Dissertation
submitted to the
Combined Faculty of Natural Sciences and Mathematics
of the Ruperto Carola University Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by
Jessica Velten, M.Sc.
Born in: Darmstadt, Germany
Oral examination: 14-04-2021

Molecular Determinants of Synaptic Specificity at the Single Cell Level

Referees:

Professor Dr. Ingrid Lohmann

Dr. Aurelio Teleman

SUMMARY

The correct wiring of neuronal circuits represents one of the most complex processes throughout development. Mistakes in their formation underly genetic neuropathies including autism and schizophrenia, but the molecular program encoded in the genome allowing each axon to precisely innervate its target cell remains poorly understood.

In this thesis I use the *Drosophila* motoneuronal system as a tractable model for studying the formation of complex neuronal circuits during development. I profiled terminally specified motoneurons (MNs) of *Drosophila* embryos in multiple biological replicates by single cell RNA-Seqencing (scRNA-Seq), mapped their spatial position within the embryo through bioinformatic approaches and high-resolution imaging, and investigated cell-specific changes in synaptic wiring induced by genetic manipulations.

I demonstrate that the combinatorial expression of specific homeodomain transcription factors (TFs) serves as a molecular coordinate system to specify the identity and position of motoneuron identities along the ventral nerve cord (VNC). This TF factor code is linked to the expression of cell-specific combinations of Immunoglobulin (Ig) domain proteins that functionally determine synaptic specificity and synaptic partner choice by cell adhesion. My data also shows that a similar mechanism acts in muscle cells, suggesting that differential cell affinities encoded by Ig proteins downstream of homeodomain TFs is a key feature of selective synaptic partner choice. This suggests that a homeo-Ig code is translated into complex neuronal wiring schemes required for establishing the structure of neuronal circuits.

Together this thesis gives insights into the molecular logic of synaptic wiring down to the single cell level. In particular, the molecular determinates defining synaptic partners and instructing cells to form cell specific synaptic connections are identified. The experimental advances in this thesis enable a systematic view on spatial organization, molecular identities and connectivity of neuronal circuits during the development of an organism.

ZUSAMMENFASSUNG

Die korrekte Verschaltung von neuronalen Schaltkreisen stellt einen der komplexesten Prozesse während der Entwicklung dar. Fehler in der Bildung dieser Schaltkreise können zu entwicklungsbedingten Nervenkrankheiten wie Autismus oder Schizophrenie führen. Dennoch sind die grundlegenden molekularen Programme nicht ausreichend verstanden, die jedes Axon zur spezifischen Innervierung des korrekten Partnerneurons führen.

In dieser Doktorarbeit habe ich das motoneuronale System von der Fruchtfliege *Drosophila* als Modelsystem verwendet, um die Bildung von komplexen neuronalen Schaltkreisen während der Entwicklung zu untersuchen. Dafür habe ich die Genexpression einzelner terminal differenzierte Motoneuronen aus dem *Drosophila* Embryo in mehreren biologischen Replikaten mit Einzelzellsequenzierung gemessen. Die Motoneuronen wurden nach ihrer räumlichen Position im Embryo mithilfe von bioinformatischen Methoden und hochauflösender Bildbearbeitung angeordnet. Schließlich wurden vorhergesagte zellspezifische Veränderungen in der synaptischen Verschaltung durch gezielte genetische Manipulationen hervorgerufen und somit validiert. Dabei konnte ich zeigen, dass ein kombinatorischer Code von bestimmten Homeodomain-Trankriptionsfaktoren als molekulares Koordinatensystem dient, um die Identität und Position einzelner Motoneuron-Identitäten entlang des ventralen Nervenstrangs festzulegen. Dieser Transkriptionsfaktor Code ist verknüpft mit der zellspezifischen Expression von Kombinationen an Immunoglobulin-Domain Proteinen, welche durch ihre Zelladhesions-Eigenschaften Axone zur Innervierung spezifischer Zielzellen anleiten und damit die synaptische Partnerwahl beeinflussen können. Meine Daten zeigen, dass ein ähnlicher Mechanismus in synaptischen Partnerzellen, dem Muskel vorhanden ist, daher liegt es nahe, dass beide synaptischen Partnern durch übereinstimmende Zellaffinitäten der Immunoglobulin Moleküle eine Verbindung eingehen und dass diese Affinitäten durch einen positionsabhängigen Homeodomain-Trandkriptionsfaktor Code reguliert werden. Diese Ergebnisse legen ein Model nahe, in dem ein zellspezifischer molekularer Homeo-Immunoglobulin Code die komplexe Verschaltung von neuronalen Netwerken definiert und reguliert. Zusammenfassend gibt diese Doktorarbeit Einblicke in die molekulare Logik von neuronalen Schaltkreisen bis herunter auf die Einzelzellebene. Insbesondere wurden molekulare Bestimmungsfaktoren identifiziert, welche einer Zelle ermöglichen spezifische synaptisch Kontakte herzustellen. Die experimentellen Fortschritte dieser Doktorarbeit ermöglichen einen systematischen Überblick über die räumliche Organisation, molekulare Identitäten und Konnektivität von neuronalen Schaltkreisen während der Entwicklung eines Organismus.

ACKNOWLEDGEMENTS

I would especially like to thank Ingrid Lohmann and the entire Lohmann Lab for the support, guidance and encouragement that made my PhD project possible.

Thanks to Ingrid Lohmann for providing me with such an exciting research question and supporting me with inspiring scientific discussions, guidance and also the freedom and trust in my own ideas. I am also very grateful to my husband Lars Velten, who collaborated on my project with his great expertise in scRNA sequencing and bioinformatics and believed in me.

I want to thank my students Rashi Agarwal and Lena Bognar for their contributions on experiments and great discussions. I wish you all the best in the future.

Thanks to the people at the EMBL genomic, proteomic and FACS facility that provided me with technology, knowledge and reagents to complete my PhD project. Especially, thanks to Malte Paulsen who's advanced FACS skills solved the technical limitations that my project encountered for almost a year. Great thanks to the people who helped and inspired me with protocol development, experiments, reagents or great discussions on the project (Patrick van Nierop, Tobias Schunke, Sebastian Sorge, Jana Friedrich, Katrin Domsch, Jonas Theelke, Kerem Yildirim and Petra Kaspar). Many thanks to my TAC committee members, Aurelio Teleman and Steffen Lemke who followed my project over the past years and helped me with great advises, input on the publication and encouraged me especially in the final phase of my project a lot. Thanks to Nick Foulkes who kindly agreed to be my fourth examiner in my thesis defense.

Finally, I want to thank my family and friends for their constant support on this journey. Especially, I want to thank my brother, my parents and my grandparents for their support and encouragement to follow my passion in science. Also, thanks to my friend's insight and outside the institute the "sandwich club", especially Linda Viol for all the personal support during PhD times and all the other friends for the joy, laughs and support.

INHALTSVERZEICHNIS

SUMMARY 3

ZUSAMMENFASSUNG 4

ACKNOWLEDGEMENTS 5

1. INTRODUCTION 8

1.1 DEVELOPMENT OF THE NEURONAL SYSTEM 8
1.1.1 THE DEVELOPMENTAL ORIGIN OF NEURONS 9
1.1.2 DEVELOPMENT OF THE CNS 11
1.1.3 HOMEODOMAIN TFs IN NERONAL DEVELOPMENT 13
1.2 SYNAPTIC SPECIFICITY AND SYNAPTIC PARTNER CHOICE 16
1.3 THE ANATOMY OF THE NERVOUS SYSTEM FOR LOCOMOTION 20
1.3.1 DEVELOPMENT OF THE MOTONEURONAL SYSTEM 22
1.4 SINGLE CELL TRANSCRIPTOMICS 24
1.4.1 SPATIAL TRANSCRIPTOMICS 25
1.5 THE MOLECULAR TOOLBOX OF FLIES 26
1.6 AIMS 28

2. RESULTS 29

2.1 A REFERENCE MAP OF MNs DURING THE SYNAPTIC WIRING PHASE 29
2.2 A HOMEO-CODE SPECIFIES SINGLE DROSOPHILA MNs DURING SYNAPTIC WIRING 35
2.3 REGIONAL SPECIFIC EXPRESSION OF THE HOMEO-CODE 40
2.4 HOMEO-TFs MODULATE SYNAPTIC TARGET SPECIFICITY IN A POSITION DEPENDENT MANNER 46
2.5 HOMEODOMAIN EFFECTS ON TARGET SPECIFICITY ARE MEDIATED BY COMBINATORIAL IG EXPRESSION 50
2.6 HOMEO-TFs MEDIATE TARGET SPECIFICITY IN BOTH MATCHING PARTNERS 55
2.7 CIRCUIT SPECIFIC ASSEMBLY AND CONNECTIVITY MECHANISMS MEDIATED BY HOMEODO-TFs 62

3. DISCUSSION 63

3.1 A SCRNA EXPERIMENT DESIGNED TO IDENTIFY MOLECULAR AND CELLULAR PROGRAMS OF INDIVIDUAL NEURONAL CELLS 63
3.2 SPATIALLY ORGANIZED HOMEO-TFS ACT IN INDIVIDUAL NEURON IDENTITIES 63
3.3 A HOMEO-CODE ACTS IN INDIVIDUAL NEURON IDENTITIES AND FINE TUNES SYNAPTIC SPECIFICITY 64
3.4 IG PROTEIN EXPRESSION IS A MEDIATOR OF SYNAPTIC SPECIFICITY 65
3.5 A MODEL FOR SYNAPTIC WIRING IN THE NEUROMUSCULAR SYSTEM 66
3.6 HOMEODO-TF EXPRESSION AND NEURONAL CIRCUIT DEVELOPMENT DURING EVOLUTION 67

4. OUTLOOK 69

5. MATERIAL AND METHODS 72

5.1 DROSOPHILA STRAINS AND EXPERIMENTAL CROSSES 72
5.2 PLASMID CONSTRUCTION AND TRANSGENESIS 73
5.3 IMMUNOHISTOCHEMISTRY 73
5.4 MICROSCOPY AND IMAGE ANALYSIS 73
5.5 DISSICIATION OF EMBRYONIC CELLS AND FLOW CYTOMETRY 74
5.6 SCRNA SEQUENCING 74
5.7 RAW DATA ROCESSING; QUALITY CONTROL AND NORMALIZATION 75
5.8 CLUSTERING AND DIMENSIONALITY REDUCTION 75
5.9 INFERENCE OF SPATIAL POSITION FROM SINGLE GENE EXPRESSION DATA 75
5.10 ZINB-WAVE ANALYSIS 76
5.11 CHIP-SEQ REANALYSIS AND IREGULON ANALYSIS 76
5.12 DATA VISUALIZATION 77
5.13 CODE AVAILABILITY 77
5.14 DATA AVAILIBILITY 77

6. RESOURCES 78

7. SUPPLEMENTARY 82

8. ABBRIVIATIONS 132

9. REFERENCES 134

1. INTRODUCTION

Each neuron of the nervous system can be connected to up to 1000 synaptic partners. In total, a human brain contains 100 trillion synaptic connections (Drachman, 2005), a thousand times more than there are stars in the universe. Genetic circuits in *Drosophila*, as well as the mammalian central nervous system (CNS), are stereotypic, thus their structure is genetically specified. Developmental programs instruct individual axons to find their appropriate synaptic partner with extraordinary precision. Precise establishment of these circuits is a prerequisite for fundamental behaviours such as walking, breathing and feeding. However, how such complex structures are precisely formed during development and in particular, how the required processes are encoded in the genome, remains poorly understood.

The basic structures of the bilaterian nervous system remains similar from *Drosophila* to humans. I will in the following introduce the development of the *Drosophila* neuronal system, and specifically, the motoneuronal system, which serves as a tractable model system for studying the establishment and specificity mechanisms of neuronal circuit formation throughout my thesis. Then, I will introduce the tools of single-cell genomics and fly genetics, which enable the work carried out in this thesis. At the end of the chapter, I will briefly describe important similarities and differences in the development of mammalian neuronal circuits.

1.1 DEVELOPMENT OF THE NEURONAL SYSTEM

To understand the specification of neuronal circuits, it is important to obtain an overview of the developmental origin of neurons. In *Drosophila*, two distinct nervous systems are built at two different developmental stages, in the larvae and in mature flies. The nervous system in adult flies originates from neurogenesis in post-embryonic stages (pupae), when animals undergo metamorphosis, a process that restructures the entire anatomy of the larvae into a mature animal (Stratoulias & Heino, 2015). In the following, I focus on the wiring phase of the nervous system in late embryos, where the first set of stereotyped connections for locomotion are established that serves as a foundation for the larval nervous system (Figure 1).

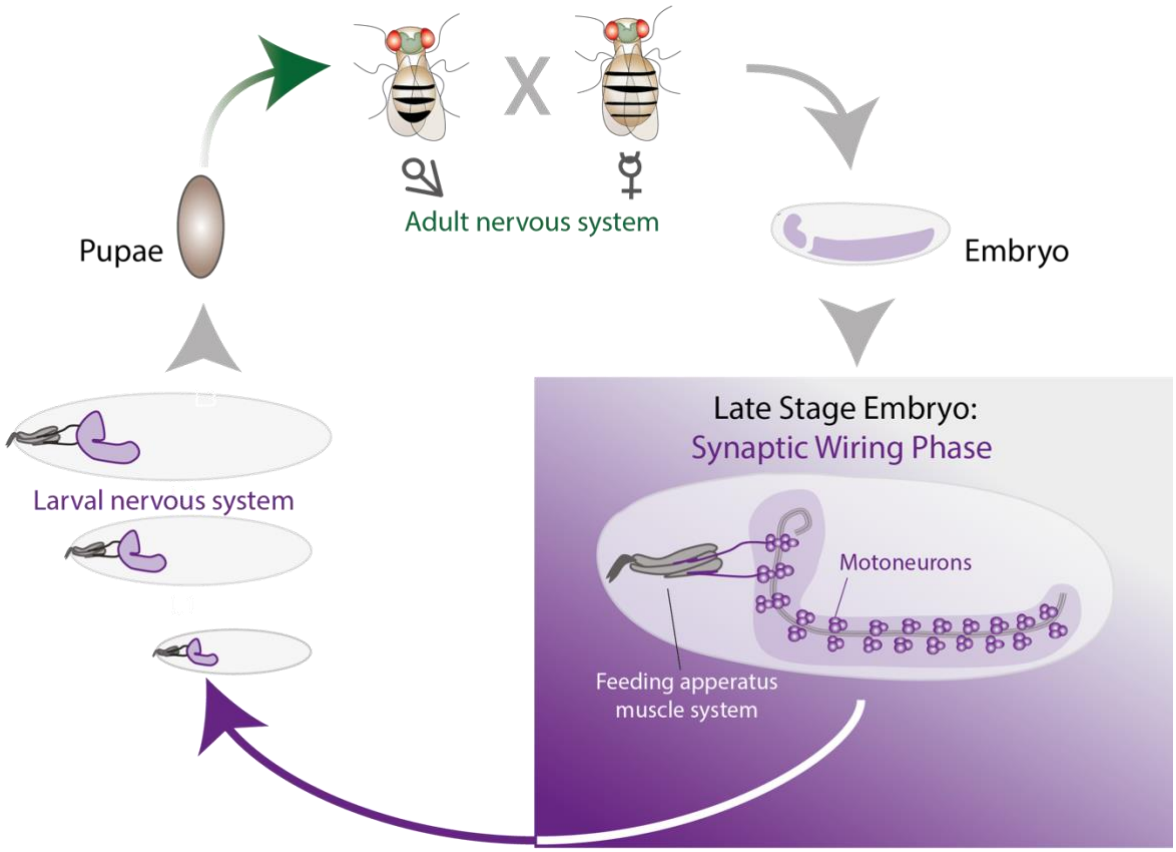


Figure 1: *Drosophila* life cycle and development of two distinct neuronal systems

After fertilization a *Drosophila* embryo takes one day to develop into an early stage 1 Larvae. In the late embryo the first round of stereotyped synaptic wiring of the prospective larval nervous system takes place, a requirement for hatching movements and further locomotions of the larvae. The Larvae extensively eats to increase in body size. Throughout larval growth the number of synaptic partners increases, thereby the Larval nervous system fully develops. When the Larvae enters Pupae stages, metamorphosis occurs, a process that includes entire restructuring of the larvae including the nervous system. A second mature nervous system develops until a mature *Drosophila* animal encloses from the Pupae. Another round of mating can take place after about 1 day of maturation.

1.1.1 THE DEVELOPMENTAL ORIGIN OF NEURONS

During the first steps of *Drosophila* development, the embryo is a syncytium, i.e. a single, polynucleated cell. At this stage, gradients of cross-repressive TFs such as maternal effector genes, segmentation genes (Riechmann & Ephrussi, 2001) sequentially expressed along the anterior to posterior (AP) axis and pattern the embryo into so-called parasegments by the expression of specific *Hox* genes (Nüsslein-Volhard & Wieschaus, 1980; Akam, 1987; Harding et al., 1985; Lewis, 1978). Parasegments are genetic units of the embryo that contain the posterior part of one segment and the anterior part of the next segment (Martinez-Arias & Lawrence, 1985; Perrimon, 1994). In parallel, opposing gradients of maternal effector genes and Toll signaling gradients along the dorsal-ventral (DV) axis pattern the embryo into zones.

While these processes are fundamental for the development of the entire embryo, they are crucial also for the specification of the nervous system, as I will describe in more detail below.

The first precursors of neuronal cells appear as soon as the embryo cellularises at blastula stage (embryonic stage 5). At this stage, neurogenic zones that later give rise to the neuronal system start to develop. During the following steps of development (gastrulation, stage 6 to 8), three distinct germ layers are specified: An inner germ layer termed endoderm that later forms inner organs; an intermediate layer termed mesoderm that develops into the muscle system; and an outer layer, termed ectoderm, that gives rise to the epidermis and the nervous system. In the ectoderm, the ventral part develops into the neuroectoderm that later gives rise to neuronal progenitor cells. These so-called neuroblasts (NBs) form the VNC, a part of the CNS (Figure 2). However, not every cell in the neuroectoderm becomes a neuronal cell, as some cells also acquire an epidermal fate.

This crucial fate decision is made through a process termed lateral inhibition, which ensures that neuronal and epidermal fates form an alternating pattern. In this process, a higher concentration of proneuronal genes is expressed in prospective NBs. The activity of these proneuronal genes, the *achaete-scute* (*As-c*) complex (J B Skeath & Carroll, 1992; James B Skeath & Carroll, 1994), promotes both the specification of cells into the neuronal fate and the expression of Delta ligand (Seugnet et al 1997). Delta binds the Notch receptor expressed on neighboring cells, where Delta-Notch signaling leads to inhibition of the neuronal fate by inhibiting proneuronal genes. Thereby, Delta-Notch signaling induces a mosaic pattern of neuronal and epidermal cells. Once specified, NBs delaminate (disconnect) from the neuroectoderm (Wheeler, 1893) and move inside the interior of the embryo. The remaining cells stay in the epithelia.

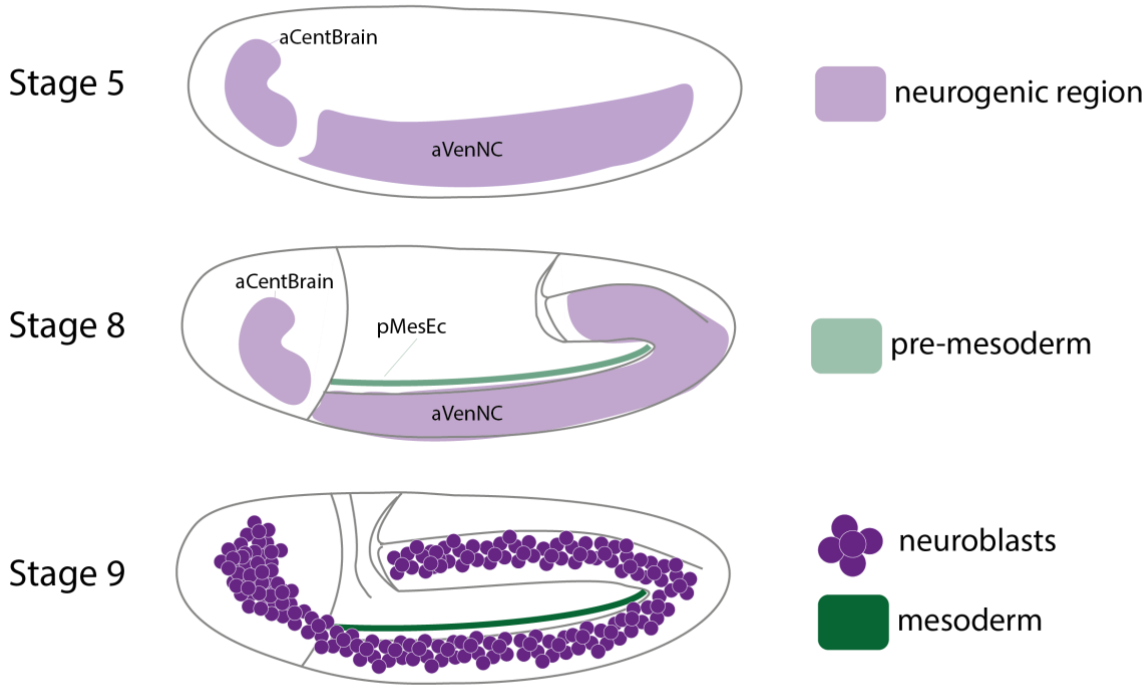


Figure 2: Development of NBs in early *Drosophila* embryos

During blastula stages neurogenic regions are primed to give rise to NBs in later developmental stages. These regions separate into a ventral neurogenic region that promotes for the prospective VNC and an anterior located neurogenic region that specifies NBs of the brain. Throughout development the pre-mesoderm divides from the ventral neurogenic region until it becomes the mesoderm in stage 9 that gives rise to the muscle system. In parallel NBs and epidermal precursor cells are specified from neurogenic regions. (Elements of the Illustration are adopted from "Atlas of *Drosophila* Development" of Volker Hartenstein)

1.1.2 DEVELOPMENT OF THE CNS

During gastrulation, the ventrally located neuroectoderm is divided into two regions; the anterior procephalic neuroectoderm that gives rise to the larval brain and the posterior, ventral neuroectoderm. The latter structure gives rise to both the subesophageal ganglion (SOG), a part of the CNS associated with feeding, and the VNC (Campos-Ortega & Hartenstein, 1997; Hartenstein & Wodarz, 2013). Importantly, cells belonging to different parasegments and dorsal-ventral "zones" contribute to the VNC and SOG. For example, the VNC, spans parasegments defined by polarization genes such as *wingless* (*wg*), *gooseberry* (*gsb*) or *engrailed* (*en*), homeodomain TFs and Hox TFs. At the same time, the VNC spans several dorsal-ventral zones defined by several columnar homeodomain TFs (*muscle segment homeobox (msh*), *intermediate NB defective* (*ind*), *ventral nervous system defective* (*vnd)*) (von Ohlen & Doe, 2000; Weiss et al., 1998). Depending on its position along these two axes, each neuronal progenitor cell within the VNC hence differs in a Cartesian grid like expression of

these factors (Allan & Thor, 2015a). The exact position of NBs can be identified by its position and the associated genetic markers (Doe, 1992) (Figure 3).

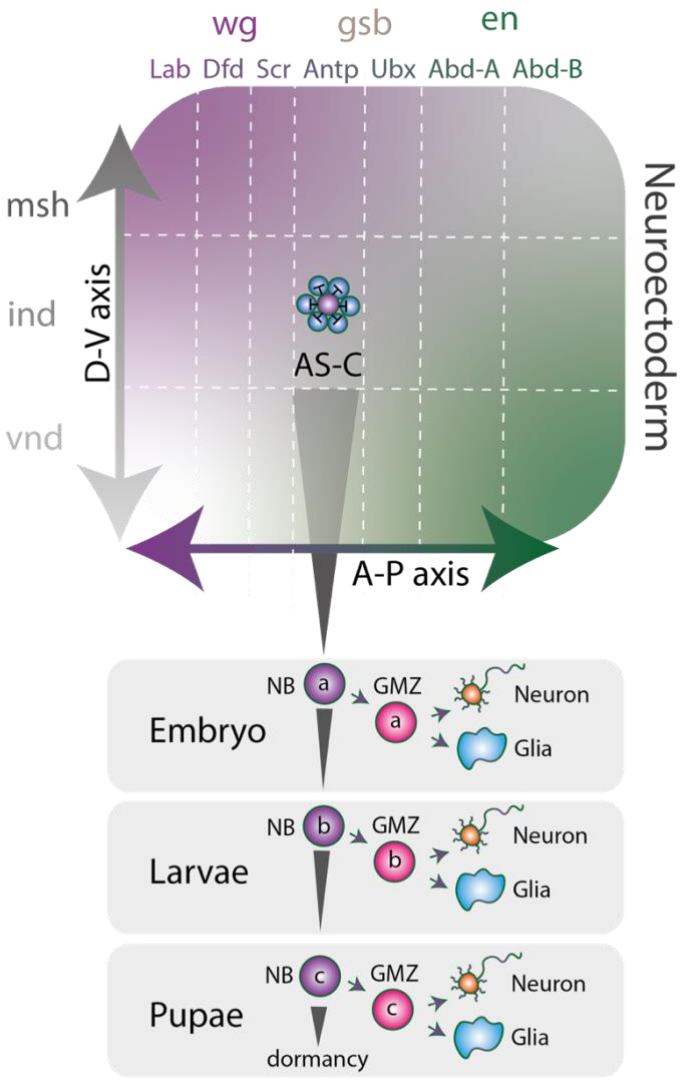


Figure 3: Development of the nervous system in *Drosophila*

The neuroectoderm provides positional information of each hemisegment by means of patterning genes. For example, the activity of segment polarization genes such as *wg*, *gsb* and *en* as well as *Hox* genes (*Lab*, *Dfd*, *Scr*, *Antp*, *Ubx*, *AbdA*, *AbdB*) define development of the AP axis, while homeotic genes such as *vnd*, *ind* and *msh* specify the DV axis. Thus, in the neuroectoderm each proneuronal cluster specified by the *AS-C* complex is defined by a unique code of patterning genes. Within every proneuronal cluster one cell becomes a NB by lateral inhibition of the neighboring cells. This cell is then capable to delaminate from the neuroectoderm and performs multiple rounds of division. Thereby, NBs give rise to ganglian mother cells (GMZ) that have the potency to develop into either a neuron or glia cell. In total three rounds of NB divisions occur at distinct developmental stages; In the late Embryo and larvae the laval nervous system is established, finally the adult nervous system is formed through NB division in the pupae. After these divisions NB undergo apoptosis or become dormant stem cells.

