systems insensitive to cell distribution/morphology, but this is a big difficulty for others whose cell functional diversity are reflected in or depend on shape, structure, niche composition, etc. If reliable transgenic drivers, covering the whole or part of the desired cell population, are available, then strong reporters such as H2B-mNG could be practical to resolve and sequencing these cell populations one piece at a time. However, not all systems are as well-equipped to the level as the armory of *Drosophila* genetics tools, and simple specific drivers for targeted cell populations are not guaranteed. One possible direction that has not been covered by this thesis study is to target cell populations with correlatable markers that could connect cell morphologies and transcriptomes together. More specifically, if cell populations can be labeled with unique barcoding system such as Bitbow, and the pre-recorded barcoded cell morphology can be directly correlate to barcoded transcriptome, then the population composition can be accurately defined by the joint modalities.

Regardless of the shortcomings of current scRNAseq methodologies, this serotonergic data set has brought in a great amount of information that awaits solid molecular validations.

With available tools targeting known and novel markers at RNA and protein levels, many unanswered questions such as the fast-acting neurotransmitter composition, 5-HT/Dopamine/Octopamine receptor distributions, and master regulator/transcription factors in defining subtypes in the serotonergic system, can be addressed in future explorations. Further, with the discovered markers, new driver lines or combinatorial driver lines in collaboration with lines like TRH-SplitGal4 can be made for more precise subtype manipulation and investigations.

4.4 Materials and Methods

Fly husbandry. Flies were reared at 25c on standard CT food. TRH-Gal4 driver line was acquired from BDSC (#38388). The reporter line UAS-H2B-2xmNeonGreen was made by first constructing human H2B cDNA and 2 copies of mNeonGreen cDNA together in the same reading frame, downstream of the UAS sequences of the pJFRC-MUH backbone (Addgene, #26213); then the transgenic fly line was established by embryo injection and PhiC-31-mediated chromosome recombination into the attP2 site (BDSC #8622, injections carried out by BestGene Inc). Progeny of the cross were synchronized to one-day birth window by only allowing the parents staying in the same vial for 24 hours. On the fifth day after the cross, only wandering larvae were collected for the further experiment, in order to minimize developmental time differences between sequenced individuals.

Dissection and dissociation. Third instar brains were dissected in 1x Rinaldini's solution (Harzer, Berger, Conder, Schmauss, & Knoblich, 2013) at room temperature within 30min before proceeding to dissociation. During the dissection, tissues attached to the brain, such as the ring gland, imaginal discs and neuron fibers connecting to the mouth hook were all carefully removed to ensure dissociation and cell collection efficiency. All dissected brains were quickly

checked under an epifluorescence microscope to confirm strong and clear TRH neuron expression pattern before proceeding to dissociation. Brains were quickly washed twice with Rinaldini's solution and briefly spun down (with ~500gcf) to the bottom of an EP tube. In Rinaldini's solution with 2 mg/ml of Papain and 2 mg/ml of Collagenase I, the brains were incubated at 37c for 1 hour for dissociation. Gentle agitations were done by tapping the EP tube every 15min during the dissociation. 20uM of the protease inhibitor E-64 was added to stop the reaction. To fully dissociated cells from the brains, quick pipetting with a siliconized P200 tip was done for ~30 times and brains could be visually confirmed to be fully dissociated. Schneider's medium with FBS at 9x volume of the cell suspension was added to fully stop the reaction, and to maintain cells at a nutritious condition while getting ready for cell sorting.

Cell sorting and sequencing. Dissociated cells were loaded onto Sony SH800S cell sorter with a 100um-nozzle chip. Sequential gates on FSC/SSC to FSC/mNG were set up using a negative control sample which were cells coming from a dissociated non-fluorescent brain. The sorting was conducted at ~7000 events/s. Cell selection was set at "Purity" mode, in which case any negative droplets or positive droplets adjacent to any contaminant would be threw out to ensure stringent selection. It was estimated that the sorting efficiency was at ~55%, i.e. 55% of mNG-positive cells were eventually sent to collection channel. The sorted cells were collected in a 1.5ml EP tube with Schneider's medium to maximize their survival rate. Small portion of the sorted cell suspension was transferred into a well of 96-well plate to confirm their mNG fluorescence, cell morphology and estimate the cell concentration. Final cell concentration was adjusted following 10X Genomics' guideline, and the 10X Chromium prep (v3 chemistry) was done by the University of Michigan Advanced Genomics Core, followed by sequencing on the Illumina NovaSeq 6000 platform.