In a next step, NBs at stereotyped positions along the VNC divide asymmetrically to form ganglion mother cells (GMC) (Zhong and Chia, 2008). This process is repeated several times to form a lineage of GMCs that differ in their so-called birth order. In each round of NB division, a distinct temporal TF is present and remains in NB progeny (starting with Hunchback (Hb), to Kruppel (Kr), to Pdm1 (POU-homeodomain), to Cas (Castor), to Grainy head (Grh) expression). These sequential “temporal selector” TF expressions indicate the stereotyped birth order of prospective progeny. Finally, a dormant NB stem cell is maintained and proliferates again at the end of first larval instar stage to generate a secondary lineage of neurons that form the secondary axon tracts in *Drosophila* (Figure 3).

GMCs, in turn, have the capability to become postmitotic neurons or glia cells. Similar as during NB specification, this fate decision is regulated by lateral inhibition, mediated by Delta-Notch signaling (Egger et al., 2008). Glia cells have supportive functions for neurons and help with processing and storage of information. Glia cells and neurons derived from the same NB stay together and form the primary axon tract.
Ultimately, each NB can therefore be uniquely specified by its position and birth order (Figure 3). This concept is important, as it provides a foundation for the establishment of single-cell specific gene expression patterns in individual neurons.

1.1.3 HOMEODOMAIN TFs IN NERONAL DEVELOPMENT

Many of the genes involved in specifying neuronal position and birth order are homeodomain TFs, which are an important gene class for diversifying neurons in mammals (Sugino et al., 2019; Zeisel et al., 2018) and in *Drosophila* (Allen et al., 2020). Furthermore, they potentially serve as a molecular identifier of single neurons: Recent work on *C. elegans* has shown that each nematode neuron class is specified by a unique expression code of homeodomain TFs (Reilly et al., 2020). While the existence of such a code in more complex organisms has not been systematically investigated, homeodomain TFs are known to play an important role during the differentiation and cell type specification of neurons also in *Drosophila* and mouse (Deneris & Hobert, 2014; Domsch et al., 2019; Reilly et al., 2020; Urbach et al., 2006, 2016). Further evidence for a unique homeo-code in *Drosophila* is provided by a study on a set of postmitotic MNs that originate all from the same NBs. This study found that a combinatorial code of six homeodomain TFs can discriminate individual cells of the same lineage and define their morphology (Enriquez et al., 2015). Thus, homeodomain TFs cannot directly regulate cell morphology, they potentially regulate target genes of neurons that modulate these processes specifically.

1.1.3.1 TARGET SPECIFICITY OF HOMEODOMAIN TFs

The 60 amino acid homeodomain allows TFs to bind the homeobox motif on the DNA that is present on a multitude of gene families. While there are about 100 homeodomain TFs in *Drosophila*, the best studied ones are the seven highly conserved Hox TFs, which specify segments and tissues (Domsch et al., 2019; Lewis, 1978; McGinnis & Krumlauf, 1992). In insects, the chromosomal order of *Hox* genes reflects their expression pattern along the AP axis (from AP: *lab, Dfd, Scr, Antp, Ubx, AbdA, AbdB*) called co-linearity (Lewis, 1978). Ectopic expression studies showed that among the approximately 1500 Hox target genes, more than one

third of these genes are regulated specifically by one individual Hox TF (Hueber et al., 2007). Although these TFs possess a high degree of specificity in vivo, systematic analysis on target specificity of Hox TF in vitro revealed that Hox TFs and most homeodomain TFs can bind to almost identical DNA recognition sites, AT-rich motifs that on average appear every 1000bp in the *Drosophila* genome (Noyes et al., 2008). A possible explanation for this phenomenon is the cooperative interaction with different sets of TFs or co-factors that modulate target specificity (Chan & Mann, 1996). Notably, most well described Hox co-factors are TALEN (three amino acid loop extension) domain homeodomain TFs, for example _Homothorax (Hth), Engrailed (En) or Extradenticle (Exd) (Chan & Mann, 1996; Gebelein et al., 2004; Mann & Affolter, 1998).

1.1.3.2 THE FUNCTION OF HOMEODOMAIN TFs IN THE NEURONAL SYSTEM

A role of homeodomain TFs is described in different spatial compartments and at multiple steps of neurogenesis. They act predominantly as "spatial selectors", a term coined by Garcia-Bellido. Selector gene expression defines spatial embryonic compartments that modulate specific developmental programs by activation and repression of target genes (Allan & Thor, 2015b; García-Bellido, 1975) (Figure 4). As described above, during gastrulation, homeodomain TFs first specify the neuroectoderm and then NBs position. Several functional studies show the importance of homeodomain TFs in compartment specification; for example, in mutants for the homeodomain TF Engrailed (En) loss of the posterior compartment and neuronal identity is observed (Kornberg, 1981; Morata & Lawrence, 1975, 1979), the Hox TF AbdA and the homeodomain TF Caudal (Cad) regulate NB development in terminal abdominal compartments (Urbach et al., 2016) and in maxillary and mandibular segments of the head, loss of the anterior Hox TF Dfd causes severe head defects (Merrill et al., 1989).

In addition to spatial roles of homeodomain TFs, they also exhibit distinct roles at different developmental time points. For example, downstream targets of the homeodomain TF Ubx changes substantially during embryo maturation (Domsch et al., 2019), indicating differential roles of homeodomain TFs during cell type specification and later phases of development such as neuronal wiring. Hessinger and colleagues showed the function of Ubx in establishing neuronal wiring of ventral projecting MNs mediated by the Wnt signaling pathway (Hessinger et al., 2017), thus Ubx is required to guide the axon to the appropriate target site. Likewise, a study using temperature sensitive Dfd mutants showed that Dfd is necessary to maintain the motor apparatus for feeding in *Drosophila* larvae (Friedrich et al., 2016). In Dfd mutants, loss of Dfd caused a general misregulation of Connectin. Connectin is a cell adhesion molecule

required for connectivity of neurons and somatic muscle targets (Gould et al., 1990) thus a critical role of Dfd in the formation of the neuromuscular system has been suggested. Interestingly, the homeodomain TF En negatively regulates Connectin expression in subsets of interneurons (Siegler & Jia, 1999), indicating that different types of homeodomain TFs can regulate cell adhesion molecules in opposite ways. In line with these findings, reduced levels of different homeodomain TFs in leg MNs originating from the same NB can switch motoneuron innervation patterns in a predictable manner (Thor et al., 1999). These studies suggest that homeodomain TF expression also impacts late developmental and cellular programs such as cell adhesion and axon guidance.

To prove the function of homeodomain TFs in synaptic wiring, effects exerted during later stages of development need to be clearly distinguished from early developmental programs such as tissue specification. Many studies used mutants to detect defects on circuit formation due to functions of these factors in early development, it is therefore not clear whether the phenotypes are primary derived from defects on circuit formation (Baek et al., 2019; Buelow et al., 2005; D'Elia & Dasen, 2018; Dasen et al., 2005; Friedrich et al., 2016; Hessinger et al., 2017; Mallo, 2014; Philippidou & Dasen, 2015; Song & Pfaff, 2005). In addition, these studies focused on the function of single homeodomain TFs, thus a comprehensive understanding of the cooperative activity of homeodomain TFs is lacking.

Taken together, mechanisms for the molecular identification of individual neurons during the wiring of neuronal circuits, a possible involvement of the homeo-code in this process, and any possible downstream mediators between this code and synaptic matching have to date not been systematically investigated, and a mechanistic understanding of neuronal circuit formation is missing.

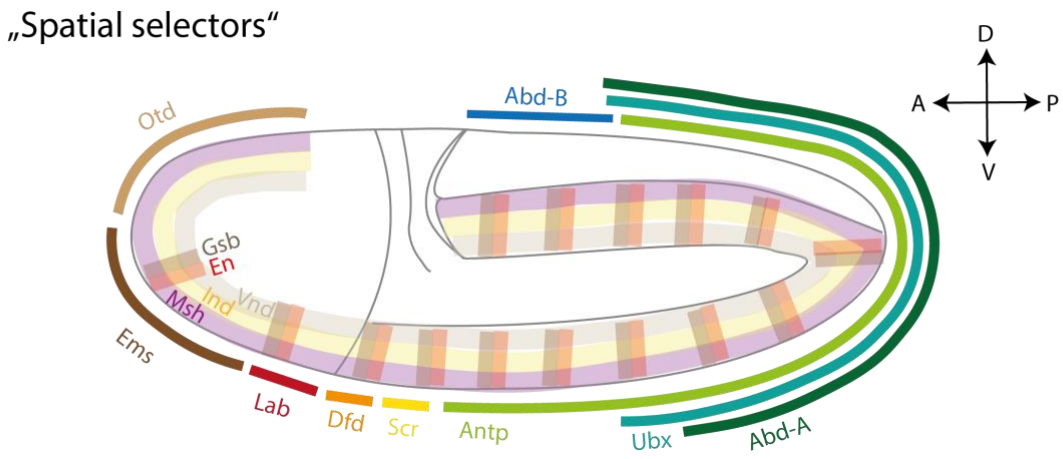


Figure 4: "Spatial selector" gene expression during gastrulation stage of embryonic development

The illustration depicts a *Drosophila* embryo during gastrulation stage, when NBs start to develop and represents expression patterns of selected homeodomain TFs. For example, combinations of Hox TFs are spatial selectors along the AP axis, from AP; labial (lab), Deformed (Dfd), Antenapedia (Antp), Ultrabithorax (Ubx), Abdominal-A (Abd-A) and Abdominal-B (Abd-B) and three homeotic genes, the *muscle specific homeobox* (*Msh*), the *intermediate nervous system defective (Ind*) and *ventral nervous system defective (Vnd)* spatially divides the DV axis. (Illustration is adapted from (Allan & Thor, 2015a).

Once each single neuron has been specified, a crucial step in the development of neuronal systems is the establishment of a highly specific set of connections between pre-defined single cells. Unlike the processes leading to cellular specification, this process is very poorly understood, and it will be the focus of my thesis.

1.2 SYNAPTIC SPECIFICITY AND SYNAPTIC PARTNER CHOICE

Neuronal circuits in the nervous system of mammals as well as in *Drosophila* are stereotypically wired for the precise execution of functional tasks critical for organismal survival. The formation of such circuits is a step-wise process, which starts with the specification of neuronal cell types and their accurate arrangements in space, followed by the correct wiring of individual cells and their final integration into a functional network. According to the 'labelled pathway hypothesis' (Sperry, 1963), the differential expression of axon guidance molecules, also called "individual identification tags", enable neurons selectively to find paths towards targets with specific chemical affinities. These axon guidance molecules are diffusible, long range signalling molecules that are sensed by receptors to either attract or repulse the growth cone of a neuron from the target cell (Kolodziej et al., 1996; Labrador et al., 2005; Matthes et al., 1995; Mitchell et al., 1996; H Sink & Whitington, 1991; Winberg et al., 1998) (Figure 5). During this process of pathfinding, neurons stochastically and transiently form contacts with many possible targets in the area of attraction. Finally, neuron growth cones and target cells with matching identification factors come in contact and establish stereotypic connections with extraordinary precision, a process called synaptic specificity (Sanes & Zipursky, 2020).

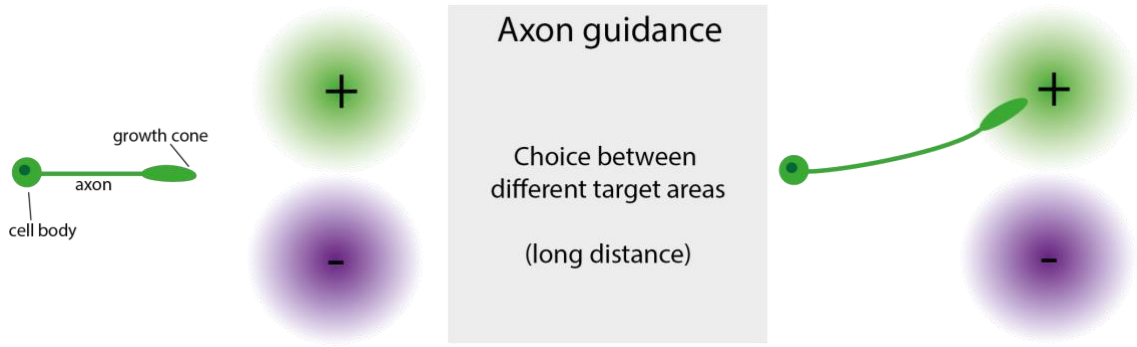


Figure 5: Axon guidance during neuronal development

During neuronal development neurons extend their axon towards target areas with attractive guidance cues. The tip of the axon called growth cones expresses axon guidance receptors that bind to diffusible ligands send by target areas. These diffusible long-range signals are termed axon guidance factors and modulate synaptic target choice.

In 1984, Goodman hypothesized that groups of cell surface proteins (CSPs) can either act as short range attractive or repellent cues to increase or decrease likelihood of synapse formation with target cells (Goodman et al., 1984) (Figure 6). These CSPs are cell adhesion molecules belonging to different molecule classes such as Ig superfamily (Thiery et al., 1977), Robos, Leucin rich repeat proteins (Kurusu et al., 2008; Nose et al., 1992, 1997; Shishido et al., 1998), Ephrins, Cadherins (Takeichi, 2018) and Semaphorins (Y Luo et al., 1993; Yuling Luo et al., 1995; Matthes et al., 1995). The large Ig domain superfamily is further divided in many subfamilies comprising Down syndrome cell adhesion molecules (Dscams), Dprs/DIPs, Beats/Sides, Contactins (Cnts), Sidekicks (Sdks) and L1 (H. Li et al., 2017). In vitro studies and studies on single CSPs in vivo demonstrate that molecules of the same class are more likely to interact with each other. Two mechanisms of interaction are distinguished: the homophilic interaction of the same type of molecule and/or heterophilic interaction with a different type of molecule. For example, Cadherin molecules exhibit a Ca2+ dependent Cadherin domain that mediates homophilic binding, while Immunoglobulin (Ig) domain proteins possess a beta-sandwich shaped external Ig repeat (Shapiro et al., 2007) for adhesive homophilic (Shapiro, 2007; Soroka et al., 2003; Zhao et al., 1998) or heterophilic interactions (Carrillo et al., 2015; Özkan et al., 2013). Homophilic interactions of Dscams and heterophilic interactions of Beat/Side molecules often prevent connections of the same type of neuron by self-avoidance mechanisms (Garrett et al., 2018; Matthews et al., 2007), while heterophilic interactions of Dpr/DIPs are more implicated in cellular connectivity mechanisms (Carrillo et al., 2015). However, little is known how these intercellular interactions by CSPs contribute to synaptic target choices in vivo. A number of studies suggested that Ig domain proteins interact with the

cytoskeleton to modulate repulsion or attraction of axon growth cones(Okumura et al., 2015; Özkan et al., 2013).

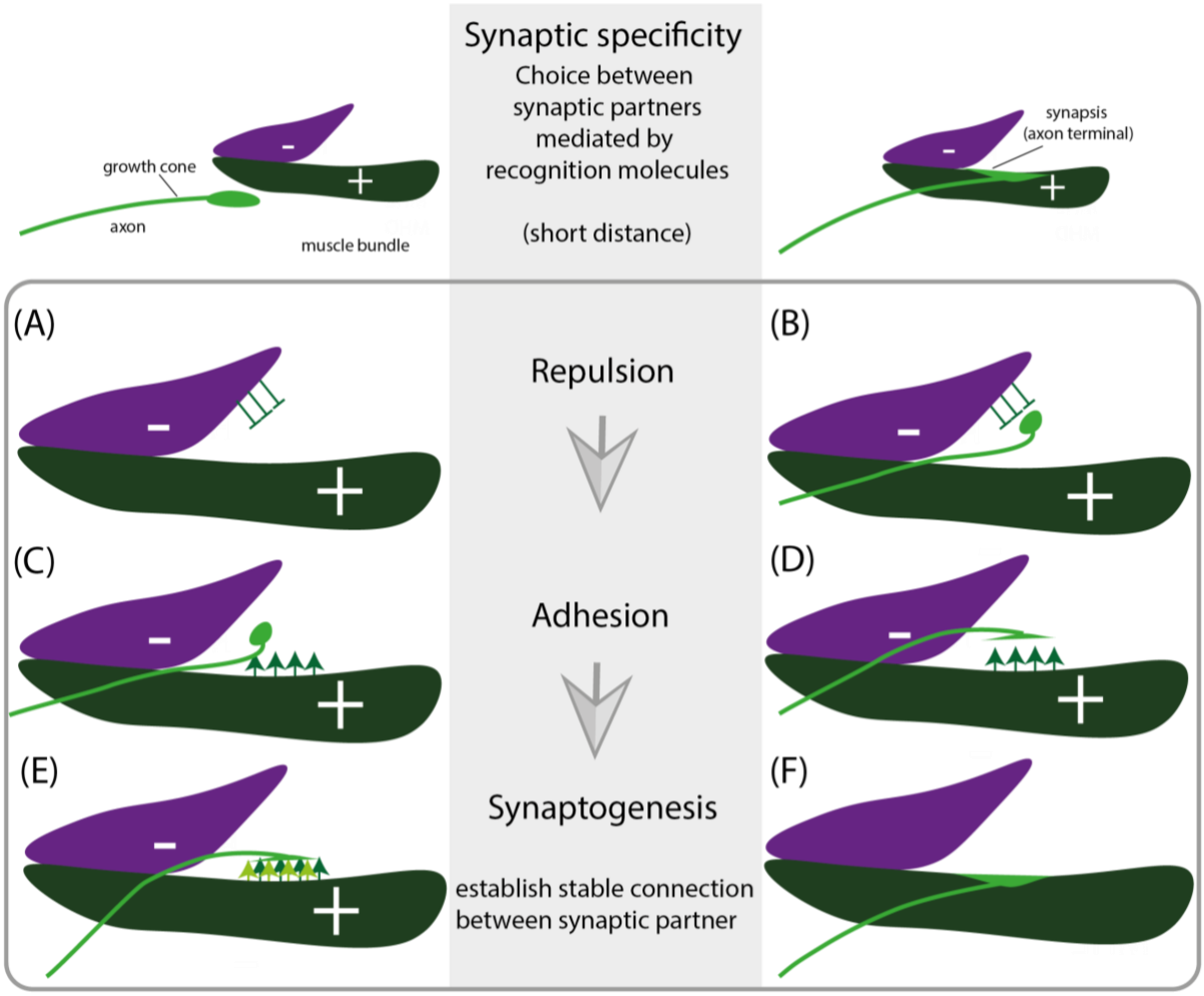


Figure 6: Synaptic partner choice mediated by cell surface molecules

Neurons select their correct synaptic partner by direct cell interactions with their axon growth cone. The growth cone senses the cell surface of a potential target cell for attractive or repulsive cell adhesion molecules to mediate a match, a process termed synaptic specificity. (A) A potential target cell (here a muscle cell in magenta) expresses cell adhesion molecules with repellent properties (B) leading to repulsion of the growth cone. The neighboring potential target cell (muscle cell in green) expresses adhesive cell surface molecules (C) mediating synaptic matching (D). Finally, more abundantly expressed cell surface molecules (E) establish stable synaptic connections between synaptic partners (F).

Nevertheless, not all Ig domain interactions can be explained by this mechanism (Cheng et al., 2019). For example, the functional role of recently discovered Dpr/DIPs is unclear. In only one example it has been shown that a pair of Dpr/DIP promoted retrograde Bone Morphogenetic Protein (BMP) signalling and thereby enhances synaptic growth (Carrillo et al., 2015). In addition, a functional role of Dpr/DIP interactions is suggested in cell survival, hence neurons lacking correct Dpr/DIP interactions are eliminated by apoptosis in the Drosophila (Carrillo et al., 2015; Xu et al., 2018). Until today, predictions on downstream signalling directed by

combinations of CSPs in vivo are challenging, because the cellular output highly depends on multiple parameters; the type of interaction (homophilic or heterophilic), the strength, cooperative interactions of different sets of CSPs and the spatial and temporal context. For example, the combinatorial interaction of cell adhesive Cadherins and repulsive Dscams on the cell surface mediates masking of adhesive properties, thereby matching of incorrect cells is prevented (Garrett et al., 2018). Thus, cell-cell interactions are mediated by a complex interaction network of CSPs with diverse functional outcomes.

A comprehensive understanding how adhesive and repellent interactions in vivo are modulated on the single cell level is still lacking. Recent work has shown that combinations of IgSF cell surface proteins (Dprs) are differentially expressed in distinct neuronal clusters and bind to specific Dpr binding proteins (DIPs) expressed in synaptic partners (Carrillo et al., 2015; Nakamura et al., 2002; Özkan et al., 2013). In the visual system, combinations of CSPs are expressed differentially in different layers (Tan et al., 2015), while in olfactory neurons, a combinatorial code of TFs and CSPs map neurons with the same olfactory receptor to the same glomerulus (Couto et al., 2005; H. Li et al., 2017, 2020). These studies indicate that a molecular mechanism is imprinted in the genome that might coordinate specific expression of CSPs. Nevertheless, all these studies explain how groups of similar neuronal cells are molecularly defined and only provide a hypothesis on how stereotypic connections to another neuronal cell type are formed.

Further molecular and cellular mechanisms are described to regulate synaptic specificity. For example, birth time order of NBs described above determines coordinated growth of synaptic partners and specifies progenitor fate that is possibly required for specific connectivity mechanisms. In addition, only neurons with correct synaptic partners are stabilized by CSPs, otherwise synapses or entire neurons are eliminated by apoptosis (Courgeon & Desplan, 2019). Until the end of embryogenesis, the first round of stereotyped synaptic wiring and subsequent differentiation into functional synapses in *Drosophila* is completed. The contacts between neurons and targets are continued to be stabilized by specific cell adhesion molecules such as Disc-large (Dlg) (Budnik et al., 1990).

In sum, many lines of evidence show that long range signalling by axon guidance molecules defines the coarse direction of axon growth and finally precise synaptic matching is mediated by a complex and not well understood combination of CSPs and other mechanisms such as

birth order, or synapse elimination. It is, however, unclear how cellular identities specific to unique single cells are translated to stereotypic wiring diagrams.

To shed further light on these processes, it will be essential to a) work with a model system with small cellular complexity and a well-known connectivity map, and b) use state-of-the art methods for measuring gene expression in individual cells and c) employ a powerful set of genetic tools for determining functional associations. In the following, I introduce the *Drosophila* motoneuronal system as a well-suited model system. Finally, I provide a survey of single-cell transcriptomics and *Drosophila* genetics, which constitute the methodological basis for my work.

1.3 THE ANATOMY OF THE NERVOUS SYSTEM FOR LOCOMOTION

The mature nervous system of *Drosophila* larvae is anatomically relatively simple, however capable of preforming complex motor behaviors similar to vertebrate organisms. The *Drosophila* larvae possesses a highly stereotypic innervation pattern of body wall muscles that consists of 30 repeated sets of somatic muscles in each segment. Sets of about 35 MNs project across three nerve projections, transverse (TN), intersegmental nerve (ISN) and segmental nerve (SN), that further divide into nerve branches and finally into single axon projections specific for each muscle cell (Matthias Landgraf et al., 1997a; Van Vactor et al., 1993). Cell bodies of these MNs originate from the VNC (Figure 7).

The basic structure of the VNC is analogous to the spinal cord in vertebrates (D Arendt & Nübler-Jung, 1996). In *Drosophila* larvae, the mature VNC is a fusion of 14 segmental ganglia that are connected via commissural axon tracs. Three of these ganglia are located in gnathal (C1-C3), three in thoracic (T1-T3) and eight in abdominal (A1-A8) segments. Each axonal tract crosses the midline stereotypically and forms a structure similar to a rope ladder. The VNC forms the integration center for sensory inputs and locomotor outputs, enabling behaviors such as walking, grooming and mating, while the SOG, a part of the VNC, integrates neuronal signaling from a higher brain region of the CNS called Mushroom Body (MB). This brain region is involved in learning and short-term memory formation. Thus, the connection of the VNC with the MB allows for more differentiated motor behaviors such as feeding in responses to pre-rated stimuli.