Bioinformatics. Raw sequencing FASTQ files were mapped to the newest version fly genome (Drosophila_melanogaster.BDGP6.95) plus the cDNA sequence of mNeonGreen, and UMI counts were generated using Cell Ranger (with help from Dr. Jingqun Ma). The cell-UMI data object was then passed into Seurat (Butler, Hoffman, Smibert, Papalexi, & Satija, 2018) for downstream analysis. Based on the distribution of number of genes and number of UMIs in the sequenced cells (Fig 4.2), cut-offs at 400 genes and then 1000 UMIs were set to filter out cells with poorer quality. UMIs in each cell were log-normalized, and the dataset was scaled by a linear regression model. Principal component analysis (PCA) was done to reduce the dimensionality. 50 PCs from the PCA were then used to conduct graph-based shared-nearest-neighbor (SNN) (Waltman & van Eck, 2013), and 2D visualization using t-distributed stochastic neighbor embedding (tSNE) (Laurens van der Maaten, 2008) or Uniform Manifold Approximation and Projection (UMAP) (McInnes, Healy, Saul, & Großberger, 2018). Markers of each cell cluster were selected with differential expression analysis based on the non-parametric Wilcoxon rank sum test; parameters describing each marker genes were generated, including the percentage of cells where the marker is detected in the defined cluster ("pct.1"), the percentage of cells where the feature is detected outside of the defined cluster ("pct.2"), and the adjusted p-value, based on bonferroni correction using all features in the dataset ("p_val_adj"). To improve the performance of clustering and marker gene detection, the gene list of the GINNAT set containing 1446 genes of GPCRs, Ion Channels, Neuropeptides, Neurotransmitter-Secretion molecules, Axon Guidance molecules, TFs was constructed from FlyBase Gene Groups, and used to generate a subset data including all cells with only these genes. Same pipeline of analysis was done to generate new cell clustering and the set of marker genes.

Chapter 5. Concluding Remarks

This dissertation elaborated the importance of gaining deeper knowledge about heterogeneous neuronal subtypes within the nervous system (Chapter 1), described an effort of generating the Bitbow tools to perform morphology and lineage tracing in the *Drosophila* nervous system (Chapter 2), and presented a multimodality investigation on the *Drosophila* serotonergic neuron subtypes using Bitbow and scRNAseq technologies (Chapter 3, 4).

Categorical study of system components is a critical step for all scientific endeavors. "What I cannot create, I do not understand." - the famous quote by the theoretical physicist Richard Feynman highlighted the reductionist philosophy that is fundamental to all branches of science. Just like having a part list of all components in a complex electronic circuit, obtaining an atlas with detailed neuron subtype descriptions will advance the understanding of the entire nervous system. Such part list provides an important entry point for depicting how a "black box" work, where measurements can be confirmed, manipulations can be traced, and responses can be correlated. This allows scientists to effectively establish, test, accept or reject hypotheses to understand the brain functions.

Neuroscience studies are propelled by advancements in tools. The boundary of our knowledge is always set at the furthest point where our last generation of methodologies could reach. Breakthroughs happen whenever the technologies have advanced to reach greater resolution or precision, explore new properties, or extend the measurements to a higher dimension or complexity. The neuron doctrine was established when the silver staining method was invented to make precise neuron anatomical characterizations possible (Glickstein, 2006).

Action potentials as the form of signal propagation was quantitatively characterized when voltage clamp was available (Hodgkin & Huxley, 1952a, 1952b, 1952c, 1952d; Hodgkin, Huxley, & Katz, 1952). Persisted lineage tracing was practical when creative transgenic designs and microscopy methods advanced to ensure higher spatial and temporal resolutions (Kretzschmar & Watt, 2012). In this dissertation study, Bitbow utilized a concise and expandable genetic design to make it suitable for generating a large number of color barcodes for morphological and lineage studies. Tunable genetic switches, either externally controllable by heat-induced transgenics, or internally programmed by neuronal specific promoters, guaranteed its flexibility with temporal and cell-population specificity. Carefully designed multispectral imaging setup and joint experimental pipeline with Expansion Microscopy ensured resolving the labeling information in spectral and spatial dimensions. It is foreseeable that Bitbow will strongly contribute to more and more advanced *Drosophila* neuronal subtype studies effectively. As the fluorescent proteins and genetic components used in Bitbow have been shown to be functional in mammalian systems (Livet et al., 2007; Cai et al., 2013), it is expected that the transgenic designs of Bitbow can be readily transferred in more complex organisms, such as zebrafish, mouse, etc.

Multi-modality study of specific neuron groups is key to accelerate our understanding of the entire nervous system. Joint investigations in neuronal morphology, lineage, transcriptome, etc. and efforts in connecting these characteristics onto the exact same cell, have been the critical but nontrivial goal for generations of neuroscientists (Bates, Janssens, Jefferis, & Aerts, 2019). In this dissertation study, *Drosophila* larval serotonergic neurons were profiled by Bitbow and scRNAseq to resolve their morphological, lineage and molecular identities. Although the results collectively suggest the presence of subtypes among the serotonergic neurons, further

investigation in other modalities, such as electrophysiological and neural functional properties are needed to strengthen the findings. More importantly, methods to directly connect these measurements to the exact same neurons are still needed to gain the most comprehensive view of their properties. Bitbow has the potential to bridging such investigations. It delivers a large number of unambiguous labeling in complex neuron populations, which effectively assign unique neuron identifications. Combining Bitbow imaging with *in situ* RNA/protein detection or single-cell sequencing after cell dissociations, it is possible to conduct and faithfully multi-modality measurement of the exact same neurons of interest. With the additional lineage and morphology (which indicating connectivity) information, we will be able to design more precise genetic interrogation that is development- and connection-specific. Starting from relatively simple subset of neurons as examples, this paradigm of multi-modality investigation is applicable to other invertebrate or vertebrate systems, and will eventually pave the way of deciphering functional relevance of neuronal subtypes in normal and disease conditions.

References

Ai, H.-W., Hazelwood, K. L., Davidson, M. W., & Campbell, R. E. (2008). Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors. *Nature Methods*, *5*(5), 401–403.

Ai, H.-W., Henderson, J. N., Remington, S. J., & Campbell, R. E. (2006). Directed evolution of a monomeric, bright and photostable version of Clavularia cyan fluorescent protein: structural characterization and applications in fluorescence imaging. *Biochemical Journal*, *400*(3), 531–540.

Alekseyenko, O. V., Lee, C., & Kravitz, E. A. (2010). Targeted manipulation of serotonergic neurotransmission affects the escalation of aggression in adult male Drosophila melanogaster. *PloS One*, *5*(5), e10806.

Bardin, A. J., Le Borgne, R., & Schweisguth, F. (2004). Asymmetric localization and function of cell-fate determinants: a fly's view. *Current Opinion in Neurobiology*, *14*(1), 6–14.

Bateman, J. R., Lee, A. M., & -ting Wu, C. (2006). Site-Specific Transformation of Drosophila via ϕC31 Integrase-Mediated Cassette Exchange. *Genetics*, *173*(2), 769–777.

Bates, A. S., Janssens, J., Jefferis, G. S., & Aerts, S. (2019). Neuronal cell types in the fly: single-cell anatomy meets single-cell genomics. *Current Opinion in Neurobiology*, *56*, 125–134.

Bayraktar, O. A., & Doe, C. Q. (2013). Combinatorial temporal patterning in progenitors expands neural diversity. *Nature*, *498*(7455), 449–455.

Becnel, J., Johnson, O., Luo, J., Nässel, D. R., & Nichols, C. D. (2011). The serotonin 5-HT7Dro receptor is expressed in the brain of Drosophila, and is essential for normal courtship and mating. *PloS One*, *6*(6), e20800.

Bhattacherjee, A., Djekidel, M. N., Chen, R., Chen, W., Tuesta, L. M., & Zhang, Y. (2019). Cell type-specific transcriptional programs in mouse prefrontal cortex during adolescence and addiction. *Nature Communications*, *10*(1), 4169.

Birkholz, O., Rickert, C., Berger, C., Urbach, R., & Technau, G. M. (2013). Neuroblast pattern and identity in the Drosophila tail region and role of doublesex in the survival of sex-specific precursors. *Development* , *140*(8), 1830–1842.

Bischof, J., Maeda, R. K., Hediger, M., Karch, F., & Basler, K. (2007). An optimized transgenesis system for Drosophila using germ-line-specific φC31 integrases. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(9), 3312–3317.

Blest, A. D. (1961). Some Modifications of Holmes's Silver Method for Insect Central Nervous Systems. *Journal of Cell Science*, *s3-102*(59), 413–417.

Bossing, T., & Technau, G. M. (1994). The fate of the CNS midline progenitors in Drosophila as revealed by a new method for single cell labelling. *Development* , *120*(7), 1895–1906.

Bossing, T., Udolph, G., Doe, C. Q., & Technau, G. M. (1996). The Embryonic Central Nervous System Lineages ofDrosophila melanogaster: I. Neuroblast Lineages Derived from the Ventral Half of the Neuroectoderm. *Developmental Biology*, *179*(1), 41–64.

Brody, T., & Odenwald, W. F. (2000). Programmed transformations in neuroblast gene expression during Drosophila CNS lineage development. *Developmental Biology*, *226*(1), 34–44.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*, *36*(5), 411–420.

Cahoon, C. K., Yu, Z., Wang, Y., Guo, F., Unruh, J. R., Slaughter, B. D., & Hawley, R. S. (2017). Superresolution expansion microscopy reveals the three-dimensional organization of the Drosophila synaptonemal complex. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(33), E6857–E6866.