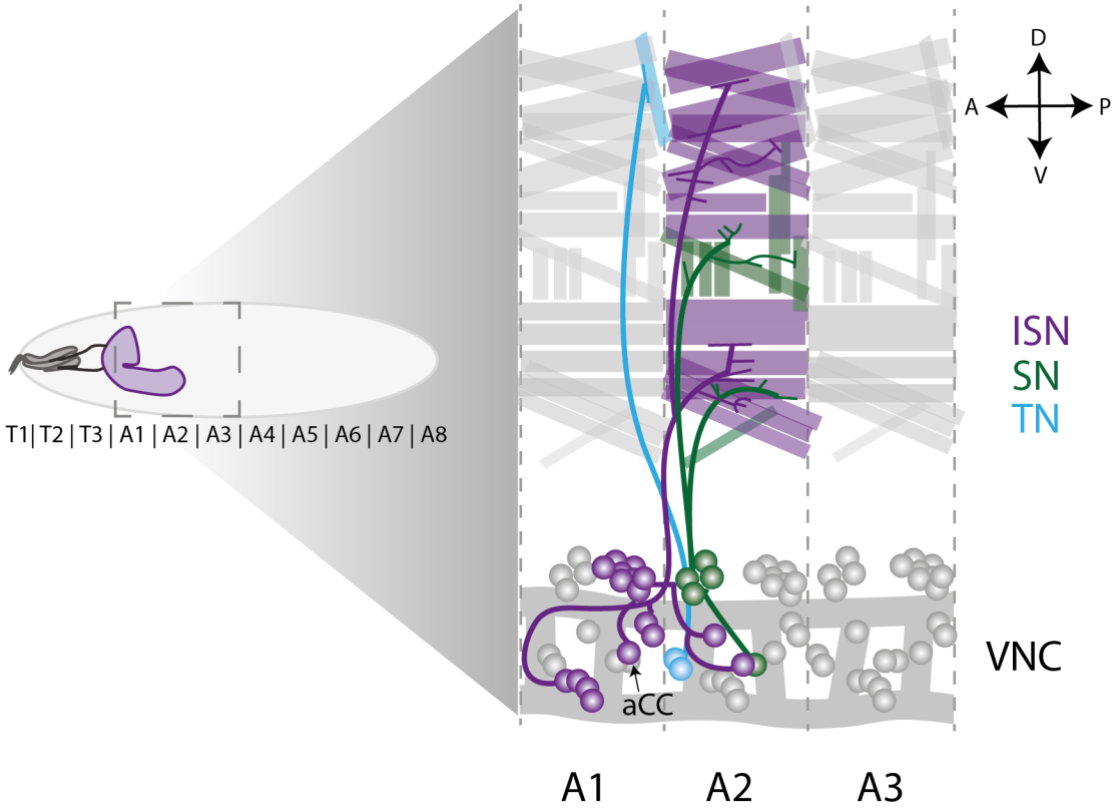


Figure 7: Stereotyped motoneuron innervation patterns in abdominal segments of the larvae

The illustration depicts repetitive innervation patterns of MNs in abdominal segments. The three major nerve projections are the ISN, the SN and the TN. Multiple MNs are bundled in each nerve projection. At muscle target sites, these axon bundles divide into single axon projections for single cell specific innervations. Cell bodies of each motoneuron originates at pre-specified positions from the same segment than the muscle target or the anterior located abdominal segment. Approximately 35 MNs innervate 30 somatic muscle cells in segments A2-A7 (Matthias Landgraf et al., 1997b). In the late embryo these connections are only partially established, while in late larval stages these innervation patterns are completed.

Feeding is a complex behaviour: *Drosophila* has to take up food extensively during larval stage to increase body size. Feeding is performed by rhythmic extension and retraction movements of the head skeleton, the so-called cephalopharyngeal skeleton (CPS). These movements are coordinated by elevation and depression of the mouth hooks and mandible-derived structures required for chopping up food (Schoofs et al., 2010). Therefore, several CNS nerves emerging from the SOG control repetitive cycles of feeding movements by means of their muscle targets: the prothoracic nerve controls the protractor muscles for tilting movements, while the maxillary nerve coordinates movements of the mouth hook elevator (MHE) and the mouth hook depressor (MHD) muscle for mouth hook elevation and depression movements and the antennal nerve regulates muscle contractions of the cibarial dilator muscle for food ingestion (Schoofs et al., 2010). Recently, a map of the feeding connectome in *Drosophila* larvae further advanced our understanding of connectivity down to the single-cell level (Miroschnikow et al., 2018) (Figure 8).

Taken together, owed to its simplicity and well understood connectivity patterns, the motoneuronal system of *Drosophila* is a well-suited model system for studying the molecular mechanisms of motor-neuronal circuit formation. In the following, I will review what is known about its development and the establishment of specific connectivity patterns in that system.

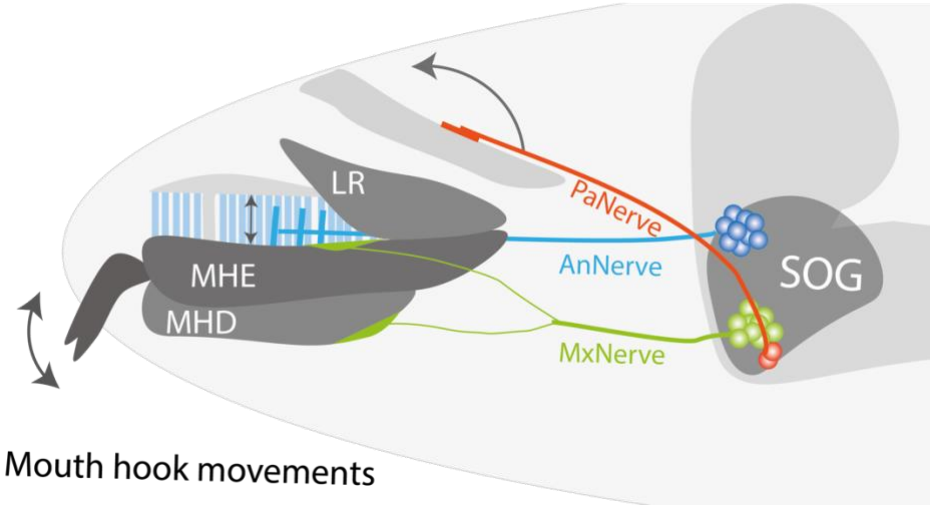


Figure 8: The laval feeding circuit that is established in the late embryo

In the larval *Drosophila* head three major nerve branches control feeding movements, the maxillary nerve for mouth hook elevation and retraction movements, the pharyngal nerve for tilling movements and the antennal nerve for pharyngal pumping project from the SOG to distinct feeding muscles by glutamergic MNs. Nine MNs belong to the maxillary nerve (MX), three of them control the MHE muscle (Friedrich et al., 2016) Two MNs belong to the pharyngal nerve (PA) and another eleven MNs belong to the antennal nerve (AN) (Hückesfeld et al., 2015). MNs of the MX nerve are fully established and control mouthhook muscles in the late embryo to enable hatching movements. Elements of the illustration are adapted from Miroschnikow et al. 2018.

1.3.1 DEVELOPMENT OF THE MOTONEURONAL SYSTEM

In the example of the *Drosophila* neuromuscular system, every single neuron within the developing embryo is defined by its position and forms a unique and stereotypic set of connections with target muscles (Allan & Thor, 2015a).

Development of such a precisely encoded network is tightly regulated and requires a stepwise execution of processes. MNs are obtained from 15 out 30 NBs of each hemisegment (Matthias Landgraf & Thor, 2006). During NB divisions as described above, a progressive cascade of temporal and spatial TFs specifies diverse cell fates and finally defines a subtype specific signature for MNs depending on position and birth order. For example MNs are specified by segment specific Hox TF's from AP (Angelini & Kaufman, 2005; Bossing et al., 1996; Schmid et al., 1999; Schmidt et al., 1997), and further defined by TFs such as LIM-homeodomain TFs along the DV axis (Broihier et al., 2004; Broihier & Skeath, 2002; Matthias Landgraf & Thor,

2006). In embryonic stage 13 to 14, MN differentiation begins, starting with anterior located anterior corner cells (aCC) in each hemisegment that are part of the ISN nerve (Figure 7). Neuron such as the aCC are called “pioneer MNs” and define the first axon tract that sets the primary path for new born MN in close proximity (Van Vactor et al., 1993). These MNs extend their nerve projections to predefined locations in all three spatial dimensions, while precise matching of muscle targets occurs earliest in stage 16 to 17 of embryonic development (Broihier et al., 2004; Broihier & Skeath, 2002; Hessinger et al., 2017; Matthias Landgraf et al., 1997b; Thor et al., 1999). Thus, the first round of MN wiring is of particular importance, because primary axon tracts are formed that are specified by the position of pioneer MNs (Van Vactor et al., 1993). In the following generations, more MNs follow these tracts and complete the formation of the entire nervous system until larval stages. Identity of these MN subtypes can be described by spatial TFs. However, little is known about the molecular factors specifying unique MN identities and pre-define synaptic connections for the exact muscle target.

In parallel to MN differentiation, somatic muscle cells are specified during development by the expression of identity genes according to position, size and innervation sites. Muscle precursor cells termed “founder myoblasts” migrate to their relative position along the AP axis that is defined by Hox TFs expression (Campos-Ortega & Hartenstein, 1997). Once founder myoblasts found their appropriate position, they fuse to a group of neighboring fusion-competent myoblasts to form individual muscle cells (Weitkunat et al., 2014). Until embryonic stage 17, muscle cells grow towards attachment sites and form muscle fibers. These muscle cells possess postsynaptic filopodia termed “myopodia” on their surface to present receptors for cell-cell interaction with MNs and are crucial for precise synaptic match making (Kohsaka & Nose, 2009; S Ritzenthaler et al., 2000; Sarah Ritzenthaler & Chiba, 2003).

The coordinated development of MNs and muscle cells and the regulation by similar spatial cues have suggested region-specific mechanisms in the specification of connectivity patterns. However, on the single cell connectome level it becomes clear that selection of single muscle targets by MNs follows more complex rules than the segmental spatial location. MNs and their muscle partners are not precisely aligned along the body axes (Couton et al., 2015; Matthias Landgraf et al., 1997a, 2003; Matthias Landgraf & Thor, 2006). Nerve projections not only bundle axons of MN originating from the same segment, but also axons from the anterior located segment. Thus, some axon projections can cross segmental boundaries for synaptic matching (Landgraf et al., 2003; Nassif et al., 1998). However, the position of each motoneuron cell body and the muscle target are invariant, indicating that intrinsic molecular identities must exist to define the position of MN along the VNC (Figure 7).

Thus, it remains unclear how unique post-mitotic neurons are specified and how this identity contribute to single cell specific connectivity patterns. The neuromuscular system in *Drosophila* provides an excellent model to study these mechanisms, due to the relative simplicity and well-described nature.

1.4 SINGLE CELL TRANSCRIPTOMICS

In preceding sections, I have introduced the complex task of a single cell to find and connect with a pre-specified wiring partner out of a large number of possibilities. In addition, I have pointed out that a multitude of factors (e.g. spatial location, birth order) confer a unique identity to single neurons (Figure 9). To investigate if these identities are associated with unique molecular profiles that drive synaptic specificity, I made use of single cell transcriptomic approaches.

In 2011, a scRNA-Seq protocol permitted for the first time the identification of transcriptional profiles of single cells from a mouse blastula (Tang et al., 2011). This novel single cell assay was based on a modified microarray protocol including a reverse transcription step, a second strand synthesis, a cDNA preamplification step, adaptor ligation for library amplification and finally next generation sequencing. During the last decade, quality and throughput of single cell sequencing techniques has improved drastically.

On one side, single cell sequencing experiments have been scaled up using microdroplet-based microfluidic technologies that can sequence thousands of transcriptomes at a time for low amounts of costs (Klein et al., 2015; Macosko et al., 2015; Zheng et al., 2017). These techniques start with tissue dissociation and a subsequent dispersion of single cells into water-oil droplets that contain a PCR barcode mix for reverse transcription and library preparation. In these high-throughput methods, individual single cells are typically sequenced at a low sequencing depth, thus especially low expressed genes such as cell adhesion molecules are captured at a lower frequency and dropout rates (technical noise) are relatively high. On the other side, all steps involved in single-cell RNA-seq have been optimized excessively in well-based formats (Hagemann-Jensen et al., 2020; Picelli et al., 2014). While these protocols have a lower throughput compared to droplet-based methods, they allow for more genes to be detected in single cells. Technical noise and dropout rates are reduced compared to droplet-based methods (Ziegenhain et al., 2017).

Concurrent with advances in single-cell RNA-seq protocols, bioinformatic routines to visualize and analyse these complex datasets have been developed. Common first steps in these routines are quality control, selection of highly variable genes (Brennecke et al., 2013), and projection to a lower dimensional space using linear methods such as principal component analysis (PCA). The latter step primarily serves to remove noise from the data by pooling information between co-expressed genes (Luecken & Theis, 2019). However, linear methods are incapable of projecting highly complex data topologies (such as cell types and their transcriptomic relationships) into a simple 2-dimensional space. It is therefore common practice to retain 10-50 dimensions from the PCA, and for visualization purposes project those to a two-dimensional space using non-linear methods, such as t-SNE (van der Maaten, 2008), uMAP (McInnes et al., 2018) or PHATE (Moon et al., 2019). Independent of visualization, clustering is performed using a variety of approaches, such as graph-based clustering or hierarchical methods.

Building on these advances in technology and bioinformatic methods, many reference maps for cell types in a multitude of tissues and organisms were created (Davie et al., 2018; Ecker et al., 2017; Hung et al., 2020; Kunst et al., 2019; Rust et al., 2020). For example, in the context of neurodevelopmental biology, work on olfactory projection neurons in *Drosophila* has shown that neuronal cell types are specified by a combinatorial molecular code of TFs (H. Li et al., 2017; J. Li et al., 2020). However, these studies have identified cell types or subtypes, i.e. presumably homogeneous groups of cells. By contrast, decomposing the unique molecular identities of individual cells has not been achieved yet.

1.4.1 SPATIAL TRANSCRIPTOMICS

A shortcoming of single cell transcriptomics is the lack of spatial information due to tissue disruption during the tissue dissociation step of single cell protocols. However, to understand cellular identities, it is necessary to retrieve the spatial localisation, because it provides information about the developmental origin, circuit architecture and cellular context (Figure 9). A number of methods have recently been developed to profile gene expression in tissues while retaining spatial information. However, these methods currently either lack single-cell resolution and are therefore not applicable in the context of identifying the unique identity of single neurons (Moncada et al., 2020; Ståhl et al., 2016) are based on microscopic technologies that are available only to a handful of laboratories worldwide (K. H. Chen et al., 2015) / MERFISH, (Lee et al., 2014) / FISSEQ, (Eng et al., 2019) / seqFISH). To overcome this bottleneck, a number of recent papers permit the superimposition of spatial information on

single-cell gene expression atlases by drawing upon existing data resources generated with classical methods such as in situ hybridisation or immunoflourescence microscopy (Achim et al., 2015; Bageritz et al., 2019; Satija et al., 2015). Thereby, they enable the identification of molecular signatures that may encode positional identity and regulators of synaptic specificity.

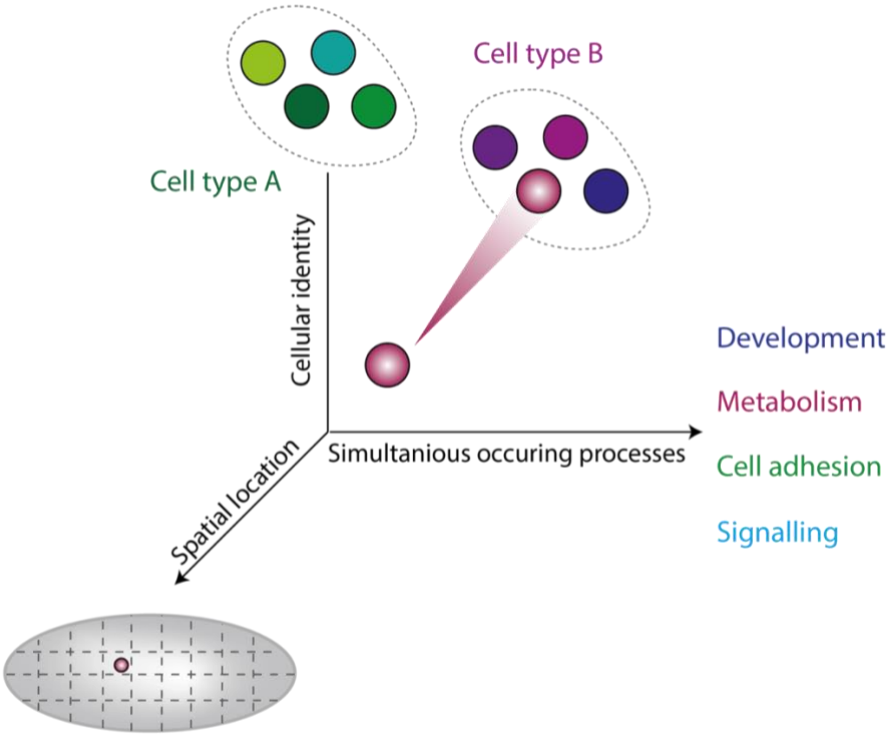


Figure 9: Single cell sequencing as tool for describing cellular identities

The illustration depicts the multitude of vectors that can be dissected by spatial single cell transcriptomics to define cellular identities. A high coverage of single cell transcriptomes allows for identification of transcriptional heterogeneity beyond the cell type level. Multiple simultaneously occurring processes in the cells can be identified, such as developmental processes mediated by TFs, specific cell adhesion properties, the metabolomic state, and simultaneously arising signaling cascades. Finally, the spatial information of a cell adds information of cellular context and position.

1.5 THE MOLECULAR TOOLBOX OF FLIES

Single cell sequencing approaches provide an ideal starting point for an unbiased identification of molecular programs involved in distinct molecular processes such as synaptic specificity, however functional insights into genes associated with these processes can only be provided by genetic approaches such as gain and loss of function studies.

The fruit fly "*Drosophila melanogaster*" became first popular as a model organism for genetics in 1910 after Thomas Morgan discovered the chromosome theory of inheritance (Morgan, 1910). He characterized thousands of generations of fruit flies to identify the exact order of genes on chromosomes (Bridges, 1935). These findings permitted him to determine genetic

distances of genes in “centimorgan”, that allows researchers to calculate the probability rate for successful chromosome recombination experiments. *Drosophila*’s well characterized genome and the ability of *Drosophila* to create stably inbred stock over multiple generations made *Drosophila* the ideal model system for genetics. Thus, an extraordinary catalog of *Drosophila* stocks with diverse genotypes, such as knockouts for every gene in the *Drosophila* genome (“a genetic toolbox”), GFP-fusion genes, etc. has been built over the last century. The genotypes of stocks are visible by “marker” mutations such as eye phenotypes. This allows to select for the right progeny that inherited the proper “marker” mutation after crossing of different genotypes. Different genotypes can be combined through simple crosses that take about two weeks. To prevent recombination in females, single chromosome alleles were modified by chromosome rearrangements and labeled by visible marker gene expression termed balancer chromosomes. This system allows for fast generation of stable fly stocks with defined genotypes (Hales et al., 2015).

In addition to mutant lines, the fly stock catalog of *Drosophila* provides an extensive number of genes under the control of the binary GAL4/UAS expression system (Brand and Perrimon 1993) that enables the expression of any gene of interest in a spatial and time-controlled manner. This binary system consists of a yeast TF “GAL4” driven by a tissue specific promotor or enhancer and an upstream activating sequence “UAS” that can activate expression of a gene of interest upon GAL4 activation. This is a modular system, that can be combined by simple fly crosses in various combinations to manipulate gene expression in many tissues. For example, the UAS/GAL4 system can be used to investigate gene function by RNA interference. Here, a short-inverted repeat RNA hairpin for a specific target gene is directed by the UAS/GAL4 system to reduce the expression of the target mRNA in the tissue of interest, called “knockdown”. The combination of these two systems is especially important, if genes possess a potential function in multiple developmental processes in a spatial and time dependent manner such as homeodomain TF. In contrast, applying mutations on these genes would cause lethality of the embryo or loss of specific tissues early in embryonic development, thus examinations of the role of these genes in late developmental stages are unreliable. Hence the versatile genetic systems in *Drosophila* ideal to study fundamental developmental processes.

In 2000, *Drosophila melanogaster* was the second multicellular organism whose entire genome was sequenced by shotgun sequencing (Adams et al., 2000). Genome predictions estimate about 13,920 (Flybase.org) or 14,692 (Brown et al., 2014) protein-coding genes until today, rather small compared to the 20,000-25,000 human protein encoding genes. However, many (75%)

human genes associated with genetic diseases possess an ortholog in *Drosophila melanogaster*, thus fundamental pathways critical for human health are conserved from humans to fly (Lloyd & Taylor, 2010; Reiter et al., 2001).

In sum, *Drosophila* is an excellent model system to study fundamental processes of development and disease with a variety of genomic and genetic tools (Hales et al., 2015).

1.6 AIMS

The aim of my PhD thesis is to gain a comprehensive understanding of interconnected molecular mechanisms driving the extraordinary precision of neuronal circuit formation. This precision goes up to the single cell level and cannot be described by current models, where axons follow coarse chemical gradients, and CSPs mediate adhesive interactions between defined groups of cells. Furthermore, we are lacking knowledge of upstream mechanisms fine-tuning the expression of recognition molecules in a coordinated manner critical for individual cell-cell interactions. The reasons are a lack of single cell specific neuronal markers, the complexity of neuronal systems and the possibly gradual and combinatorial nature of CSP expression.

Therefore I here studied these processes using an experimental system with reduced neuronal complexity that allowed me to capture neuronal diversity (1) on a single cell level; (2) exactly at the time when stereotypic connections are formed, so as to identify critical molecular regulators for this process; (3) with an unbiased scRNA-Seq approach to define cell specific markers for neuronal identity and capture the major driving forces diversifying cells; (4) including a spatial mapping approach based on Hox gene expression to locate MNs along their AP position and thereby gain insight into the role of spatial mechanisms during wiring.

2. RESULTS

Text and figures of the following chapter are adapted from a preprint deposited on biorxiv:
Jessica Velten, Rashi Agarwal, Patrick van Nierop, Lena Bognar, Malte Paulsen, Lars Velten, Ingrid Lohmann: The molecular logic of synaptic wiring at the single cell level. doi: 10.1101/2020.11.30.402057v1
As stated in the preprint, I performed the experiments myself with contributions by R.A., L.B. and M.P. Bioinformatic data analysis was performed jointly by myself and L.V. I wrote the text and generated the figures presented below myself, with contributions from I.L. and L.V.

2.1 A REFERENCE MAP OF MNs DURING THE SYNAPTIC WIRING PHASE

To generate a spatially resolved molecular map of MNs during the synaptic wiring phase, I performed scRNA-Seq of cells marked by the motoneuron-specific *OK371*-GAL4 driver (Mahr & Aberle, 2006) controlling the expression of the UAS-*RFP* transgene. The glutamergic driver line (*Ok371*-GAL4) is based on a regulatory element of the glutamergic neurotransmitter receptor (VGlut). As soon as synaptic connections are established, expression of the glutamate neurotransmitter is turned on for neurotransmission in MNs. It is critical to map molecular profiles in this tight time frame, since according to (H. Li et al., 2017) neurons diversify most during synaptic wiring phase and become indistinguishable upon completion of neuronal connectivity. Hence, the onset phase of *Ok371* activity in the embryo is ideal to select MNs specifically in the process of synaptic wiring.
Development of synaptic connections in the larval nervous system is a step-wise process, beginning with a low number of stable synaptic connections in the late embryo and significant increase number of connections throughout larval development (Couton et al., 2015; Mahr & Aberle, 2006). In this study I focus on the first round of connectivity, therefore a significant lower number of Ok371-positive MNs wiring with a target cell are expected than the final set of 35 MNs in each hemisegment of the larval nervous system. Using microscopy, I found that approximately 140 cells are positive for *Ok371* activity at the early synaptic wiring phase (stage 17).

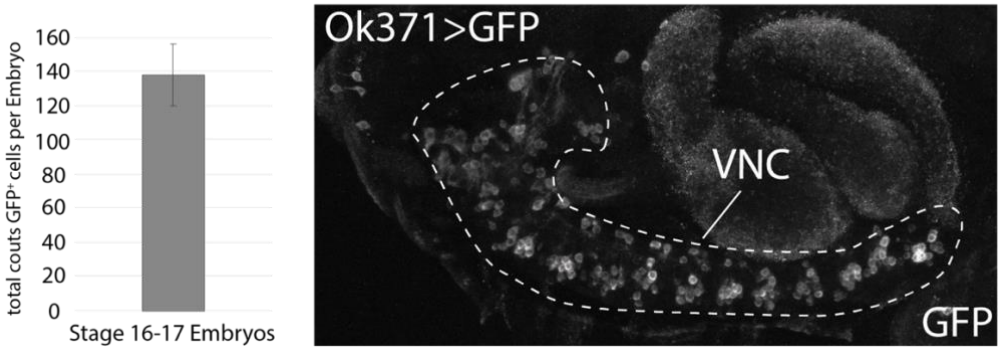


Figure 10: About 140 MNs are labelled by reporter gene expression of the Ok371-GAL4 line that drives at the onset of synaptic connectivity in late-stage *Drosophila* embryos

Quantification of the average number of terminally differentiated *OK371>GFP* positive MNs in stage 16-17 embryos (n=3) along the VNC.

Due to this low number of putative glutamatergic MNs, I considered single-cell sequencing to be the ideal method for analysis, allowing for a high coverage of single cells (Figure 10) (Mahr & Aberle, 2006). Single *RFP*-expressing MNs were sorted from a pool of precisely staged embryos by Fluorescence-Activated Cell Sorting (FACS) (Figure 11, Methods). Importantly, I implemented a custom modification of the SMART-Seq2 protocol for single-cell transcriptomics that permits to increase representation of *Hox* genes as spatial markers. *Hox* genes are known to be expressed in a consecutive order along the AP axis of *Drosophila*. I used this property to precisely locate single cell transcriptomes along the AP axis as further described below. To this end, I implemented a custom modification of the SMART-Seq2 protocol by adding primers targeting each *Hox* gene to the reverse transcription (RT) and preamplification step that permits an increased representation of *Hox* gene expression as spatial markers (Figure 11, see Materials and Methods). Then, I added primers targeting each *Hox* gene to the reverse transcription (RT) and preamplification step of the SMART-Seq2 protocol. After library preparation of each single cell, I performed next generation sequencing to obtain single cell transcriptomes of putative glutamergic MNs in the *Drosophila* embryo.