Cai, D., Cohen, K. B., Luo, T., Lichtman, J. W., & Sanes, J. R. (2013). Improved tools for the Brainbow toolbox. *Nature Methods*, *10*(6), 540–547.

Cembrowski, M. S., & Spruston, N. (2019). Heterogeneity within classical cell types is the rule: lessons from hippocampal pyramidal neurons. *Nature Reviews. Neuroscience*, *20*(4), 193–204.

Chen, J., & Condron, B. G. (2008). Branch architecture of the fly larval abdominal serotonergic neurons. *Developmental Biology*, *320*(1), 30–38.

Cloonan, N., Forrest, A. R. R., Kolle, G., Gardiner, B. B. A., Faulkner, G. J., Brown, M. K., … Grimmond, S. M. (2008). Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nature Methods*, *5*(7), 613–619.

Croset, V., Treiber, C. D., & Waddell, S. (2018). Cellular diversity in the Drosophila midbrain revealed by single-cell transcriptomics. *eLife*, *7*. https://doi.org/10.7554/eLife.34550

Darmanis, S., Sloan, S. A., & Zhang, Y. (2015). A survey of human brain transcriptome diversity at the single cell level. *Proceedings of the*. Retrieved from https://www.pnas.org/content/112/23/7285.short

Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, Ł., … Aerts, S. (2018). A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. *Cell*, *174*(4), 982–998.e20.

Davis, M. W., Morton, J. J., Carroll, D., & Jorgensen, E. M. (2008). Gene activation using FLP recombinase in C. elegans. *PLoS Genetics*, *4*(3), e1000028.

Dymecki, S. M., & Kim, J. C. (2007). Molecular Neuroanatomy's "Three Gs": A Primer. *Neuron*, *54*(1), 17–34.

Evans, C. J., Olson, J. M., Ngo, K. T., Kim, E., Lee, N. E., Kuoy, E., … Banerjee, U. (2009). G-TRACE: rapid Gal4-based cell lineage analysis in Drosophila. *Nature Methods*, *6*(8), 603–605.

Fischbach, K.-F., -F. Fischbach, K., & Dittrich, A. P. M. (1989). The optic lobe of Drosophila

melanogaster. I. A Golgi analysis of wild-type structure. *Cell and Tissue Research*, Vol. 258. https://doi.org/10.1007/bf00218858

Fishilevich, E., & Vosshall, L. B. (2005). Genetic and functional subdivision of the Drosophila antennal lobe. *Current Biology: CB*, *15*(17), 1548–1553.

Frise, E., Knoblich, J. A., Younger-Shepherd, S., Jan, L. Y., & Jan, Y. N. (1996). The Drosophila Numb protein inhibits signaling of the Notch receptor during cell-cell interaction in sensory organ lineage. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(21), 11925–11932.

Fuzik, J., Zeisel, A., Máté, Z., Calvigioni, D., Yanagawa, Y., Szabó, G., … Harkany, T. (2016). Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. *Nature Biotechnology*, *34*(2), 175–183.

Gaspar, P., Cases, O., & Maroteaux, L. (2003). The developmental role of serotonin: news from mouse molecular genetics. *Nature Reviews. Neuroscience*, *4*(12), 1002–1012.

Gasque, G., Conway, S., Huang, J., Rao, Y., & Vosshall, L. B. (2013). Small molecule drug screening in Drosophila identifies the 5HT2A receptor as a feeding modulation target. *Scientific Reports*, *3*, srep02120.

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., 3rd, & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–345.

Glickstein, M. (2006). Golgi and Cajal: The neuron doctrine and the 100th anniversary of the 1906 Nobel Prize. *Current Biology: CB*, *16*(5), R147–R151.

Golgi, C. (1885). *Sulla fina anatomia degli organi centrali del sistema nervoso*. S. Calderini.

Greig, L. C., Woodworth, M. B., Galazo, M. J., Padmanabhan, H., & Macklis, J. D. (2013). Molecular logic of neocortical projection neuron specification, development and diversity. *Nature Reviews. Neuroscience*, *14*(11), 755–769.

Groth, A. C., Fish, M., Nusse, R., & Calos, M. P. (2004). Construction of Transgenic Drosophila by Using the Site-Specific Integrase From Phage φC31. *Genetics*, *166*(4), 1775–1782.

Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., … Zhang, Z. (2018). Publisher Correction: Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nature Medicine*, *24*(10), 1628.

Hadjieconomou, D., Rotkopf, S., Alexandre, C., Bell, D. M., Dickson, B. J., & Salecker, I. (2011). Flybow: genetic multicolor cell labeling for neural circuit analysis in Drosophila melanogaster. *Nature Methods*, *8*(3), 260–266.

Hampel, S., Chung, P., McKellar, C. E., Hall, D., Looger, L. L., & Simpson, J. H. (2011). Drosophila Brainbow: a recombinase-based fluorescence labeling technique to subdivide neural expression patterns. *Nature Methods*, *8*(3), 253–259.

Hanesch, U., Fischbach, K.-F., & Heisenberg, M. (1989). Neuronal architecture of the central complex in

Drosophila melanogaster. *Cell and Tissue Research*, *257*(2), 343–366.

Harris, R. M., Pfeiffer, B. D., Rubin, G. M., & Truman, J. W. (2015). Neuron hemilineages provide the functional ground plan for the Drosophila ventral nervous system. *eLife*, *4*. https://doi.org/10.7554/eLife.04493

Harzer, H., Berger, C., Conder, R., Schmauss, G., & Knoblich, J. A. (2013). FACS purification of Drosophila larval neuroblasts for next-generation sequencing. *Nature Protocols*, *8*(6), 1088–1099.

Hodgkin, A. L., & Huxley, A. F. (1952a). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, *117*(4), 500–544.

Hodgkin, A. L., & Huxley, A. F. (1952b). Currents carried by sodium and potassium ions through the membrane of the giant axon of Loligo. *The Journal of Physiology*, *116*(4), 449–472.

Hodgkin, A. L., & Huxley, A. F. (1952c). The components of membrane conductance in the giant axon of Loligo. *The Journal of Physiology*, *116*(4), 473–496.

Hodgkin, A. L., & Huxley, A. F. (1952d). The dual effect of membrane potential on sodium conductance in the giant axon of Loligo. *The Journal of Physiology*, *116*(4), 497–506.

Hodgkin, A. L., Huxley, A. F., & Katz, B. (1952). Measurement of current-voltage relations in the membrane of the giant axon of Loligo. *The Journal of Physiology*, *116*(4), 424–448.

Hofbauer, A., & Campos-Ortega, J. A. (1990). Proliferation pattern and early differentiation of the optic lobes in Drosophila melanogaster. *Roux's Archives of Developmental Biology: The Official Organ of the EDBO*, *198*(5), 264–274.

Holmes, W. (1943). Silver staining of nerve axons in paraffin sections. *The Anatomical Record*, Vol. 86, pp. 157–187. https://doi.org/10.1002/ar.1090860205

Homem, C. C. F., & Knoblich, J. A. (2012). Drosophila neuroblasts: a model for stem cell biology. *Development* , *139*(23), 4297–4310.

Honig, M. G., & Hume, R. I. (1986). Fluorescent carbocyanine dyes allow living neurons of identified origin to be studied in long-term cultures. *The Journal of Cell Biology*, *103*(1), 171–187.

Honig, M. G., & Hume, R. I. (1989). Dil and diO: versatile fluorescent dyes for neuronal labelling and pathway tracing. *Trends in Neurosciences*, *12*(9), 333–335, 340–341.

Huser, A., Rohwedder, A., Apostolopoulou, A. A., Widmann, A., Pfitzenmaier, J. E., Maiolo, E. M., … Thum, A. S. (2012). The serotonergic central nervous system of the Drosophila larva: anatomy and behavioral function. *PloS One*, *7*(10), e47518.

Isshiki, T., Pearson, B., Holbrook, S., & Doe, C. Q. (2001). Drosophila neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. *Cell*, *106*(4), 511–521.

Ito, K., Awano, W., Suzuki, K., Hiromi, Y., & Yamamoto, D. (1997). The Drosophila mushroom body is a quadruple structure of clonal units each of which contains a virtually identical set of neurones and glial cells. *Development* , *124*(4), 761–771.

Ito, K., & Hotta, Y. (1992). Proliferation pattern of postembryonic neuroblasts in the brain of Drosophila melanogaster. *Developmental Biology*, *149*(1), 134–148.

Ito, K., Shinomiya, K., Ito, M., Armstrong, J. D., Boyan, G., Hartenstein, V., … Insect Brain Name Working Group. (2014). A systematic nomenclature for the insect brain. *Neuron*, *81*(4), 755–765.

Jenett, A., Rubin, G. M., Ngo, T.-T. B., Shepherd, D., Murphy, C., Dionne, H., … Zugates, C. T. (2012). A GAL4-driver line resource for Drosophila neurobiology. *Cell Reports*, *2*(4), 991–1001.

Johnson, O., Becnel, J., & Nichols, C. D. (2011). Serotonin receptor activity is necessary for olfactory learning and memory in Drosophila melanogaster. *Neuroscience*, *192*, 372–381.

Kanca, O., Caussinus, E., Denes, A. S., Percival-Smith, A., & Affolter, M. (2014). Raeppli: a whole-tissue labeling tool for live imaging of Drosophila development. *Development* , *141*(2), 472–480.