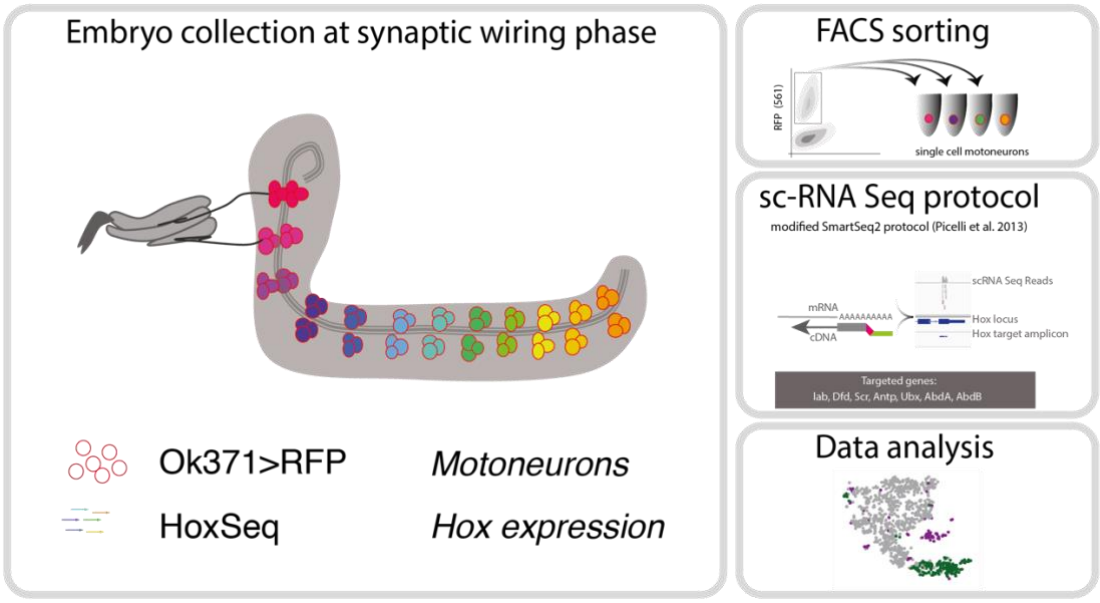


Figure 11 A: ScRNA sequencing of embryonic *Drosophila* MNs at synaptic wiring

Experimental design. Schematic drawing depicts the lateral view on an early stage 17 *Drosophila* embryo. MNs expressing RFP under control of *Ok371*-GAL4 are color coded along the VNC according to different patterns of *Hox* gene expression. ScRNA-Seq protocol of MNs was performed using a targeted protocol to enrich for *Hox* genes (*Hox*Seq).

In total, 1536 cells were sequenced by SMART-Seq2 (Picelli et al., 2014). After filtering, 999 single-cell transcriptomes were retained, i.e. each biologically unique RFP positive cell was covered on average 7 times in our dataset. A median of 1202 genes were observed per cell (Figure 12 A) and a negligible fraction of 0.1% of reads mapped to the mitochondrial genome, demonstrating that the experimental treatment hardly induced any stress responses on the transcriptional level. Despite the low expression of *Hox* genes in late embryonic stages, I identified at least one *Hox* gene to be expressed in 75% of the MNs (Figure 12 B). Thus, coverage of Hox genes is sufficient for the use as a marker for AP position in our data set. Finally, all three populations showed abundant expression of motoneuronal marker genes like *Vesicular glutamate transporter (VGlut)* and *embryonic lethal abnormal vision (elav)*, thereby confirming their motoneuronal identity (Figure 13.1).

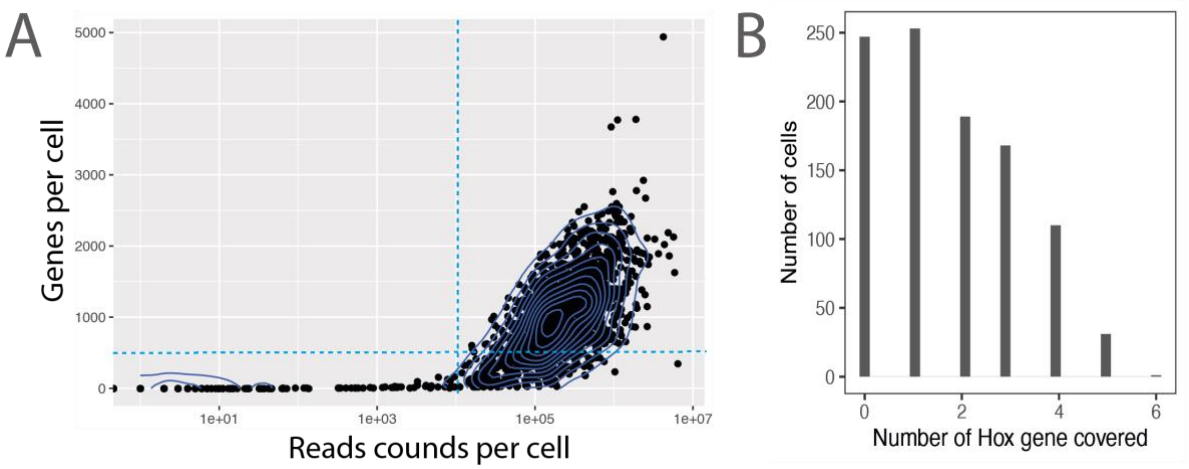


Figure 12: High quality single-cell RNA sequencing of embryonic *Drosophila* MNs

(A) Visualization of filtering criteria for single cells (dashed blue line, see Materials and Methods). Density dot plot represents the total reads (library size) versus genes observed per cell (library quality, diversity). Each dot represents a motoneuronal cell (total of 1536 cells). In total, 999 cells passed the filtering criteria indicated by the dotted lines (see Methods). **(B)** Bar chart represents the number of cells expressing *Hox* genes. 749 cells of 999 cells express at least one *Hox* gene (~75% *Hox* gene coverage).

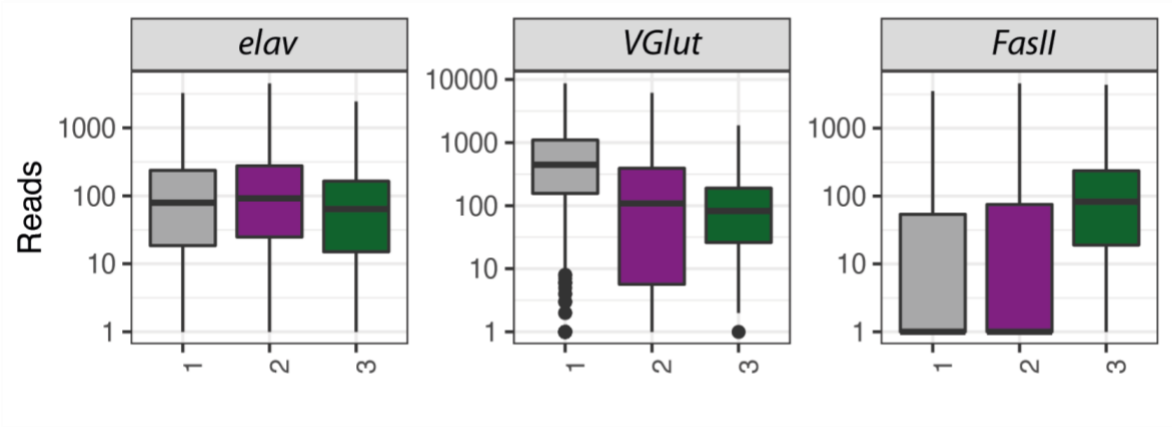


Figure 13.1: Known neuronal markers expressed in MNs

Single motoneuronal cells (columns) are hierarchically clustered using the 20 most variable expressed genes (rows) following methodology of Li et al. (2017). Hierarchical clustering was performed using ward linkage on an euclidean distance metrics. Color code represents gene expression levels (see Methods).

To explore the molecular diversity of MNs, I performed two independent unsupervised analyses, t-distributed neighbor embedding (tSNE) and hierarchical clustering (Figure 13 A, B). Both methods identified three major clusters, corresponding to modulator neurons (VUMs, 8% of the cells) expressing marker genes *Vesicular monoamine transporter (Vmat)* and *Dopamine transporter (DAT)*, and two large yet distinct clusters of MNs that differ in the expression of the marker genes *rhea* and *target of wit* (*twit*) (Figure 13 B, Figure 14). VUM MNs belong to a very distinct motoneuron subtype, the type II glutamatergic/octopaminergic MNs that exhibit modulator roles in taste responses (Matthias Landgraf et al., 1997b; Siegler & Jia, 1999; Helen Sink & Whitington, 1991; Stagg et al., 2011), while the *twit*^{low} and *twit*^{high}

cluster can be assigned to the abundant glutamatergic type I motoneuron class(Hoang & Chiba, 2001; M. D. Kim et al., 2009). In particular, *in situ* hybridizations of late stage embryos localized *twit* transcripts in median and lateral clusters of posterior located MNs (N. C. Kim & Marqués, 2012), suggesting that those two groups of MNs represent indeed two distinct motoneuron subtypes with different locations.

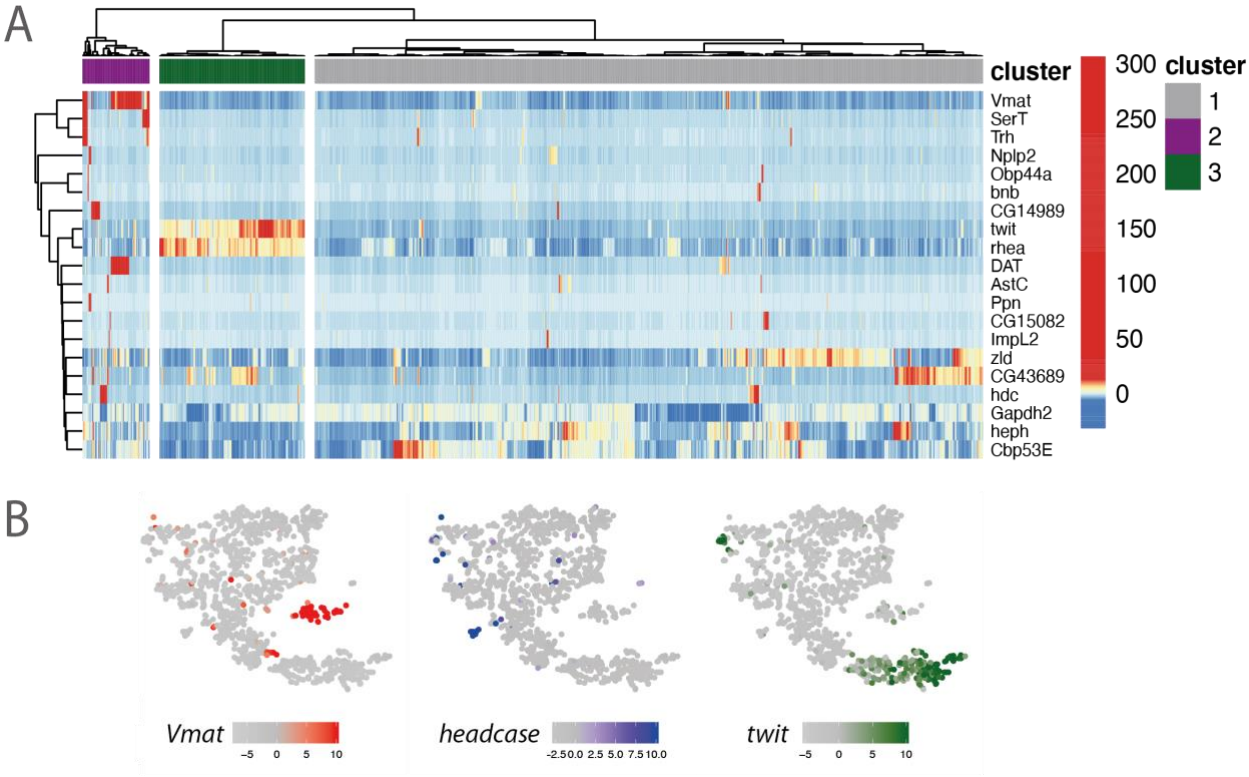


Figure 13.2: Three populations of embryonic *Drosophila* MNs

(A) Single motoneuronal cells (columns) are hierarchically clustered using the 20 most variable expressed genes (rows) following methodology of (H. Li et al., 2017). Hierarchical clustering was performed using ward linkage on an euclidean distance metrics. Color code represents gene expression levels (see Methods). **(B)** Expression level of key marker genes are highlighted on the t-SNE of Figure 1B. **(C)** Single motoneuronal cells (columns) are hierarchically clustered using the 20 most variable expressed genes (rows) following methodology of Li et al. (2017). Hierarchical clustering was performed using ward linkage on an euclidean distance metrics. Color code represents gene expression levels (see Methods).

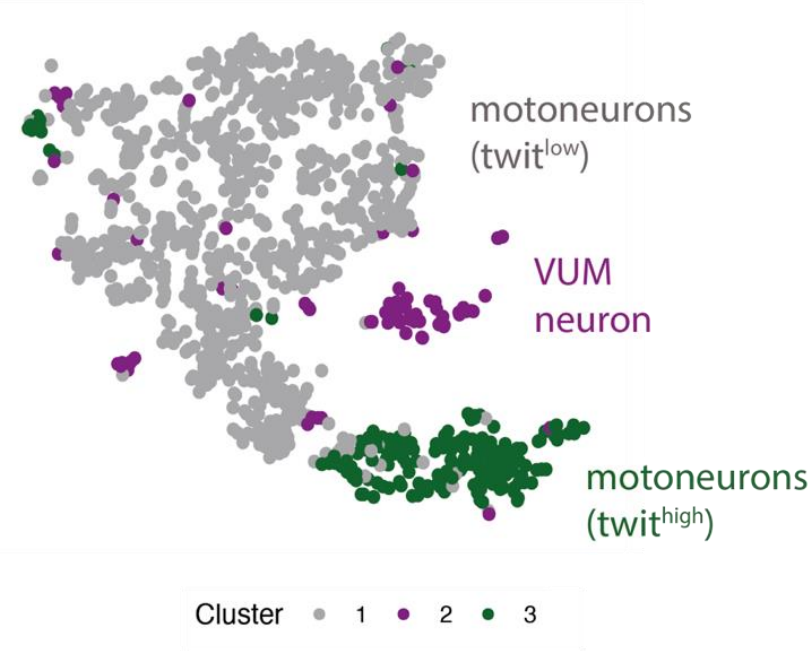


Figure 14: Three major clusters of MNs can be assigned to three different types of MNs

Three major clusters of *Ok371>GFP+* MNs are highlighted on the t-SNE. The color code corresponds to hierarchical clustering using the 20 most variable expressed genes following methodology of Li et al. (2017) (see Figure 13 A) and assignment of the major population of MNs (*Twit*^{low}), the *Twit*^{high} cluster of MNs and modulator MNs (VUMs) has been performed according to marker gene expression (see Figure 13 A).

Although the motoneuronal population in the largest cluster (*twit*^{high}) appears rather homogenous, our data set confirms mutual exclusive expression of marker genes, expressed in subsets of dorsally and ventrally projecting MNs (Garces & Thor, 2006; Kania & Jessell, 2003; M Landgraf et al., 1999) In particular, the dorsal marker gene, *even skipped (eve)* and the ventral marker genes *Nkx6* or *Hb9* are expressed in distinct populations. Likewise, the dorsal marker gene *grainy head (grn)* or *LIM homeobox 1 (Lim1)* and the ventral marker gene *islet (isl)* are expressed in different subsets (Figure 15A). A more detailed analysis was performed to assign cells to their NB origin by clustering on lineage markers (Figure 15B). Although unsupervised methods did not identify these combinations of lineage markers, supervised clustering successfully categorized these cells to lineages.

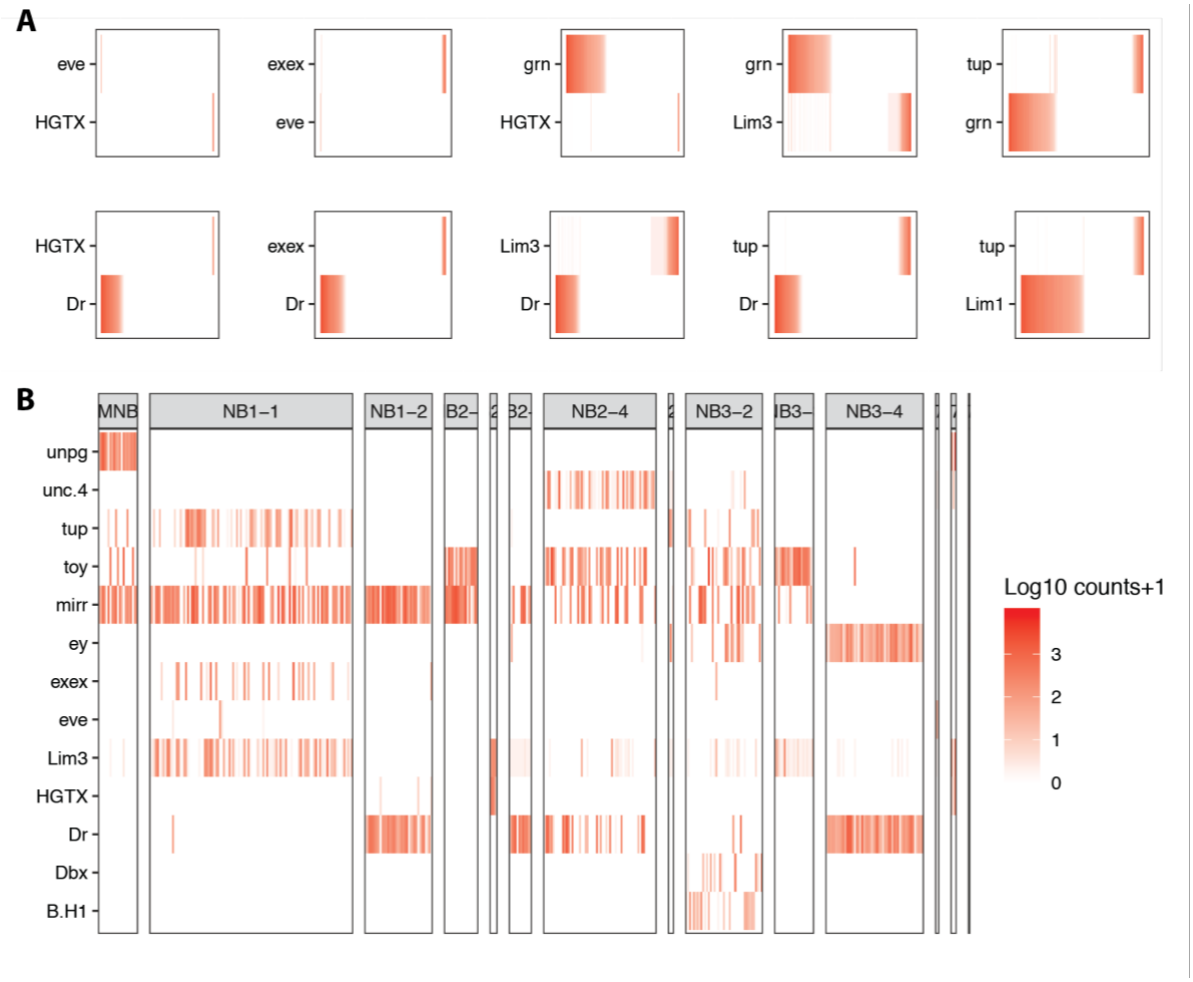


Figure 15: Comparison of single motoneuron transcriptomes (twit^{high} cluster) with with molecular markers described in literature

(A) Expression of gene pairs described in literature to be expressed in a mutually exclusive pattern. Columns correspond to single cells. (B) Expression patterns of NB lineage markers. Cells were assigned to a lineage if at least 2 lineage markers genes (Urbach et al., 2016) were expressed. All other cells are excluded.

Taken together, these analyses showed that the dataset generated in this study was of high quality, faithfully reflected known biology, and provided an approximately 7-fold cellular coverage of each biologically unique motoneuron at the synaptic wiring stage.

2.2 A HOMEO-CODE SPECIFIES SINGLE DROSOPHILA MNs DURING SYNAPTIC WIRING

To investigate processes required for motoneuron diversity during the synaptic wiring phase, I focused on the largest cluster of MNs expressing low levels of *twit* (*twit*^{low}). Our reasoning was that this cluster contains the majority of glutamergic type I MNs that are equally distributed along the VNC, rather than specialized subtypes of MNs such as the VUMs or *twit*^{high} clusters that are unevenly distributed along this embryonic body axis. This is particularly important,

because in this analysis I will focus on processes required for synaptic wiring rather than on molecular programs distinguishing cell subtype identities. As a first step, I used unsupervised identification of highly variable genes and principal component analysis (PCA) of the *twit*^{low} cells and identified variability in several biological processes. While processes related to metabolism and cell cycle dominated the first two principal components (Figure 16.1 A, B), notably genes with high loadings on PC3 were enriched in homeodomain TFs (p=2.0e-13) as well as genes encoding CSPs (p=3.5e-04) involved in synaptic matching and axon guidance (Figure 16.1 C, D). In particular, the expression of known mediators of synaptic specificity, for instance members of the Ig superfamily like Dpr protein encoding genes (Carrillo et al., 2015), was highly variable within the *twit*^{low} cluster, and co-varied with homeodomain TF gene expression (Figure 16.2 A).

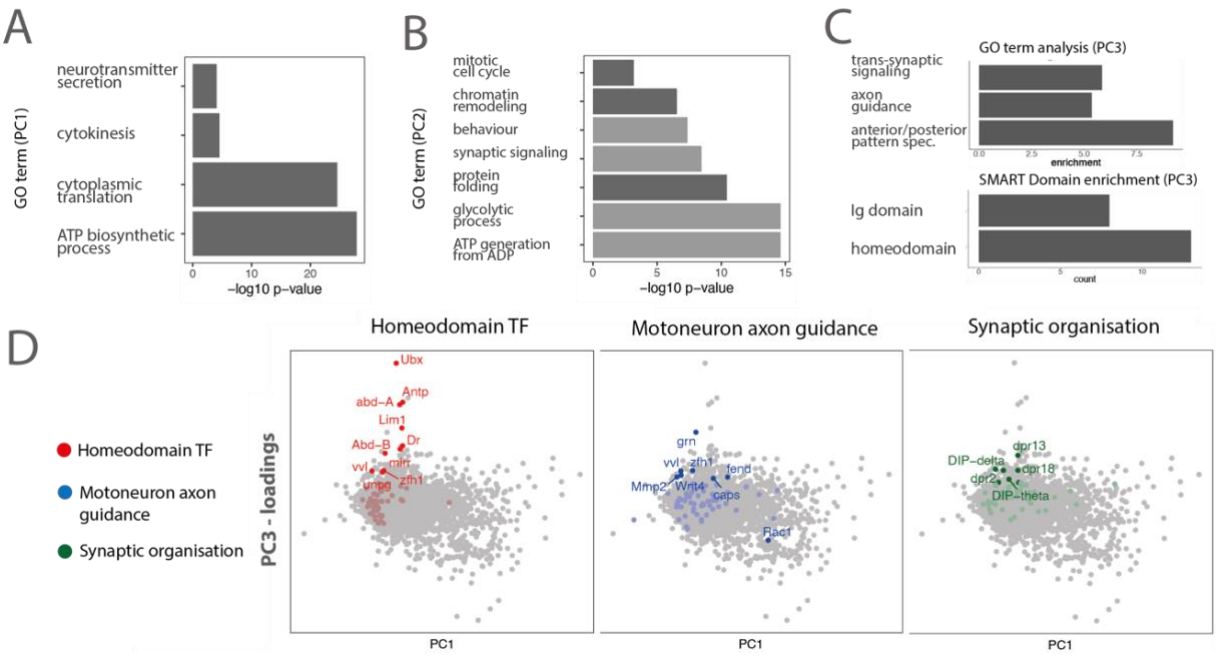


Figure 16.1: Identification of variable processes in embryonic MNs by principal component analysis

(A, B, C) Principal component analysis (PCA) was performed on cells from the *twit*^{low} cluster. GO term analysis for biological processes was performed on the top 10% genes with highest loadings on principal component 1, PC1 (log 10 p-value) (A), principal component 2, PC2 (B) and principal component 3, PC3 (C). GO term and SMART domain analysis was performed on the top 300 genes representing the most enriched candidates in the PCA. Dark grey indicates processes enriched among genes with positive loadings, light grey indicates processes enriched among genes with negative loadings. Together these analyses indicated that PC1 and PC2 are associated with metabolic processes, cellular differentiation and/or technical variation, while PC3 is associated with anterior, posterior patterning processes and synaptic specificity processes. **(D)** Principal component loadings plots highlighting homeodomain TFs (red), genes associated with the GO terms motoneuron axon guidance (blue) and synaptic organisation (green). Points with label correspond to the highest 5% of loadings. Identification of highly variable genes using the method by (Brennecke et al., 2013): homeodomain TFs (p=2.0e-13), Ig (p=3.5e-04).