Kasture, A. S., Hummel, T., Sucic, S., & Freissmuth, M. (2018). Big Lessons from Tiny Flies: Drosophila melanogaster as a Model to Explore Dysfunction of Dopaminergic and Serotonergic Neurotransmitter Systems. *International Journal of Molecular Sciences*, *19*(6). https://doi.org/10.3390/ijms19061788

Kepecs, A., & Fishell, G. (2014). Interneuron cell types are fit to function. *Nature*, *505*(7483), 318–326.

Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., … Kirschner, M. W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*, *161*(5), 1187–1201.

Knoblich, U., Huang, L., Zeng, H., & Li, L. (2019). Neuronal cell-subtype specificity of neural synchronization in mouse primary visual cortex. *Nature Communications*, *10*(1), 2533.

Kretzschmar, K., & Watt, F. M. (2012). Lineage tracing. *Cell*, *148*(1-2), 33–45.

Lacar, B., Linker, S. B., Jaeger, B. N., Krishnaswami, S. R., Barron, J. J., Kelder, M. J. E., … Gage, F. H. (2016). Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nature Communications*, *7*, 11022.

Lacin, H., Chen, H.-M., Long, X., Singer, R. H., Lee, T., & Truman, J. W. (2019). Neurotransmitter identity is acquired in a lineage-restricted manner in the CNS. *eLife*, *8*. https://doi.org/10.7554/eLife.43701

Lai, S.-L., Awasaki, T., Ito, K., & Lee, T. (2008). Clonal analysis of Drosophila antennal lobe neurons: diverse neuronal architectures in the lateral neuroblast lineage. *Development* , *135*(17), 2883–2893.

Larsen, C., Shy, D., Spindler, S. R., Fung, S., Pereanu, W., Younossi-Hartenstein, A., & Hartenstein, V. (2009). Patterns of growth, axonal extension and axonal arborization of neuronal lineages in the developing Drosophila brain. *Developmental Biology*, *335*(2), 289–304.

Laurens van der Maaten, G. H. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research: JMLR*, *9*, 2579–2605.

Lee, T., & Luo, L. (1999). Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron*, *22*(3), 451–461.

Li, H., Horns, F., Wu, B., Xie, Q., Li, J., Li, T., … Luo, L. (2017). Classifying Drosophila Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing. *Cell*, *171*(5), 1206–1220.e22.

Lillesaar, C., Tannhäuser, B., Stigloher, C., Kremmer, E., & Bally-Cuif, L. (2007). The serotonergic phenotype is acquired by converging genetic mechanisms within the zebrafish central nervous system. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *236*(4), 1072–1084.

Liu, Y., Jiang, Y. 'ai, Si, Y., Kim, J.-Y., Chen, Z.-F., & Rao, Y. (2011). Molecular regulation of sexual preference revealed by genetic studies of 5-HT in the brains of male mice. *Nature*, *472*(7341), 95–99.

Livet, J., Weissman, T. A., Kang, H., Draft, R. W., Lu, J., Bennis, R. A., … Lichtman, J. W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature*, *450*(7166), 56–62.

Li, X., Erclik, T., Bertet, C., Chen, Z., Voutev, R., Venkatesh, S., … Desplan, C. (2013). Temporal patterning of Drosophila medulla neuroblasts controls neural fates. *Nature*, *498*(7455), 456–462.

Luan, H., Peabody, N. C., Vinson, C. R., & White, B. H. (2006). Refined spatial manipulation of neuronal function by combinatorial restriction of transgene expression. *Neuron*, *52*(3), 425–436.

Lundell, M. J., & Hirsh, J. (1994). Temporal and spatial development of serotonin and dopamine neurons in the Drosophila CNS. *Developmental Biology*, *165*(2), 385–396.

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., … McCarroll, S. A. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, *161*(5), 1202–1214.

Markstein, M., Pitsouli, C., Villalta, C., Celniker, S. E., & Perrimon, N. (2008). Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nature Genetics*, *40*(4), 476–483.

McInnes, L., Healy, J., Saul, N., & Großberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, Vol. 3, p. 861. https://doi.org/10.21105/joss.00861

McLeod, M., Craft, S., & Broach, J. R. (1986). Identification of the crossover site during FLP-mediated recombination in the Saccharomyces cerevisiae plasmid 2 microns circle. *Molecular and Cellular Biology*, *6*(10), 3357–3367.

Molyneaux, B. J., Arlotta, P., Menezes, J. R. L., & Macklis, J. D. (2007). Neuronal subtype specification in the cerebral cortex. *Nature Reviews. Neuroscience*, *8*(6), 427–437.

Muzumdar, M. D., Luo, L., & Zong, H. (2007). Modeling sporadic loss of heterozygosity in mice by using mosaic analysis with double markers (MADM). *Proceedings of the National Academy of Sciences of the United States of America*, *104*(11), 4495–4500.