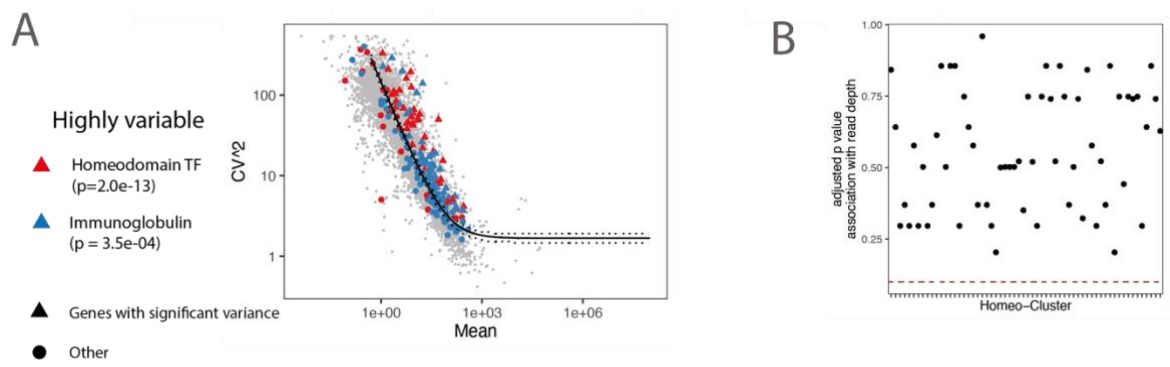


Figure 16.2: Identification of highly variable genes in MNs

(A) Identification of highly variable genes using the method by (Brennecke et al., 2013). Scatter plot depicts for each gene the mean expression and squared coefficient of variation across twil^{ow} cells. The solid line indicates the fit, dashed lines the 95% confidence interval. Genes with a significantly elevated variance are shown as triangles, other genes as circles. Different gene classes are color coded. P values shown are from a hypergeometric test for enrichment of the respective gene class among highly variable genes. **(B)** Scatter plot depicting for each homeo-cluster from main Figure 1C the strength of association with a technical covariate (sequencing depth). P values are from a wilcoxon test contrasting sequencing depth in cells from that cluster, and all other cells. The dotted read line indicates the p value required for significance (0.05). All associations are therefore non-significant.

Using a detailed analysis of homeodomain TF expression, I observed that 60 small groups of cells were defined by a unique expression pattern, an expression 'code' (Figure 17, see supplementary table 2). These patterns were statistically independent of technical covariates such as sequencing depth. In our experimental design, we focused in particular on the subset of type I MNs (*twit*^{low} cluster) whilst each biologically unique cell was covered approximately seven times, raising the possibility that unique homeodomain TF expression patterns correspond to symmetrical pairs of MNs with biologically unique identities (2x60).

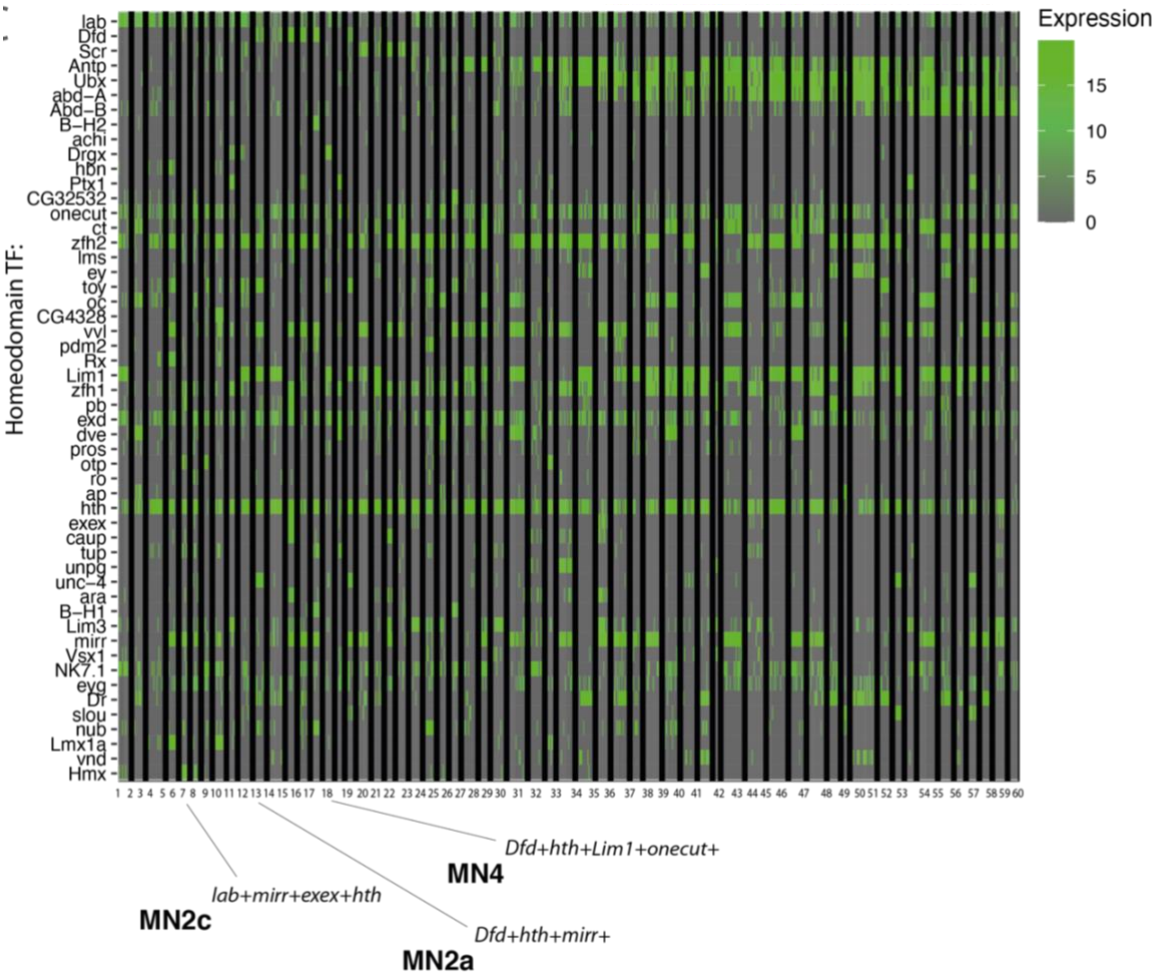


Figure 17: scRNA-Seq identifies homeo-codes as a major determinant of motoneuronal heterogeneity

Heatmap depicting the expression of homeodomain encoding genes (rows) across n= 758 single *twit*^{low} MNs (columns). Rows and columns are arranged by hierarchical clustering. Normalized expression levels are color-coded. Arrows indicate selected clusters for follow up studies shown in Figure 10.

Since a statistic specification of cluster numbers from scRNA-Seq data remains an unresolved issue in the field (Luecken & Theis, 2019; Zhu et al., 2019), I used immunofluorescence stainings in selected examples of homeodomain TFs in MNs to validate the single-cell specificity of the observed code. For example, scRNA-Seq data suggested the existence of a motoneuron that co-expresses *Deformed (Dfd)*, *homothorax (hth)* and *mirror (mirr)* (Figure 17). This neuron should be located in the Dfd+ region (i.e. the maxillary segment of the AP axis) and in the ventral region along the DV axis (Mirr). In fact, the co-expression of these factors was confirmed in a motoneuron, which projects via the maxillary nerve to the MHE muscle and emerges ventrally to the antennal nerve of the maxillary segment (Figure 18 and Figure 19, Figure 8 of introduction). Based on its similar position in *Calliphora vicina*, I termed this axon projection MN2a (Schoofs et al., 2010). A motoneuron lacking the expression of the ventral markers Mirr projects to the dorsal bundle of the MHD in larvae (termed MN4 based on (Schoofs et al., 2010)). Finally, the MN2c motoneuron co-expresses the same set of genes

as MN2a but lacks the segment defining Dfd expression (Figure 18). Thus, it projects anterior to the labial retractor (LR) muscle in larvae.

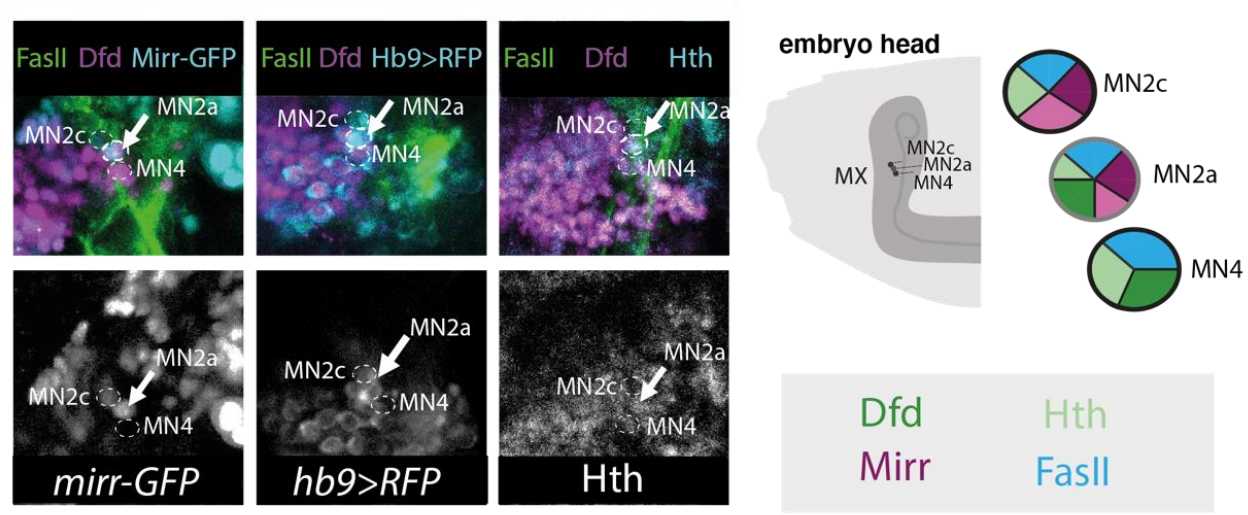


Figure 18: Validation of homeo-code cluster with unique cell identities

Visualization of homeodomain TFs by immunohistochemistry in the head of a stage 17 *Drosophila* embryo, lateral view. Pie charts represent single cells with corresponding homeodomain TF expression (color code). The combinatorial expression of the four homeodomain TFs Dfd, Mirr (*mirr-GFP*), Exex (*Hb9>RFP*) and Hth is exclusive to the MN2a motoneuron (FasII^{+}). The homeodomain TF Dfd is absent in motoneuron A (MN2c), while the homeodomain TFs Mirr and Exex are absent in motoneuron IV.

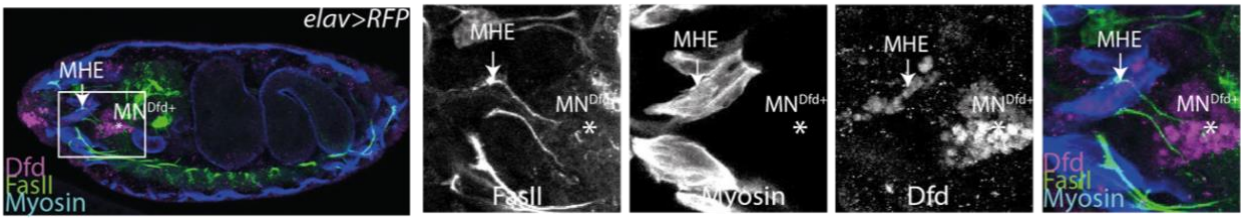


Figure 19: MN2a expressing Dfd projects via the maxillary nerve to the MHE muscle

Confocal image analysis of the lateral view of early stage 17 *Drosophila* embryos to visualize Dfd expression (magenta) in FasII-positive MNs (green) and Myosin-positive muscles (blue) for the identification of the exact target destination of the Dfd^{+} MN2a motoneuron. The MN2a axon (asterisk, Figure 17) projects to the MHE (arrow).

As further examples of homeodomain co-expression patterns, I confirmed the existence of a single motoneuron originating from the A2 segment with high levels of Ultrabithorax (Ubx) and intermediate (mid) levels of Mirr, as well as a single motoneuron with high levels of Mirr and intermediate (mid) levels of Ubx originating in the A3 segment (Figure 20). In total I confirmed the single-cell specific homeodomain TF expression patterns of five biologically unique MNs.

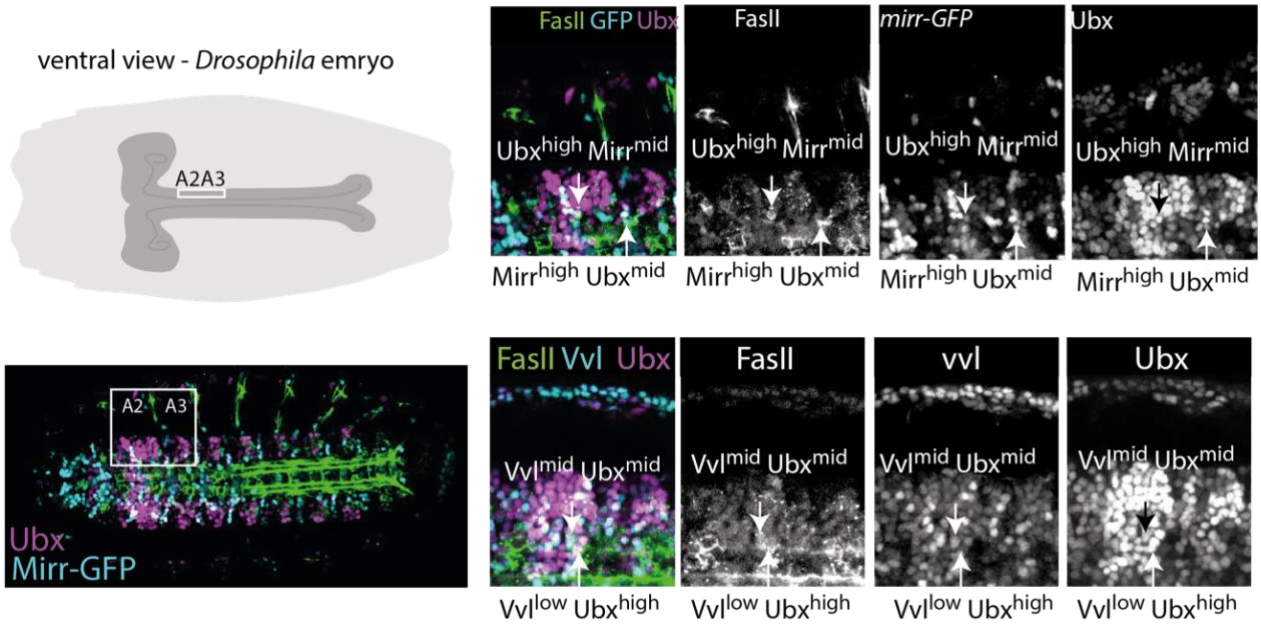


Figure 20: Validation of the homeo-code in abdominal segments

Left panel: Schematic drawing and immunofluorescence image depict the ventral view on an early stage 17 *Drosophila* embryo and highlight the A2 and A3 segment in the VNC. *Right panel:* Confocal images depict the ventral view of early stage 17 *Drosophila* embryos, zooms on the abdominal A2 and A3 segments and highlights the combinatorial expression of homeodomain TFs (Mirr, Vvl and Ubx) expression in individual cells (white arrows). MNs are labelled by FasII (green).

In sum, these analyses revealed a homeodomain TF 'code' as unique identifier of cellular identity during the synaptic wiring phase.

2.3 REGIONAL SPECIFIC EXPRESSION OF THE HOMEO-CODE

The identification of a homeo-code raised the possibility that cells possess a molecular memory of their spatial position that later coordinates synaptic wiring. To investigate this hypothesis, I mapped single MNs in space. For mapping neurons along the AP axis, I created a high-resolution map of Hox protein expression using immunofluorescence (Figure 21A, B).

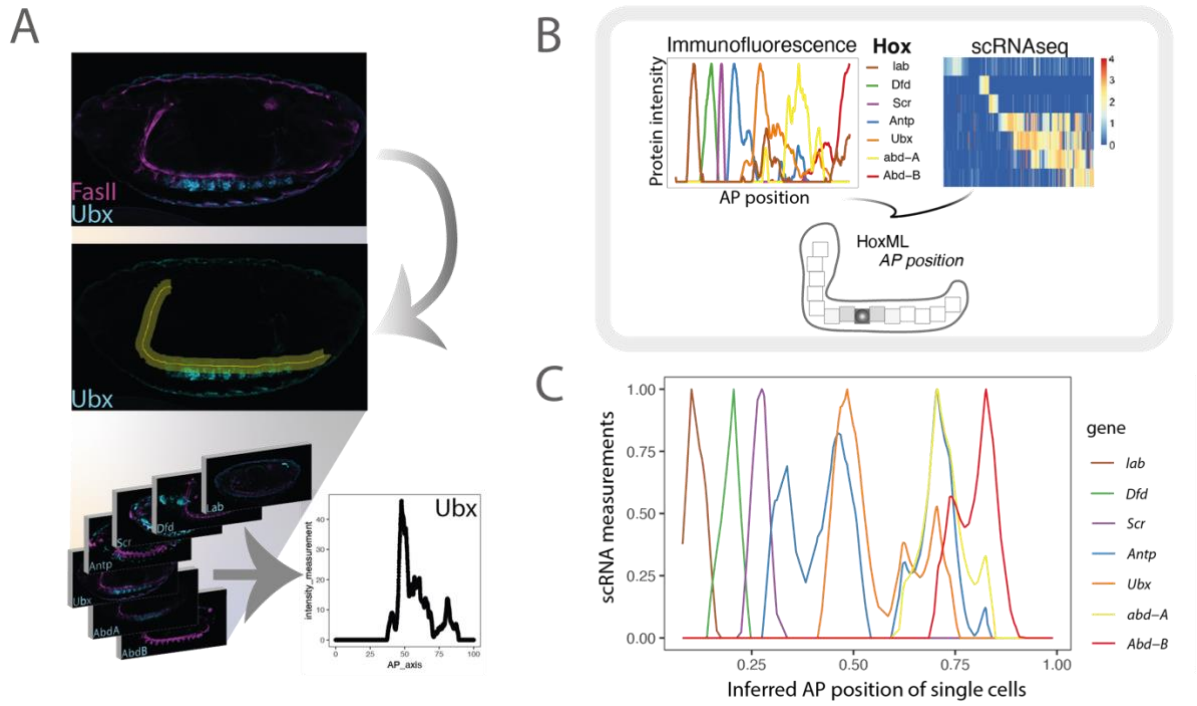


Figure 21: Spatial reconstruction of AP axis based on Immunofluorescence measurements of Hox proteins

(A) Pipeline for protein intensity measurements of the seven Hox TFs (from anterior to posterior: Lab, Dfd, Scr, Antp, Ubx, AbdA and AbdB) expressed along the AP axis of the VNC. *Upper panel:* The motoneuronal marker (FasII) in magenta was used as reference to measure Hox TF expression patterns in different embryos in a standardized manner (see Materials and Methods). *Lower panel:* Procedure of translating different fluorescence intensity measurements into standardized graphs by Fiji. **(B)** Outline of strategy to map single MNs to position along the AP axis. *Left panel:* Intensities of Hox protein expression along the VNC measured by immunofluorescence, see also Figure S3A, S3B. *Right panel:* Co-expression pattern of *Hox* gene transcripts measured by scRNA-Seq, see also Figure S3C. Expression level is color-coded; columns correspond to n=758 single *twit*^{low} MNs. *Bottom panel:* AP position is inferred form scRNA-Seq data by probabilistically mapping *Hox* gene expression pattern in each individual cell to the immunofluorescence reference data. See Materials and Methods. **(C)** Normalized and smoothened single cell mRNA measurements of *Hox* gene expression (color code) arranged along the inferred AP position (See Figure 22, Materials and Methods).

Anterior Hox proteins were expressed in clearly defined stripes, whereas posterior Hox protein expression patterns partly overlapped. The same co-expression patterns were observed in our scRNA-Seq data (Figure 21), allowing us to probabilistically map each cell from the scRNA-Seq data set to a position along the AP axis (Figure 21C and 22, Materials and Methods).

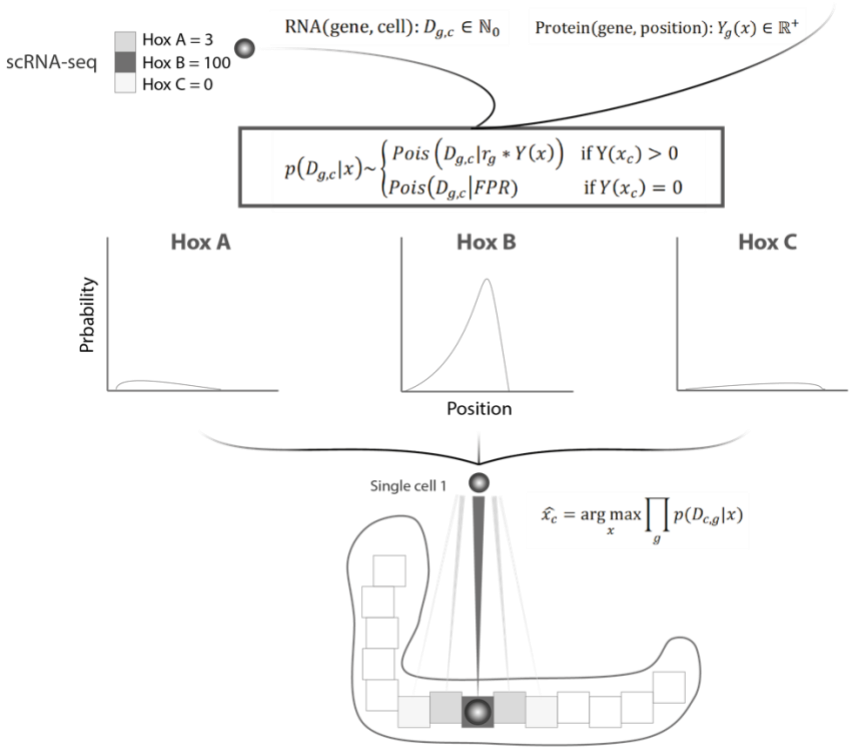


Figure 22: Computational modelling of AP position

Immunoflourescence data for each Hox gene (g) was calculated as a function of position (Y_g (X) € (0,1)) along the animal AP axis and integrated with single cell gene expression data across each gene and cell (Dg,c € N_0). Then the probabilistic position of each cell was inferred based on gene expression profiles (see Material and Methods).

I validate this mapping strategy by immunofluorescence. To this end, I used the inferred AP position to estimate the expression pattern of every gene along the AP axis. I thereby identified candidate genes with differential expression along this axis.

Importantly, these candidates were not used for constructing the model. For two such candidates, *Frq1* and *hth*, I compared the predicted expression pattern to immunofluorescence data and observed a high agreement (Figure 23).

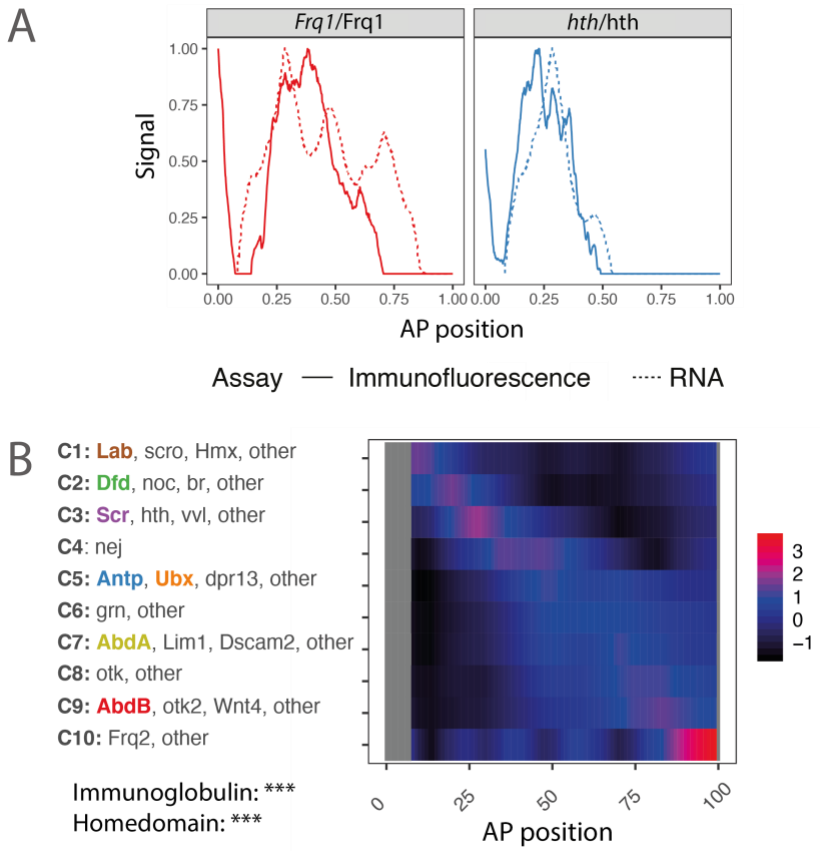


Figure 23: Unsupervised analysis of scRNA-Seq data identifies homeodomain gene expression patterns along the AP axis

(A) Genes with significant variation along the AP axis were identified and clustered into 10 groups of distinct expression pattern (Materials and Methods). Heatmap shows average gene expression per cluster (rows) across single cells (columns). See Supplementary Table 1 for a complete list of genes in each cluster. Asterisks indicate p-value of a hypergeometric test for enrichment of protein domains, ***: $p < 0.001$. (B) Genes with significant variation along the AP axis were identified and clustered into 10 groups of distinct expression pattern (Methods). Heatmap shows average gene expression per cluster (rows) across single cells (columns). See Supplementary Table 1 for a complete list of genes in each cluster. Asterisk indicate p-value from a hypergeometric test for enrichment of protein domains, ***: $p < 0.001$.