Nern, A., Pfeiffer, B. D., & Rubin, G. M. (2015). Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(22), E2967–E2976.

Niens, J., Reh, F., Çoban, B., Cichewicz, K., Eckardt, J., Liu, Y.-T., … Riemensperger, T. D. (2017). Dopamine Modulates Serotonin Innervation in the Brain. *Frontiers in Systems Neuroscience*, *11*, 76.

Ohyama, T., Schneider-Mizell, C. M., Fetter, R. D., Aleman, J. V., Franconville, R., Rivera-Alba, M., … Zlatic, M. (2015). A multilevel multimodal circuit enhances action selection in Drosophila. *Nature*, *520*(7549), 633–639.

Pandey, S., Shekhar, K., Regev, A., & Schier, A. F. (2018). Comprehensive Identification and Spatial Mapping of Habenular Neuronal Types Using Single-Cell RNA-Seq. *Current Biology: CB*, *28*(7), 1052–1065.e7.

Pascual, A., & Préat, T. (2001). Localization of long-term memory within the Drosophila mushroom body. *Science*, *294*(5544), 1115–1117.

Pearson, B. J., & Doe, C. Q. (2003). Regulation of neuroblast competence in Drosophila. *Nature*, *425*(6958), 624–628.

Peng, H., Ruan, Z., Long, F., Simpson, J. H., & Myers, E. W. (2010). V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets. *Nature Biotechnology*, *28*(4), 348–353.

Pfeiffer, B. D., Ngo, T.-T. B., Hibbard, K. L., Murphy, C., Jenett, A., Truman, J. W., & Rubin, G. M. (2010). Refinement of tools for targeted gene expression in Drosophila. *Genetics*, *186*(2), 735–755.

Pfeiffer, B. D., Truman, J. W., & Rubin, G. M. (2012). Using translational enhancers to increase transgene expression in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(17), 6626–6631.

Pollen, A. A., Nowakowski, T. J., Shuga, J., Wang, X., Leyrat, A. A., Lui, J. H., … West, J. A. A. (2014). Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nature Biotechnology*, *32*(10), 1053–1058.

Raj, B., Wagner, D. E., McKenna, A., Pandey, S., Klein, A. M., Shendure, J., … Schier, A. F. (2018). Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nature Biotechnology*, *36*(5), 442–450.

Ramon y Cajal, S. (1911). Histologie du système nerveux de l'homme et des vertébrés. *Maloine, Paris*, *2*, 153–173.

Richier, B., & Salecker, I. (2015). Versatile genetic paintbrushes: Brainbow technologies. *Wiley Interdisciplinary Reviews. Developmental Biology*, *4*(2), 161–180.

Roossien, D. H., Sadis, B. V., Yan, Y., Webb, J. M., Min, L. Y., Dizaji, A. S., … Cai, D. (2019). Multispectral tracing in densely labeled mouse brain with nTracer. *Bioinformatics* , *35*(18), 3544–3546.

Sakaue-Sawano, A., Kurokawa, H., Morimura, T., Hanyu, A., Hama, H., Osawa, H., … Miyawaki, A. (2008). Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell*, *132*(3), 487–498.

Schlake, T., & Bode, J. (1994). Use of mutated FLP recognition target (FRT) sites for the exchange of

expression cassettes at defined chromosomal loci. *Biochemistry*, *33*(43), 12746–12751.

Shaner, N. C., Lambert, G. G., Chammas, A., Ni, Y., Cranfill, P. J., Baird, M. A., … Wang, J. (2013). A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. *Nature Methods*, *10*(5), 407–409.

Shcherbo, D., Murphy, C. S., Ermakova, G. V., Solovieva, E. A., Chepurnykh, T. V., Shcheglov, A. S., … Chudakov, D. M. (2009). Far-red fluorescent tags for protein imaging in living tissues. *Biochemical Journal*, *418*(3), 567–574.

Sitaraman, D., Zars, M., Laferriere, H., Chen, Y.-C., Sable-Smith, A., Kitamoto, T., … Zars, T. (2008). Serotonin is necessary for place memory in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(14), 5579–5584.

Skeath, J. B., & Doe, C. Q. (1998). Sanpodo and Notch act in opposition to Numb to distinguish sibling neuron fates in the Drosophila CNS. *Development* , *125*(10), 1857–1865.

Sodhi, M. S. K., & Sanders-Bush, E. (2004). Serotonin and brain development. *International Review of Neurobiology*, *59*, 111–174.

Spana, E. P., Kopczynski, C., Goodman, C. S., & Doe, C. Q. (1995). Asymmetric localization of numb autonomously determines sibling neuron identity in the Drosophila CNS. *Development* , *121*(11), 3489–3494.

Strauss, R. (2002). The central complex and the genetic dissection of locomotor behaviour. *Current Opinion in Neurobiology*, *12*(6), 633–638.