Inferred AP position was highly correlated with principal component 3 and 4, indicating that AP position profoundly affects the entire transcriptome of the cell. More importantly, the above described homeo-code is aligned with the AP position, but on the fine scale the homeo-code shows more variability independent of AP position (Figure 24).

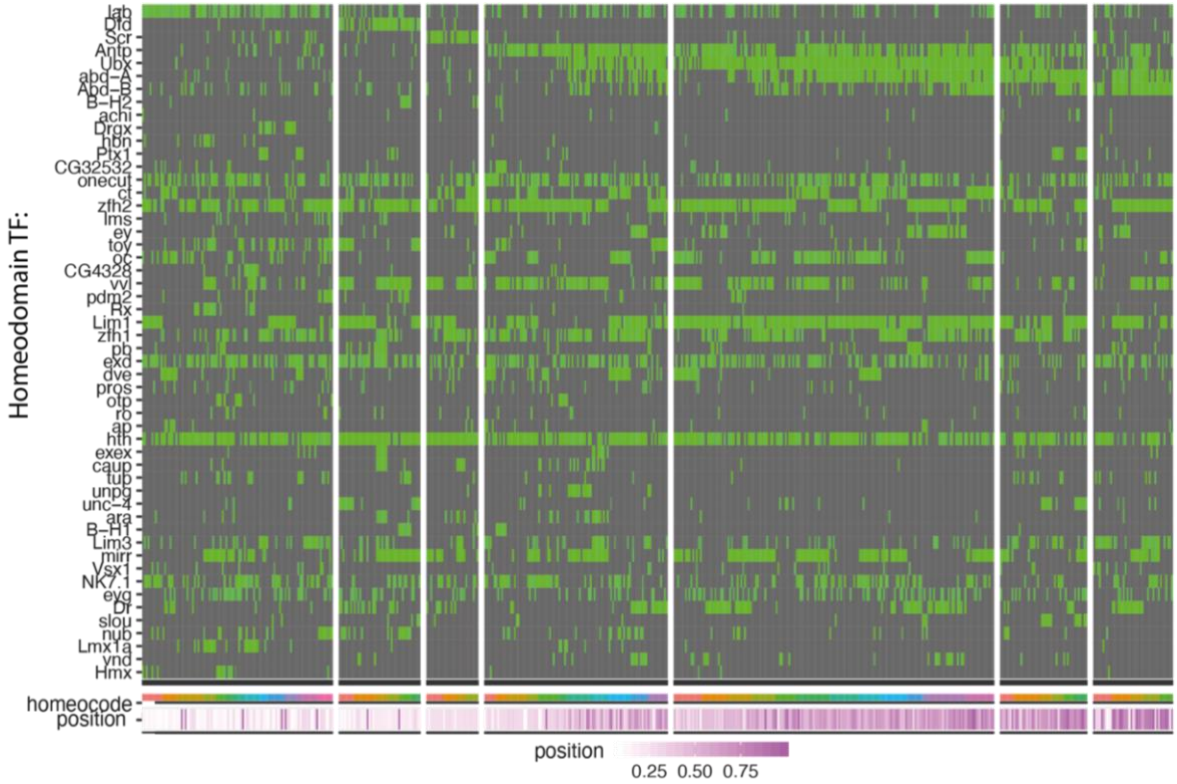


Figure 24: Analysis of scRNA-Seq data identifies variability of the homeo-code independent of AP position

Heatmap depicting the expression of homeodomain encoding genes (rows) across n= 758 single *twit*^{low} MNs (columns). Normalized expression levels are color-coded in green. Rows and columns are arranged by *Hox* gene expression from anterior (A) to posterior (P) (see Figure 21), inferred position is color coded in magenta. Multi-color code indicates individual homeo-clusters as defined in Figure 17.

In order to identify these additional processes, I used ZINB-WaVE (Risso et al., 2019) to separate scRNA-Seq data into variability linked to the known covariates (AP position and technical variability), and processes statistically independent thereof (Figure 25 A). Thereby, I identified on the first component of AP-independent variability one group of genes known as marker for DV position (Bhat, 1999; J B Skeath, 1999; Urbach et al., 2006). These genes were ordered according to their localization in the embryo from dorsal to ventral (Figure 25 B). Immunofluorescence experiments of two ventral marker genes, *mirror (mirr)* and *ventral veins lacking (vvl)*, confirmed the predicted DV position (Figure 25 C-F) as example in two different compartments of the *Drosophila* embryo, in the maxillary segment (Figure 25 C, D) and in abdominal segments (Figure 25 E, F).

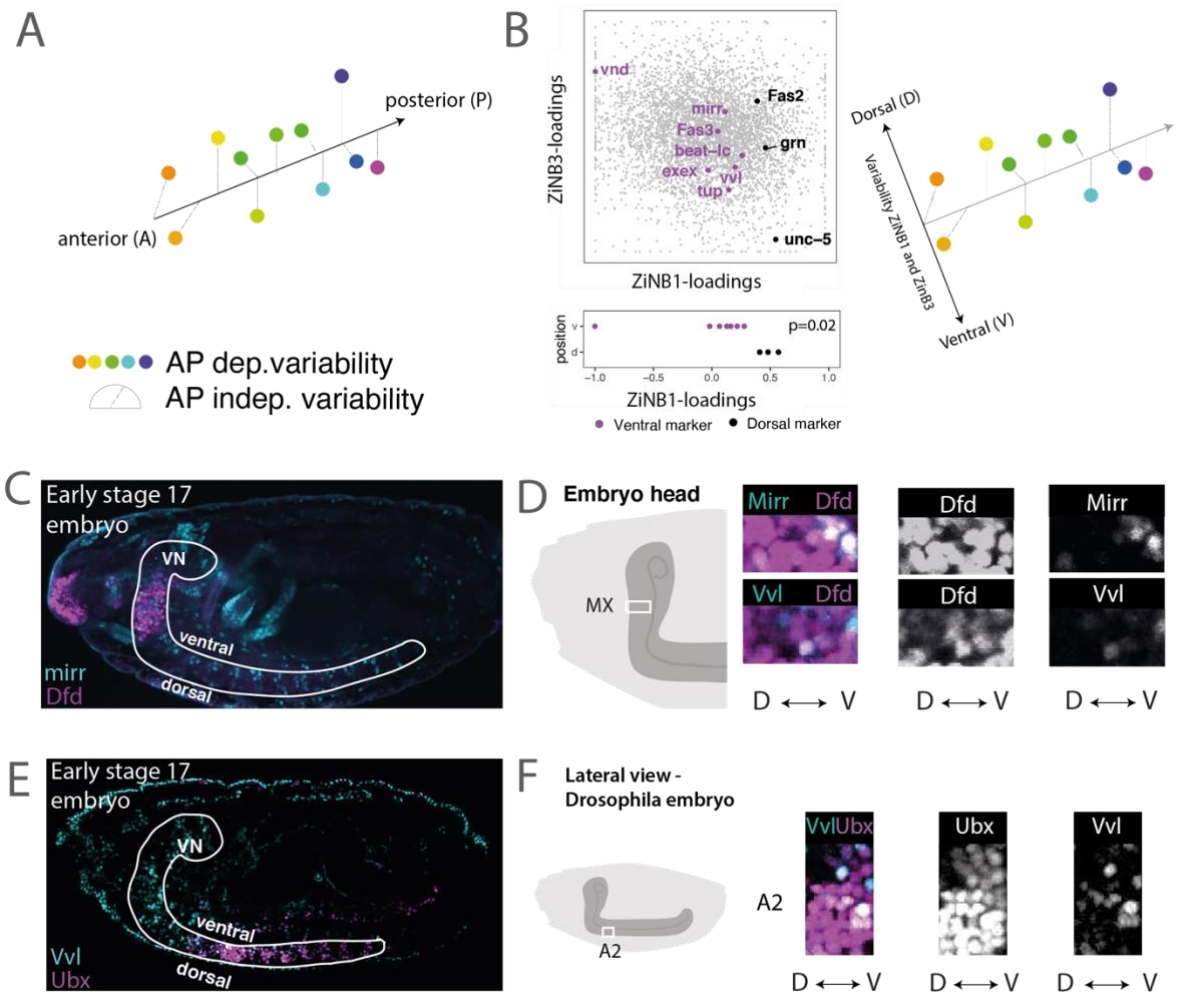


Figure 25: Unsupervised analysis of scRNA-Seq data identifies heterogeneity depending on the DV-axis

(A) ZiNB-WaVE (Risso et al., 2019) was used to statistically separate gene expression variability into parts linked to AP position and parts independent thereof. **(B)** Scatter plot of ZINB-WaVE loadings separates known dorsal and ventral marker genes on ZiNB-WAVE component 1 **(C-F)** Representative confocal picture of an early stage 17 *Drosophila* embryo, lateral view. Dorsal and ventral site of elongated VNC are highlighted. Visualization of two ventral homeodomain TF markers (Mirr, Vvl) in the maxillary segment (MX).

Again, homeodomain TFs and Ig surface proteins were among the most variable genes on this AP-independent axis of variability (Figure 26), suggesting that distinct expression patterns of homeodomain TFs and Ig domain molecules correlating with position (DV and AP axis) are diversifying MNs most during synaptic wiring.

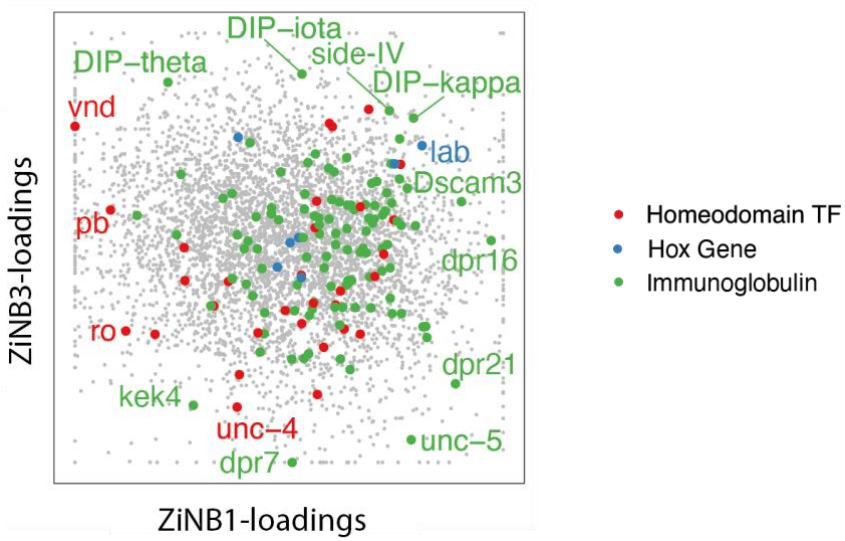


Figure 26: Unsupervised analysis of scRNA-Seq data identifies heterogeneity of homeodomain gene expression patterns independent of AP axis

Genes encoding homeodomain TFs and genes encoding Ig domain molecules (see color code) show high loadings on ZINB-WaVE component 1 and 3, demonstrating high variability independent of AP position.

Following upon the finding that distinct homeo-codes are expressed in MNs with individual cell identities (Figure 18), these analyses suggested that the position along the animal body axis determines the homeo-code and thereby defines cellular identity.

2.4 HOMEO-TFs MODULATE SYNAPTIC TARGET SPECIFICITY IN A POSITION DEPENDENT MANNER

Are homeodomain TFs functionally linked to processes mediating synaptic specificity? It is described that this class of TFs is involved in maintaining the neuronal lineage (Deneris & Hobert, 2014; Domsch et al., 2019; Friedrich et al., 2016; Reilly et al., 2020). However, it is so far unclear whether sets of homeodomain TFs instruct specific programs for precise matching of MNs and their targets. To address this question, I interfered with *Dfd*, *mirr* and *hth* by RNAi using the pan-neural driver *elav*-GAL4 (L. Luo et al., 1994) and examined synaptic defects in early stage 17 *Drosophila* embryos. Despite the activity of the driver in NB stages, RNAi mediated knockdown is realized earliest in stage 13-14 of motoneuronal axonogenesis (Figure 27). Thus, the driver is ideally suited for temporal interference of proteins during the synaptic wiring phase.

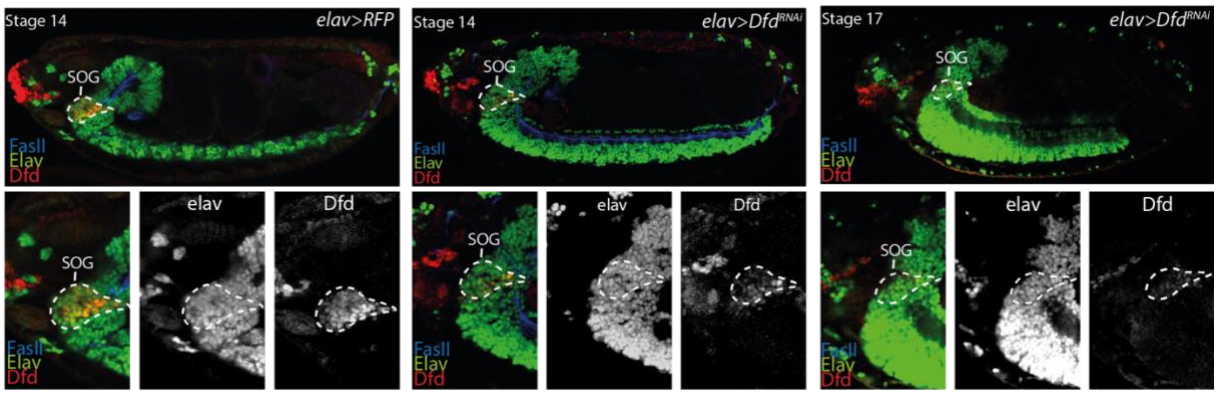


Figure 27: Characterisation of the Dfd RNAi line.

Left panel: Representative confocal image of a stage 14 *Drosophila* embryo, which depicts the lateral view of the Dfd expression pattern in the SOG (white circle) in *elav>RFP* control animals. *Middle panel:* Representative confocal picture of a stage 14 embryos under *Dfd* knockdown conditions (*elav>Dfd*^{RNAi}), showing a strong reduction of Dfd neuronal expression at that stage. *Right panel:* Representative confocal image of a stage 17 embryo displaying full depletion of *Dfd* under knockdown conditions in neurons.

To complement this approach, I used in addition a more specific driver, *OK6*-GAL4, which is active in glutamatergic MNs only (Sanyal, 2009).

Defects induced by RNAi mediated gene silencing were classified into two distinct categories, wiring defects (i.e. mistargeting of neurons to the wrong muscle), and terminal defects (i.e. abnormal synaptic morphologies at axon terminal sites). I focused my analyses on the MN2a motoneuron (Figure 28 A), since it can be easily identified by the co-expression of Dfd and FasII. MN2a innervates the MHE muscle, an interior muscle of the *Drosophila* feeding apparatus that also controls hatching movements (Figure 28.1 A, Figure 8 of introduction) (Friedrich et al., 2016). Thus, an assay of hatching rates is a relatively simple readout to examine the impact of wiring defects on behaviour. Embryos depleted for *Dfd*, *mirr* and *hth* were all characterized by significant morphological changes of the MN2a motoneuron (Figure 28.1 B, C, F). 70% of *hth* depleted embryos displayed severe wiring defects, including mistargeting of the MN2a to ventral and anterior muscles of the outer skeleton of *Drosophila* embryos (Figure 28.1 F), while 29% of the *Dfd* depleted embryos showed mistargeting of MN2a to the LR muscle located directly anterior to the MHE in the Lab^{+}, Dfd^{-} segment (Figure 28.1 F, Figure 8 of introduction), leading to a decreased hatching rate (Figure 28.2 A). Survivors of the hatching rate assay showed proper targeting of MN2a to the MHE, but abnormal synaptic morphologies, such as an increased number of synapses (Figure 28.2 B). In the case of *mirr* knockdown, wiring effects were less pronounced and frequent. Together, these results suggested that loss of homeodomain TF expression leads to synaptic targeting defects.

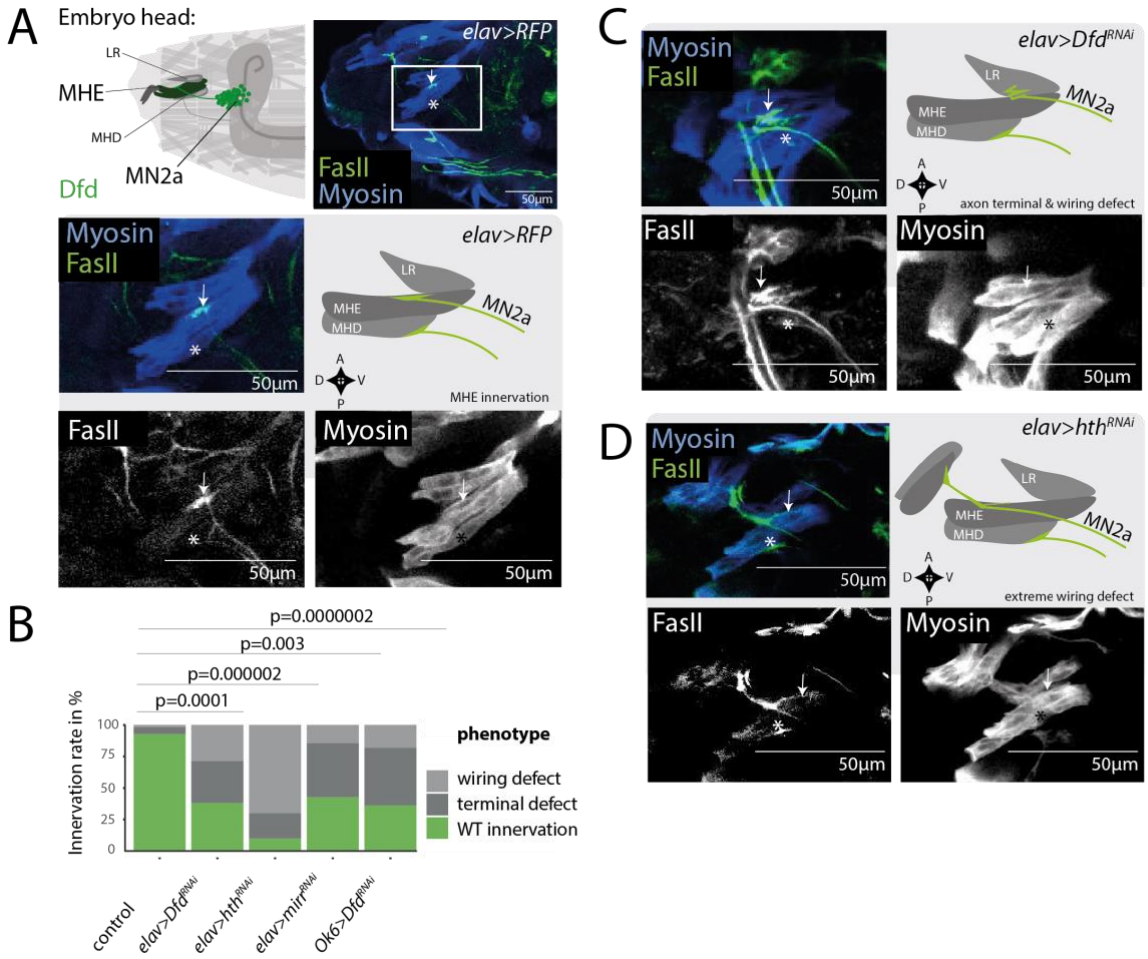


Figure 28.1: Homeodomain TFs modulate target specificity of MNs

(A) *Left panel:* Schematic drawing depicts lateral view of an early stage 17 *Drosophila* embryonic head. The MN2a motoneuron expressing Dfd (light green) innervates the MHE (dark green). The MHE controls elevation movements of mouth hooks (dark grey) that are required for hatching from the eggshell (Friedrich et al., 2016). *Right panel:* Representative confocal picture of an early stage 17 *Drosophila* embryonic head depicts the innervation site (white arrow) on the MHE muscle (blue) by MN2a (green). The innervation site on the MHD muscle (blue) is highlighted by an asterisk as reference for orientation in the embryo. **(B-D)** Visualization of three independent genetic experiments using the pan-neural *elav*-GAL4 driver to control the expression of UAS-*RNAi* constructs of homeodomain TFs i. e. UAS-Dfd^{RNAi}, UAS-hth^{RNAi} and UAS-*RFP* as control. Zoom on MHE of an early stage 17 *Drosophila* embryo. *Right panels:* Schematic drawing depicts the innervation phenotype on MHE by the MN2a motoneuron. Phenotypes are classified for abnormal synaptic morphologies at axon terminal sites: 'terminal defects` and incorrect innervation of target muscle 'wiring defects'. In wild-type conditions, the MN2a motoneuron projects to the MHE, while in *elav>*Dfd^{RNAi} animals MN2a MNs display axon 'terminal defects' and 'wiring defects', i.e. incorrect targeting to the LR muscle and in *elav>*hth^{RNAi} animals incorrect targeting of an outer skeleton muscle, termed 'extreme wiring defects'. **(B)** Phenotypic penetrance for MN2a projections 'innervation rate' was calculated by the rate of correct axon projections of MN2a to MHE (mock; n=56) compared to abnormal synaptic morphologies at axon terminals 'terminal defects' and wiring defects displayed by animals expressing RNAi constructs for homeodomain TFs i. e. *OK6>*Dfd^{RNAi} (n=11), *elav>*Dfd^{RNAi} (n=21), *elav>*$mirr^{RNAi}$ (n=7) and *elav>*hth^{RNAi} (n=10). Note, each genetic experiment was performed in parallel to an adequate control experiment using the same driver line crossed to a line that controls expression of either UAS-*RFP* or UAS-GFP^{RNAi}. For quantifications these control experiments were summarized and termed 'mock' experiments. p-values between two genetic conditions were calculated by a two-sided hypergeometric test.

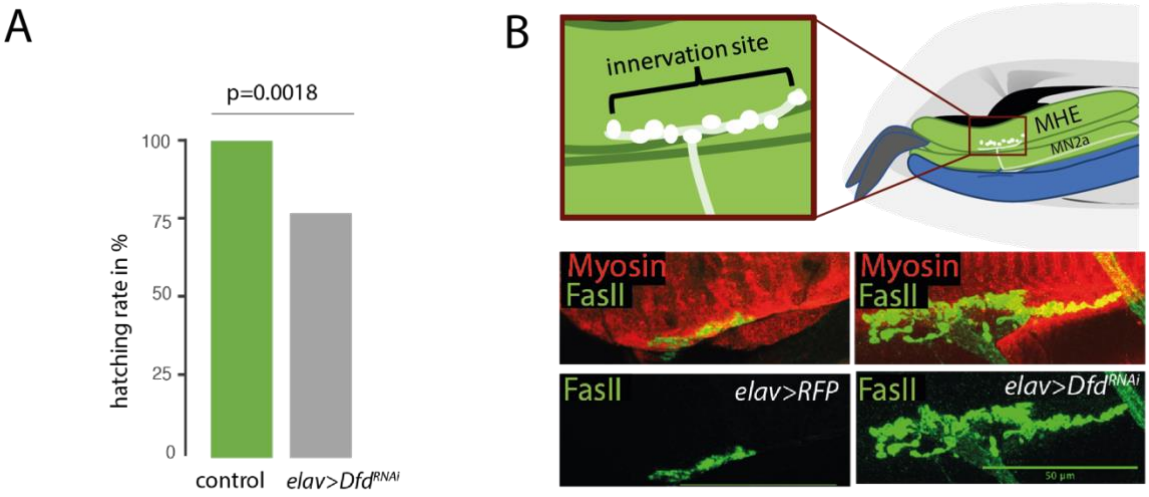


Figure 28.2: Loss of homeodomain TFs in neurons decreases hatching movements and leads to aberrant synapsis phenotypes in survivors

(A) Correct MHE innervation is required for hatching of *Drosophila* embryos from the eggshell (Friedrich et al., 2016). Hatching rate was calculated based on the number of L1 larvae observed after 24h in genetic crosses, depleted of Dfd (UAS-*Dfd*^{RNAi}) in neurons (*elav*-GAL4; n=217) compared to crosses with control animals (mock = *elav>RFP*; n=156). (B) Representative confocal picture depicts the innervation site on the MHE muscle in L3 larvae. Animals depleted of Dfd in neurons (*elav>Dfd*^{RNAi}) show enlarged innervation sites and increased amounts of boutons and branching phenotypes (right panel) compared to control animals (*elav>RFP*), (left panel).

The hypothesis of a homeodomain 'code' implies that not only the loss of factors but also their ectopic expression should cause wiring defects. I therefore ectopically expressed the *Hox* genes *labial* (*lab)* as well as *Ubx* in all neurons using the *elav*-GAL4 driver (Figure 29 A-C). Lab is expressed directly anterior to Dfd, while Ubx is active in abdominal segments. Pan-neural and motoneuron-specific misexpression of *lab* led to mistargeting of MN2a to the LR muscle (Figure 29 A), reminiscent of the mistargeting phenotype observed in *Dfd* knockdown conditions (Figure 28 C). By contrast, *Ubx* mis-expression typically elicited terminal defects but not a wiring phenotype (Figure 29 B).