Struhl, G., & Adachi, A. (1998). Nuclear access and action of notch in vivo. *Cell*, *93*(4), 649–660.

Sultan, K. T., & Shi, S.-H. (2018). Generation of diverse cortical inhibitory interneurons. *Wiley Interdisciplinary Reviews. Developmental Biology*, *7*(2). https://doi.org/10.1002/wdev.306

Suzuki, T., Kaido, M., Takayama, R., & Sato, M. (2013). A temporal mechanism that produces neuronal diversity in the Drosophila visual center. *Developmental Biology*, *380*(1), 12–24.

Tang, F., Barbacioru, C., Bao, S., Lee, C., Nordman, E., Wang, X., … Surani, M. A. (2010). Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. *Cell Stem Cell*, *6*(5), 468–478.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., … Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods*, *6*(5), 377–382.

Taniguchi, H., He, M., Wu, P., Kim, S., Paik, R., Sugino, K., … Huang, Z. J. (2011). A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron*, *71*(6), 995–1013.

Tan, X., & Shi, S.-H. (2013). Neocortical neurogenesis and neuronal migration. *Wiley Interdisciplinary Reviews. Developmental Biology*, *2*(4), 443–459.

Tillberg, P. W., Chen, F., Piatkevich, K. D., Zhao, Y., Yu, C.-C. J., English, B. P., … Boyden, E. S. (2016). Protein-retention expansion microscopy of cells and tissues labeled using standard

fluorescent proteins and antibodies. *Nature Biotechnology*, *34*(9), 987–992.

Truman, J. W., Moats, W., Altman, J., Marin, E. C., & Williams, D. W. (2010). Role of Notch signaling in establishing the hemilineages of secondary neurons in Drosophila melanogaster. *Development* , *137*(1), 53–61.

Turan, S., Kuehle, J., Schambach, A., Baum, C., & Bode, J. (2010). Multiplexing RMCE: versatile extensions of the Flp-recombinase-mediated cassette-exchange technology. *Journal of Molecular Biology*, *402*(1), 52–69.

Urbach, R., & Technau, G. M. (2003). Segment polarity and DV patterning gene expression reveals segmental organization of the Drosophila brain. *Development* , *130*(16), 3607–3620.

Urbach, R., Volland, D., Seibert, J., & Technau, G. M. (2006). Segment-specific requirements for dorsoventral patterning genes during early brain development in Drosophila. *Development* , *133*(21), 4315–4330.

Vallés, A. M., & White, K. (1988). Serotonin-containing neurons in Drosophila melanogaster: development and distribution. *The Journal of Comparative Neurology*, *268*(3), 414–428.

Venken, K. J. T., He, Y., Hoskins, R. A., & Bellen, H. J. (2006). P [acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. *Science*, *314*(5806), 1747–1751.

Volkert, F. C., & Broach, J. R. (1986). Site-specific recombination promotes plasmid amplification in yeast. *Cell*, *46*(4), 541–550.

Waltman, L., & van Eck, N. J. (2013). A smart local moving algorithm for large-scale modularity-based community detection. *The European Physical Journal B*, Vol. 86. https://doi.org/10.1140/epjb/e2013-40829-0

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews. Genetics*, *10*(1), 57–63.

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., … Tang, F. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nature Structural & Molecular Biology*, *20*, 1131.

y Cajal, S. R. (1888). *Estructura del cerebelo*.

y Cajal, S. R. (1910). Las fórmulas del proceder del nitrato de plata reducido y sus efectos sobre los factores integrantes de las neuronas. *Trab. Lab. Invest. Biol. Univ. Madr.*, *8*, 1–26.

Ye, B., Zhang, Y., Song, W., Younger, S. H., Jan, L. Y., & Jan, Y. N. (2007). Growing dendrites and axons differ in their reliance on the secretory pathway. *Cell*, *130*(4), 717–729.

Yuan, Q., Joiner, W. J., & Sehgal, A. (2006). A sleep-promoting role for the Drosophila serotonin receptor 1A. *Current Biology: CB*, *16*(11), 1051–1062.

Yu, H.-H., Chen, C.-H., Shi, L., Huang, Y., & Lee, T. (2009). Twin-spot MARCM to reveal the developmental origin and identity of neurons. *Nature Neuroscience*, *12*(7), 947–953.

Zeisel, A., Muñoz-Manchado, A. B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., …

Linnarsson, S. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science*, *347*(6226), 1138–1142.

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*, *8*, 14049.

Zhou, F., Wang, R., Yuan, P., Ren, Y., Mao, Y., Li, R., … Tang, F. (2019). Reconstituting the transcriptome and DNA methylome landscapes of human implantation. *Nature*, *572*(7771), 660–664.