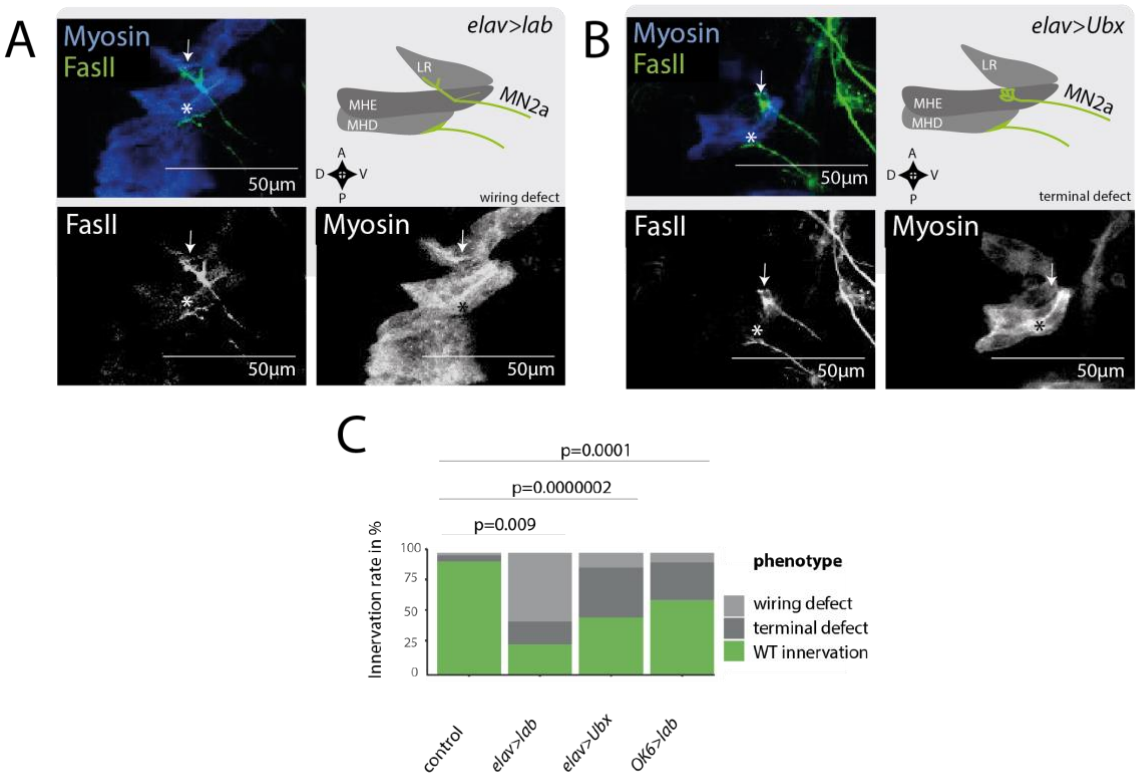


Figure 29: Homeodomain TFs modulate target specificity of MNs

(A-B) *Right panel:* Representative confocal images of two independent genetic experiments using the pan-neural *elav*-GAL4 driver to control the expression of the homeodomain TFs Ubx and Lab i. e. UAS-*Ubx*, UAS-*lab*. The control experiment (*elav>RFP*) is depicted in Figure 29A. *Left panel:* Zoom on MHE of an early stage 17 *Drosophila* embryo. *Right panel:* Schematic drawing depicts the innervation phenotype on MHE by the MN2a motoneuron. See Figure 28 B for classification of phenotypes. In wild-type conditions the MN2a motoneuron projects to the MHE, while in *elav>lab* animals MN2a MNs display 'wiring defects', i.e. incorrect targeting to the LR muscle and in *elav>Ubx* animals show aberrant morphologies at axon terminals, termed 'terminal defects' **(C)** Innervation rates of *OK6>lab* (n=13), *elav>lab* (n=16) and *elav>Ubx* (n=17) animals are calculated as described in Figure 28 B.

In sum, stage specific genetic interference with homeodomain TFs resulted in changes in target specificity of MNs, supporting the idea that these factors modulate synaptic specificity in late embryonic stages when synaptic connections are established.

2.5 HOMEODOMAIN EFFECTS ON TARGET SPECIFICITY ARE MEDIATED BY COMBINATORIAL IG EXPRESSION

Our single-cell analysis and functional follow-ups revealed a critical role for homeodomain TFs in controlling synaptic specificity in the neuromuscular system, motivating an investigation of potential downstream effectors. Unsupervised analysis of gene classes associated with the homeo-code revealed that Ig encoding genes most strongly correlated with homeodomain clusters (Figure 30, Figure 31).

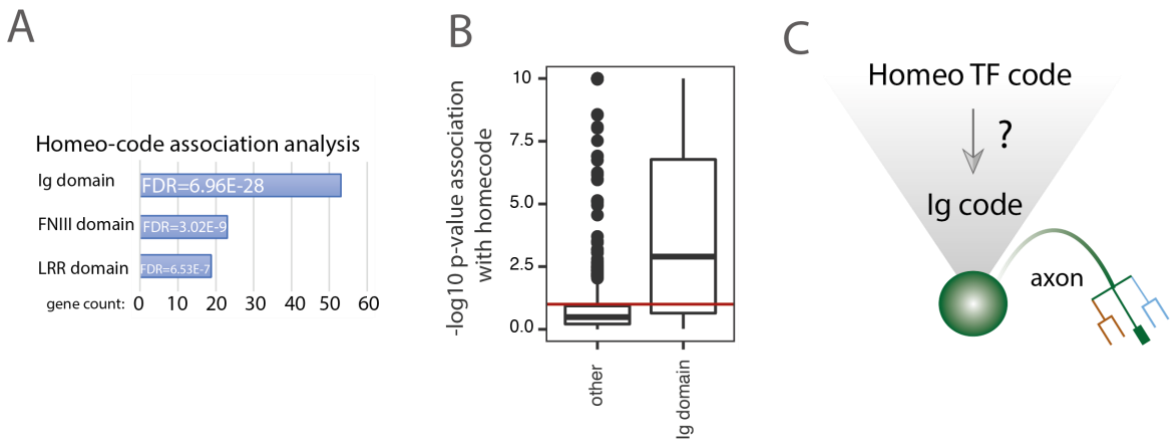


Figure 30: Association of homeo-code with Ig-code

(A, B) Investigation of gene classes associated with the homeo-code. For each gene in the dataset, normalised, scaled expression in single cells was modelled as a function of homeo-code cluster identity, and significance of the association was determined using an F-test. Panel (**A)** depicts the number of genes with significant homeo-code association falling into distinct gene classes. P values are from a Fisher test. Panel **(B)** contrasts the P values for homeo-code association between Ig domain genes and other genes using a boxplot. **(C)** Schematic drawing of the relationship between the homeo-code and the Ig-code in an individual neuron.

Thus, our data set gives us a comprehensive overview about the distribution of Ig encoding genes across different cellular identities. For example, segment specific expression of Ig encoding genes is observed for *off-track 2 (otk2)* that is enriched in posterior segments. Some other Ig encoding genes, such as genes of the *Down syndrome cell adhesion molecule (Dscam)* or *kekkon (kek)* gene family are expressed in most homeodomain clusters, but are occasionally switched off. In other cases, such as genes of the *DIP* family, the *Ig* gene is specifically expressed in one or few homeodomain clusters, which I thus term “ON code”. More detailed analysis allowed us to predict specific Ig-codes (Figure 31 A, B, see supplementary table 3) for every homeodomain cluster. This code is the most specific Ig-code observed in our data set. Together these results indicate that homeodomain TFs might play a direct role regulation of *Ig* gene expression.

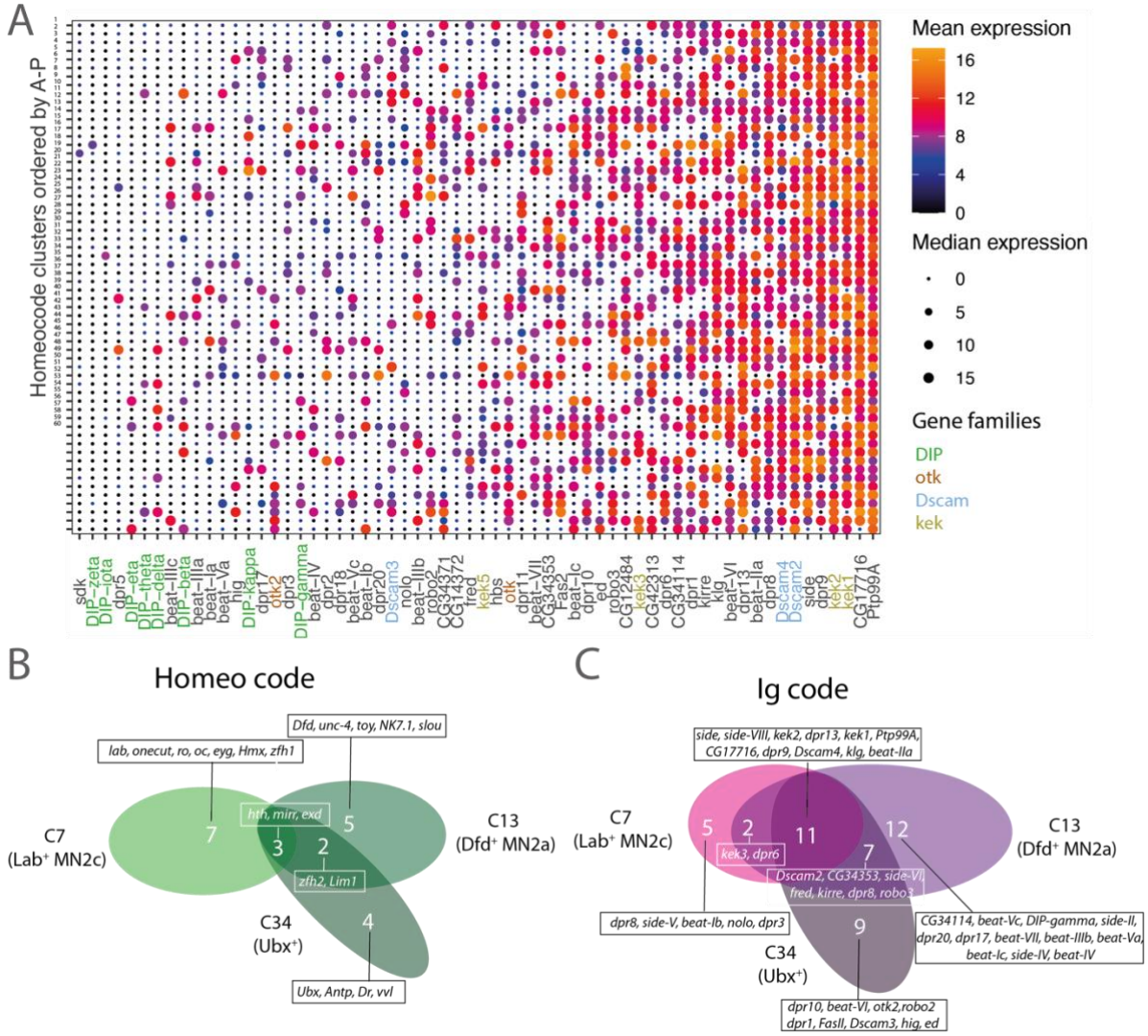
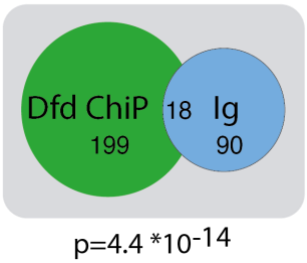


Figure 31: Homeo-codes are linked to Ig expression

(A) Single cells were grouped into 60 clusters according to their expression of genes encoding homeodomain proteins (see Figure 1C) and arranged along the inferred AP position (see Figure 2A). For each cluster, the mean expression of Ig encoding genes was computed. Heatmap depicts Ig encoding genes in columns and clusters in rows; mean expression is color-coded and median expression level is visualized by circle size. Gene families are highlighted in color codes. **(B-C)** Venn diagram displays the homeo-code (E) or Ig-code (F) in three different Homeodomain clusters (see cluster analysis Figure 17) each of the homeo-clusters is associated with a different Hox gene.

To provide further evidence for the hypothesis that an Ig-code is controlled by unique combinations of homeodomain TFs, I analysed previously generated Dfd ChiP-Seq data (Sorge et al., 2012) and found that Ig encoding genes were enriched among the Dfd bound targets (Figure 32 A). In addition, homeodomain TF binding sites were overrepresented within regulatory sequences associated with *Ig* genes (Figure 32 B), strongly supporting a direct regulation of *Ig* expression by homeodomain TF.

A



B

NES	Name	predicted TF's	Ig targets	Motif
3.268	yetfasco-204	pdm2, nub	61 (108)	
3.165	encode-UW.Motif	Scr, Dfd, ftz, XRCC1	64 (108)	
3.089	homer-M00659	vvl	60 (108)	

Figure 32: Homeodomain TFs regulate Ig expression

(A) Venn diagram displaying the relationship between genes expressed in MNs, which are bound by Dfd in a whole embryo ChIP (Sorge et al., 2012) and Ig genes. p-value was calculated using a hypergeometric test, n=307 genes. **(B)** iRegulon analysis was used to identify TF motifs enriched in the vicinity of Ig encoding genes (see Materials and Methods). 3 of the top 15 highest- ranked motifs of TFs are shown; the predicted targets of these motifs are homeodomain TFs.

To investigate the functional role of Ig domain proteins, I focused on the $Dfd^{+}Hth^{+}Mirr^{+}$ MN2a motoneuron, which highly expressed the Ig domain proteins DIP-k and DIP-y (Figure 33 A-B). RNAi-mediated knockdown of *DIP-k* in neuronal cells caused both mistargeting of the LR muscle (6%) and terminal defects (41%) with an increased number of synapses formed between MN2a and the target muscle (Figures 33A, 33B, 33C, 33E), similar to the phenotype observed in *Dfd* knockdown embryos (Figures 33B, 33E). In addition, I analysed the hatching behaviour of embryos depleted for *DIP-k* or *DIP-y* and found that their hatching rates were decreased by 93% and 26% respectively when compared to control animals (Figures 33F). This result indicated that the observed wiring defects are sufficient to impair motor behaviour. Similar defects were not observed with embryos depleted for the Ig domain encoding genes *tutl* or *DIP-alpha* that do not correlate with *Dfd* expression. Next, I examined the effects of *Ig* gene mis-expression in neurons on synaptic target selection. A member of the Dpr network, *Dpr1*, is normally not expressed in MN2a MNs according to our single cell data (Supplementary Figures 33E). Interestingly, misexpression of *Dpr1* in Mn2a resulted in both terminal defects (22%) and wiring defects (22%) such as ectopic targeting of the LR muscle, indicating an altered preference in target choice (Figures 33D, 33E).

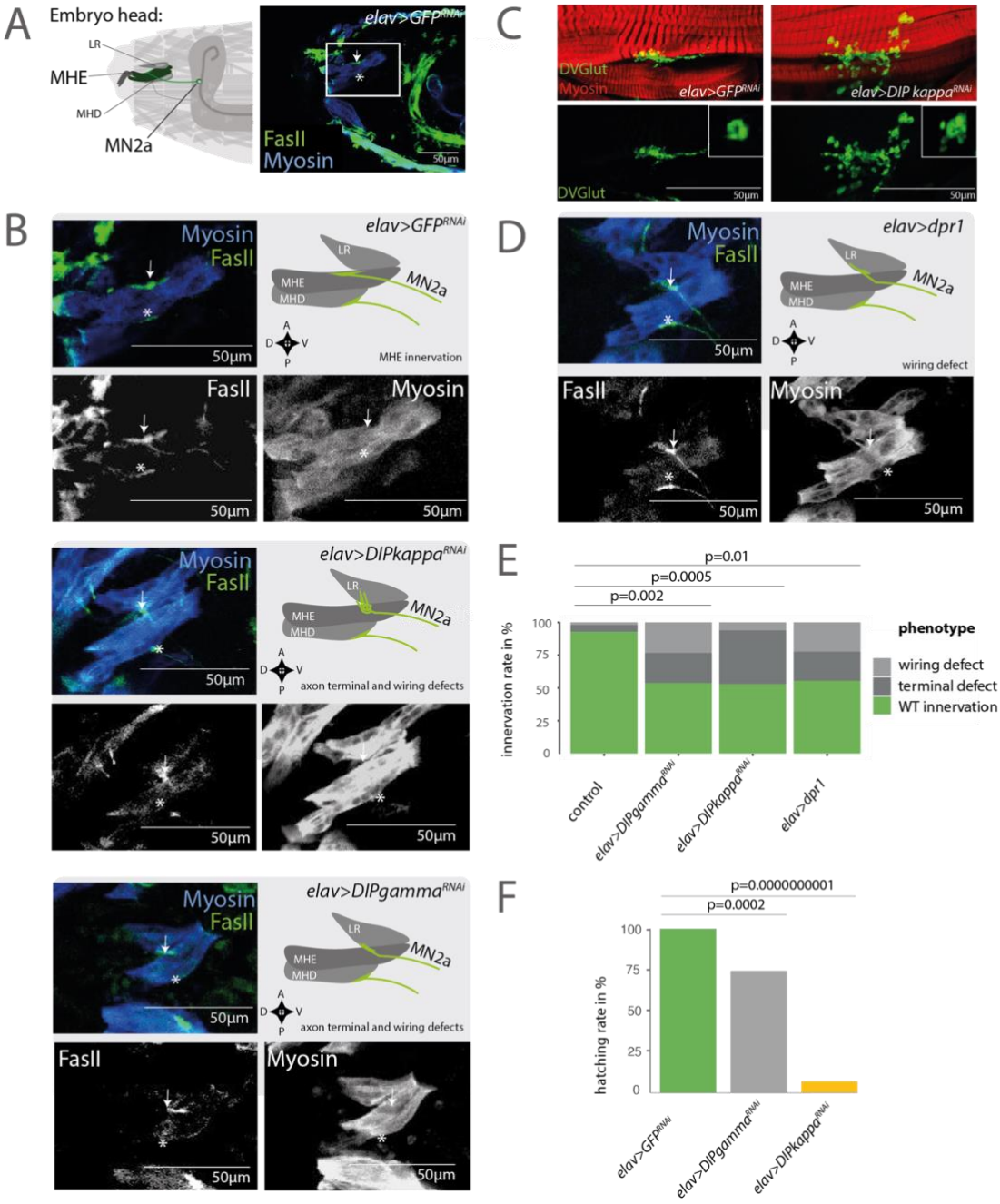


Figure 33: Ig proteins are effectors of synaptic specificity

(A) *Left panel:* Schematic drawing depicts the MN2a motoneuron (light green) that innervates the MHE (dark green) in an early stage 17 *Drosophila* embryo. *Right panel:* Representative confocal picture of an early stage 17 *Drosophila* embryo (*elav>GFP*^{RNAi}) depicts the MN2a motoneuron (arrow) that innervates the MHE. The innervation site on the MHD muscle is highlighted by an asterisk as reference for orientation in the embryo. **(B)** Visualization of three independent genetic experiments using the pan-neural *elav*-GAL4 driver to control the expression of UAS-*RNAi* constructs of homeodomain TFs i. e. UAS-*DIPkappa*^{RNAi}, UAS-*DIPgamma*^{RNAi} and UAS-*GFP*^{RNAi} as control. *Left panel:* Zoom on MHE of an early stage 17 *Drosophila* embryo. Note the upper panel shows a zoom of the image depicted in Figure 33 A as control. *Right panel:* Schematic drawing depicts the innervation phenotype on MHE by the MN2a motoneuron. See classification of phenotypes in Figure 28 B. In UAS-*DIPkappa*^{RNAi}, UAS-*DIPgamma*^{RNAi} animals MN2a MNs display axon 'terminal defects' and 'wiring defects', i.e. incorrect targeting to the LR muscle. **C)** Representative confocal picture depicts the innervation site on the MHE muscle in L3 larvae. In animals depleted of *DIPkappa* in neurons (*elav>DIPkappa*^{RNAi}) enlarged innervation sites are observed and an increased

number of boutons and branching phenotypes (right panel) compared to control animals (*elav>GFP^{RNAi}*) (left panel). **(D)** Representative confocal images of the pan-neural *elav*-GAL4 driver controlling the expression of the Ig domain protein dpr1 that is not expressed in Mn2a. The control experiment (*elav>GFP^{RNAi}*) is depicted in Figure 5B. *Left panel:* Zoom on MHE of an early stage 17 *Drosophila* embryo. *Right panel:* Schematic drawing depicts the innervation phenotype on MHE by the MN2a motoneuron. See Figure 28 B for classification of phenotypes. In wild-type conditions the MN2a motoneuron projects to the MHE, while in *elav>dpr1* animals MN2a MNs display 'wiring defects', i.e. incorrect targeting to the LR muscle. **(E)** Innervation rates for MN2a projections are calculated for *elav>DIPkappa^{RNAi}* (n=17), *elav>DIPgamma^{RNAi}* (n=13) and *elav>dpr1* (n=9) animals as described in Figure 28 C. **(F)** Correct MHE innervation is required for hatching of *Drosophila* embryos from the eggshell (Friedrich et al., 2016). Hatching rate was calculated based on the number of L1 larvae observed after 24h in genetic crosses, depleted of DIPkappa (*elav>DIPkappa^{RNAi}*; n=165) and DIPgamma (*elav>DIPgamma^{RNAi}*; n=210) in neurons compared to control animals (*elav>GFP^{RNAi}*; n=211).

Together these results demonstrate that Ig domain proteins mediate synaptic target specificity in the neuromuscular system downstream of homeodomain TFs.

2.6 HOMEO-TFs MEDIATE TARGET SPECIFICITY IN BOTH MATCHING PARTNERS

The homeodomain TFs Dfd, Mirr, and Hth are co-expressed in both functionally connected cells of the feeding motor unit, the MN2A neuron and the MHE muscle (Figure 34). By contrast, the adjacent LR and MHD muscles do not express these homeodomain TFs. This observation led us to hypothesize that combinations of homeodomain TFs label defined synaptic partners at aligned positions.

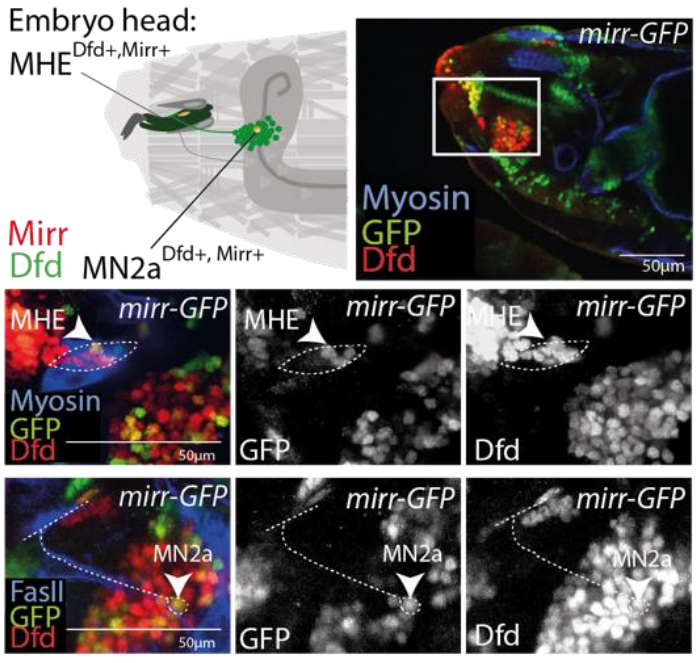


Figure 34: A matching homeo-code in neurons and muscles

Left panel: Schematic drawing depicts the MN2a motoneuron (yellow) expressing the homeodomain TFs Dfd (green) and Mirr (red) that innervates the MHE (yellow nucleus) expressing the same set of homeodomain TFs in an early stage 17 *Drosophila* embryo. *Right panel:* Representative confocal picture of an early stage 17 *Drosophila* embryo (*mirr-GFP*). The mid panel and lower panel show two independent experiments i. e. mid panel displays expression of the homeodomain TFs Mirr and Dfd in the MHE (arrow), while the lower panel depicts expression of Mirr and Dfd in the MN2a motoneuron that innervates the MHE. The MHD muscle underneath the MHE is negative for Mirr and Dfd.

To complement the dataset on MNs, I have used a fly stock expressing endogenously *GFP* tagged *Myosin heavy chain* (*Mhc-TAU-GFP*, (E. H. Chen & Olson, 2001)) to sort for somatic muscle, and performed single-cell RNA-seq with enrichment for *Hox* genes, as described above for MNs. Analysis of this dataset by means of t-SNE and clustering indicates the existence of six relatively distinct subtypes of somatic muscle (Figure 35 A). Using previously described marker genes, these were tentatively identified as Dr+ dorsal somatic muscle, lms+ lateral somatic muscle, mid+ and well as Poxm+ ventral and lateral somatic muscle, Ptx1+ ventral somatic muscle, as well as esg+ara+ progenitor cells. While *Hox* gene expression in most clusters was excessively low, coverage in the $Poxm^{+}$ cluster was similar to the coverage observed in the motoneuronal dataset. Thus, I focused on the Poxm+ cluster to identify highly variably expressed genes (Brennecke et al., 2013) and demonstrated that homeodomain TFs and Ig domain encoding genes were highly variably expressed within otherwise homogeneous somatic muscle subtypes (Figure 35 B), most other clusters showed similar results (Figure 35 C). This finding supports the hypothesis that homeodomain clusters are associated with *Ig* gene expression, following a similar logic than the motoneuronal data set. Nevertheless, spatial mapping of somatic muscle cells could not be performed equally as in the motoneuronal data set, due to the low representation of *Hox* genes in the scRNA muscle data set and due to technical limitations of Hox protein measurements in somatic muscle cells. However, for a small number of cells region specific gene expression patterns were identified based on the described order of *Hox* gene expression along the AP- axis (*Lab, Dfd, Scr, Antp, Ubx, Abd-A, Abd-B*). Thereby I identified regional specific patterns of cell adhesion molecules expressed in specific regions of embryonic muscle cells (Figure 36).

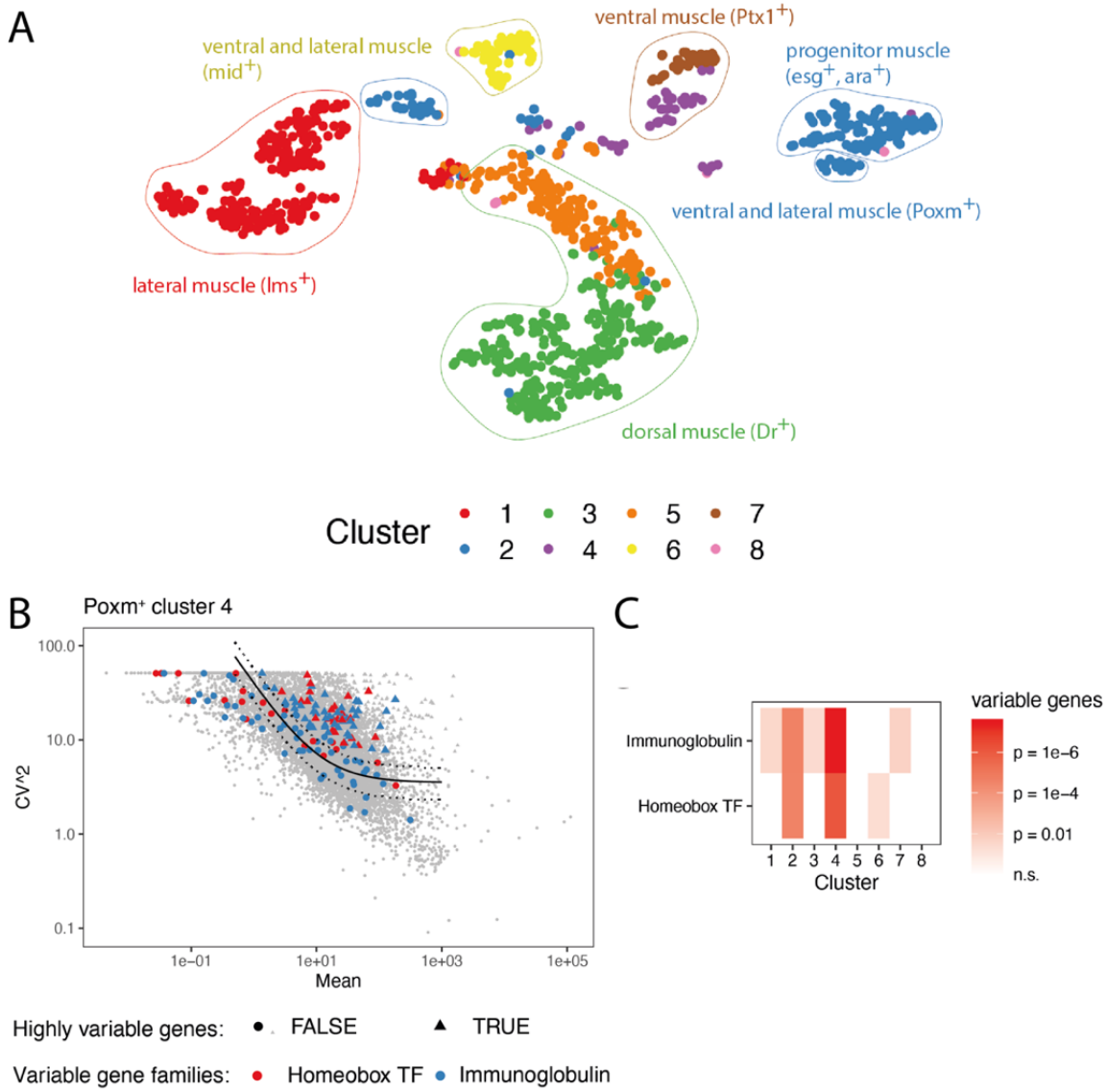


Figure 35: A homeo-Ig-code diversifies somatic muscle cells

A) T-distributed stochastic neighbor embedding (t-SNE) plot of single-cell transcriptomes of GFP expressing somatic muscle cells sorted from a *Mhc-TAU-GFP* expressing fly line. Colors correspond to clusters identified using hierarchical clustering that were annotated using marker gene expression of muscle subtypes. **(B)** Identification of highly variable genes using the method by (Brennecke et al., 2013). Scatter plot depicts for each gene the mean expression and squared coefficient of variation across cells from the Poxm+ cluster 4. The solid line indicates the fit, dashed lines the 95% confidence interval. Genes with a significantly elevated variance are shown as triangles, other genes as circles. Different gene classes are color coded. **(C)** Statistical model of Brennecke et al. (2013) was used to identify highly variable genes (see panel B for an example). Color code displays enrichment of Igs or homeobox TFs among variable genes according to a hypergeometric test.

I then investigated if muscle and neuron from the same segment express cell adhesion receptors and their respective cognate binding partners (Özkan et al., 2013) (see supplementary table 4). Therefore, I compared regional specific gene expression programs of *Ig* genes between the motoneuronal data set and the muscle data set. According to the observations above, matching synaptic partners comprise a similar homeo-code. This homeo-code is associated with an Ig-code. Thus, I correlated *Ig* genes in each tissue with *Hox* gene expression (Figure 36). Thereby,

I identified regional specific *Ig* expression patterns associated with specific *Hox* signatures. In particular heterophilic cell adhesion molecules such as candidates of the dpr/DIP family and beaten path members show regional specific correlations. However, comparison of these candidates with in vitro interactome studies (Özkan et al., 2013) suggested very few interactions of *Ig* genes among similar regions such as the interaction of *otk/otk2* in *Abd-B* expressing MNs and muscle cells (see Figure 36 and supplementary table 1). Most predicted interactions are identified across different regions i. e. *DIP-k* is enriched in Dfd+ MNs, while their predicted interaction partner *dpr7* or *dpr1* are expressed in lab+ or Scr+ muscle cells. Importantly, the poor identification of interaction partners can be due to the low representation of *homeotic* genes in single cells of the muscle data set (~25%) or the identification of a homeo-Ig-code in individual cellular identities is necessary to identify a matching Ig connectivity code between synaptic partners. However, the poor sample quality of the muscle data set and the lack of immunofluorescence data for spatial mapping, did not allow for this analysis.

			Lab	Dfd	Scr	Antp	Ubx	AbdA	AbdB
Lab	Motoneuron	*dpr2*	**0,15**	0,01	0,09	0,10	0,09	0,09	-0,01
	Muscle	*kek5*	**0,14**		0,02	0,04	0,10	0,09	0,06
		dpr7	**0,12**		0,08	0,01	0,04	0,06	0,06
		dpr18	**0,13**		0,07	0,02	0,09	0,10	0,05
Dfd	Motoneuron	*beat-IV*	-0,05	**0,16**	0,01	-0,04	0,10	0,03	0,00
		DIP-k	0,04	**0,15**	0,01	-0,05	-0,08	-0,10	-0,07
	Muscle								
Scr	Motoneuron	*lea*	0,11	0,19	**0,20**	0,02	0,03	0,05	0,01
		beat-IIIa	0,05	-0,04	**0,15**	0,09	0,02	-0,03	0,01
		CG18268	0,15	0,00	**0,33**	0,04	0,07	0,04	0,04
	Muscle	*dpr4*	0,11	0,19	**0,20**	0,02	0,03	0,05	0,01
		dpr1	0,05	-0,04	**0,15**	0,09	0,02	-0,03	0,01
Antp	Motoneuron	*dpr13*	-0,03	0,08	0,03	**0,23**	**0,23**	0,14	0,08
	Muscle								
Ubx	Motoneuron	*kek5*	0,01	0,03	0,01	0,12	**0,15**	0,08	-0,01
		plum	-0,02	-0,08	-0,07	0,13	**0,18**	0,12	-0,01
		DIP-θ	0,04	-0,06	-0,02	0,11	**0,15**	0,08	0,05
		DIP-δ	0,03	-0,04	0,09	0,17	**0,29**	0,18	0,06
		Bsg	-0,02	-0,01	0,03	0,19	**0,20**	0,18	0,05
	Muscle	*Dscam1*	0,09		0,22	0,16	**0,24**	-0,01	0,10
		Dscam4	0,03		0,06	0,05	**0,27**	0,06	0,03
		DIP-alpha	0,08		0,07	0,03	**0,20**	-0,01	0,03
		dpr13	0,07		0,14	0,11	**0,16**	0,13	0,14
Abd-A	Motoneuron	*fas*	-0,06	-0,04	-0,05	0,14	0,12	**0,16**	0,07
		lar	-0,08	-0,07	0,04	0,08	0,14	**0,17**	0,10
		Dscam2	0,05	-0,06	-0,03	0,12	0,19	**0,24**	0,10
		dpr18	0,01	-0,01	-0,07	0,12	0,14	**0,15**	0,09
		Hasp	-0,04	0,00	-0,05	0,19	0,23	**0,26**	0,11
		DIP-n	-0,03	-0,03	-0,02	0,07	0,10	**0,16**	0,06
		babos	0,04	0,00	0,01	0,07	0,08	**0,14**	0,13
		otk2	0,03	-0,02	-0,05	0,20	0,30	**0,38**	0,33
	Muscle	*Dscam2*	0,05		0,10	0,10	0,15	**0,22**	0,20
		otk	0,11		0,05	0,01	0,03	**0,13**	0,05
		dpr8	0,05		0,07	0,04	0,12	**0,16**	0,07
Abd-B	Motoneuron	*otk*	0,05	0,04	-0,05	0,19	0,28	0,34	**0,43**
	Muscle	*otk2*	0,05		0,07	0,05	0,09	0,16	**0,23**
		dpr17	0,03		0,07	0,07	0,03	0,12	**0,26**
		dpr10	0,02		0,05	0,07	0,05	0,14	**0,18**

Figure 36: Differential Ig domain encoding gene expression across the Drosophila AP axis

Hox genes were Pearson correlated with all genes expressed in the motoneuronal and muscle data set. In the heatmap are the Ig candidates depicted that are correlating significantly with *Hox* genes FDR > 0.1. Correlation between *Hox* genes and differential expressed Ig domain encoding genes are color coded (red = significant negative Pearson correlation; green = significant positive Pearson correlation). Ig domain encoding genes are clustered according to highest correlation with a *Hox* gene (Columns). *Hox* genes are sorted for AP localisation from anterior to posterior (rows and columns). In addition, Supplementary table 4 shows predicted *Ig* protein interactions of candidates expressed in the motoneuronal and muscle scRNA data set.

To further study the role of homeodomain TFs in both muscles and neurons, I interfered with the expression of *Dfd* in a tissue dependent manner to examine synaptic defects of stage 17 embryos. Hence, I knocked down *Dfd* (UAS-*Dfd*^{RNAi}) individually in neurons (*elav*-GAL4) and in muscles (*Mef2*-GAL4) or in both tissues (*elav*-GAL4; *Mef2*-GAL4;). In all three conditions I observed significant synaptic wiring defects and terminal defects (Figure 37A-F).

Interestingly, the strength and type of wiring defects differed in each tissue. *Dfd* depletion in neurons and muscles alone showed a wiring defect in 29% and 48% of the embryos respectively, while knockdown of *Dfd* in both tissues increased the rate to 70% of cases, indicating that homeodomain TFs are required in both tissues neurons and muscle to coordinate synaptic partner matching. *Dfd* depletion in MN2a led to mistargeting of the LR muscle (Fig. 37C), while *Dfd* depletion in muscles led to innervation of the MHE muscles by MNs that were supposed to innervate the MHD muscles (Fig. 37D). Knockdown of *Dfd* in both muscle and neurons caused aberrant innervations of the MHE muscles and loss of the stereotypic innervation (Fig. 37E), suggesting that homeodomain TF expression in both tissues is required to regulate repellent and adhesive factors for synaptic target selection thereby preventing mistargeting. In all three conditions, hatching rates were significantly decreased: by 23% when *Dfd* was depleted in neurons, by 29% when *Dfd* levels were reduced in muscle, and by 99% when *Dfd* was silenced in both tissues (Fig. 37F). These results demonstrated the importance of homeodomain TF expression in both neurons and muscles for synaptic wiring and behavior.

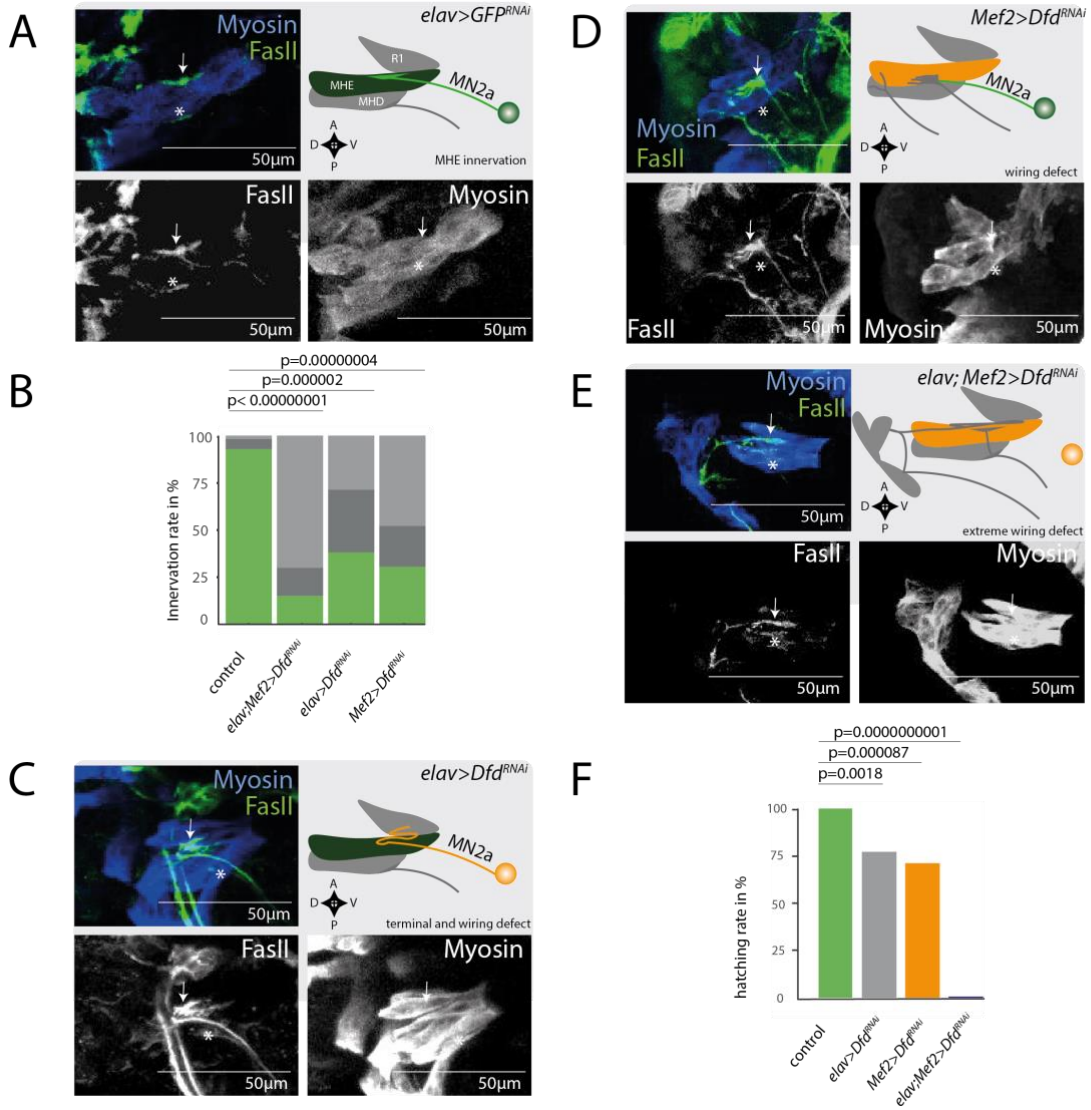


Figure 37: Synaptic targeting requires the homeodomain TF Dfd in neurons and muscles

(A-E) Visualization of three independent genetic experiments using the pan-neural *elav*-GAL4 driver, the muscle driver (*Mef2*-GAL4) and both of these drivers combined (*elav*-GAL4*; Mef2*-GAL4) to control the expression of a UAS-*Dfd*^{RNAi} construct. *Left panel:* Zoom on MHE of an early stage 17 *Drosophila* embryo. Note panel 6b and 6c show the same images as depicted in Figure 5B. *Right panel:* Schematic drawing depicts the innervation phenotype on MHE by the MN2a motoneuron. See classification of phenotypes in Figure 3B. In *elav>Dfd*^{RNAi} animals MN2a MNs display axon terminal defects and ectopic targeting to the LR muscle, while *Mef2>Dfd*^{RNAi} and *elav;Mef2>Dfd*^{RNAi} show extreme wiring defects and ectopic innervations on the MHE muscle target. **(B)** Innervation rates for MN2a projections are calculated for *elav>Dfd*^{RNAi} (n=21), *Mef2>Dfd*^{RNAi} (n=23) and *elav;Mef2>Dfd*^{RNAi} (n=20) animals as described in Figure 3C. **(F)** Hatching rates of animals depleted of Dfd in neurons (*elav>Dfd*^{RNAi}*;* n=217), in muscle (*Mef2>Dfd*^{RNAi}; n=284) in neurons and in muscle (*elav;Mef2>Dfd*^{RNAi}; n=2403) compared to control animals (mock; n=2898) was calculated as described in Figure 33 F.

In sum, the results showed that combinations of homeodomain TFs are required to align synaptic partners in the neuronal and muscle tissues and coordinate their preferences in synaptic target choice.

2.7 CIRCUIT SPECIFIC ASSEMBLY AND CONNECTIVITY MECHANISMS MEDIATED BY HOMEO-TFs

A recent study in *C. elegans* revealed that individual combinations of homeodomain TF describe every class of neuron in the entire nervous system (Reilly et al., 2020). Strikingly, these homeo-codes neither correlate with cell type (sensory neurons, MNs, interneurons) nor with neurotransmitter expression or lineage identity, raising the question whether this code is required for connecting functionally related cells of one circuit.

In line with this hypothesis, immunofluorescence staining's of the homeodomain TF Dfd showed expression in all functionally connected cell types of one neuro-muscular circuit including the sensory input system (olfactory cells), a specific part of the CNS associated with feeding (SOG), a higher order brain region for olfactory learning and memory (Kenyon cells) and the neuro-muscular output system (Figure 38).

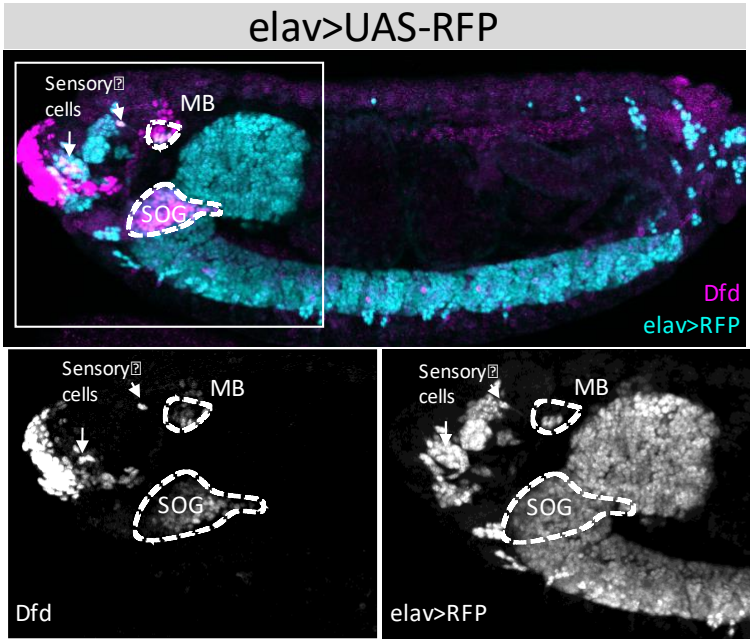


Figure 38: Dfd is expressed in functionally connected cell types of the feeding circuit

Confocal picture depicts the Dfd expression pattern and co-expression pattern with the pan-neuronal *elav*-GAL4 driver on a late stage *Drosophila* embryo. The lower left panel visualizes Dfd expression in sensory cells (white arrows), the MB and the SOG. All three regions are labelled by the pan-neuronal *elav*-GAL4 driver (lower right panel).

Together with the data above, this indicates that homeodomain TFs can indeed label and regulate the formation of functionally connected neuronal circuits.

3. DISCUSSION

Each neuron in the nervous system chooses a single target cell from a high number of possible interactions, which is a prerequisite for the stereotypic formation of neuronal circuits. This extraordinary degree of precision is thought to be mediated by a matching code of cell recognition and adhesion molecules such as CSPs. However, up to date a systematic scrutinization of this code and upstream mechanisms fine-tuning its expression are missing.

3.1 A scRNA EXPERIMENT DESIGNED TO IDENTIFY MOLECULAR AND CELLULAR PROGRAMS OF INDIVIDUAL NEURONAL CELLS

Our experimental design allowed us to overcome several challenges that have previously impeded the identification of mechanisms mediating cellular individuality during synaptic wiring. First, by focusing on one neuronal subtype, the motoneuronal population in *Drosophila* embryos, we were able to reduce neuronal complexity. Second, we investigated this cell population exactly at the time when they form stereotypic connections with their muscle targets, which allowed us to identify molecular cues critical for synaptic wiring. Third, MNs form highly cell-specific connections with muscles and are present in a relatively small number per embryo. Thus, we were able to use a single-cell genomic approach with a high number of biological replicates of every biologically unique cell. Thereby, we identified novel markers specific to individual cell identities or small groups of cells that in turn permit the identification of transcriptome signatures relevant for wiring. And finally, we implemented a spatial mapping approach based on *Hox* gene expression to locate MNs along their AP position to gain insights into the role of spatial mechanisms during synaptic wiring.

3.2 SPATIALLY ORGANIZED HOMEO-TFs ACT IN INDIVIDUAL NEURON IDENTITIES

Our scRNA-Seq data revealed that a homeo-code acts as major determinant for transcriptional heterogeneity during the wiring phase of MNs. Since scRNA-Seq, despite the use of a high number of biological replicates for each unique cell, cannot be used to unanimously identify biologically unique cells, we used imaging to demonstrate its specificity to single cells in five out of five candidate single-cell specific homeo-codes. Together, these data allowed us to conclude that a homeo-code defines cellular identities in *Drosophila* MNs, possibly down to the level of single cells. Recent work demonstrated the existence of a similar code in *C. elegans*, suggesting that homeodomain TFs might be utilized to specify unique neuronal cell identities throughout evolution. Indeed, homeodomain TFs are known to play a role in the specification

of mammalian neuronal subtypes as well (Buelow et al., 2005; Dasen, 2018; Mallo, 2014; Song & Pfaff, 2005).

Interestingly, the homeo-code shows clear ON/OFF states, reminiscent of data from mouse, where low transcriptional noise of homeo-TF expression serves to robustly define cell *populations,* but not single cells (Sugino et al., 2019). Epigenetic mechanisms might underlie the robustness of the observed OFF state of homeodomain TFs: studies in mouse demonstrated that about 100 homeodomain TFs are bound by Polycomb group proteins (Montavon & Soshnikova, 2014; Sugino et al., 2019), while more detailed analysis in *Drosophila* indicated that the Polycomb complex mediates an inaccessible chromatin state in *Hox* genes and target genes leading to lineage commitment (Bantignies et al., 2011; Cheutin & Cavalli, 2018; Domsch et al., 2019).

To gain insight into the developmental origin of the homeo-code, we spatially mapped all MNs along the AP axis using *Hox* gene expression as markers. We validated the high accuracy of our mapping approach, and showed that our strategy requires a much smaller number of markers compared to existing approaches for the spatial mapping of single-cell transcriptomes (Achim et al., 2015; Bageritz et al., 2019; Satija et al., 2015). We found that position along the AP axis decisively impacts the expression of homeodomain TFs at the synaptic wiring stage. Furthermore, genes associated with the DV position of MNs highly co-vary statistically independent of AP position. These results suggest that during development, molecular position is imprinted early on and affects cellular function during synaptic wiring. However, we also observed substantial additional variability of homeodomain TF expression, suggesting that beyond such spatial mechanisms, other processes induce the specification of unique cellular identities. Neuronal birth order is a possible candidate process for future follow-up (Kulkarni et al., 2016; H. Li et al., 2017).

3.3 A HOMEO-CODE ACTS IN INDIVIDUAL NEURON IDENTITIES AND FINE TUNES SYNAPTIC SPECIFICITY

A recent study in *C. elegans* revealed that a combinatorial homeo-code in single neuronal classes neither correlates with cell type (sensory neurons, MNs, interneurons) nor with neurotransmitter expression or lineage identity (Reilly et al., 2020). Instead, they found an association of the homeo-code with functionally related neurons. Earlier data from Ingrid Lohmann's lab has shown that the *Hox* gene *Dfd* labels specifically the motoneuronal circuit that controls feeding movements (Friedrich et al., 2016). Together, these studies suggest that homeo-codes might label neuronal circuits. This thesis provides conclusive evidence that

homeo-codes functionally specify neuronal circuits. Our data demonstrates that homeodomain TF expression is not only highly specific to cellular identities, but is also associated with profound differences in the entire transcriptome and in particular the expression of Ig domain CSPs as possible effectors for circuit wiring. Chromatin Immunoprecipitation experiments of homeodomain TFs show binding on regulatory sequences of *Ig* genes and conversely, *Ig* genes exhibit motifs of homeodomain TFs in their regulatory elements, suggesting a putative direct regulation of homeodomain TFs on *Ig* expression. Importantly, manipulation of homeodomain TF expression during the wiring phase by knockdown or ectopic expression leads to changes in target preferences, which are phenocopied by manipulations of their *Ig* targets. These effects are predictable (i.e. mistargeting to a neighboring segment as in the case of *Dfd* knockdown and ectopic *lab* expression), or extreme (such as in the case of *hth* knockdown). Previous studies showed that distinct complexes of homeodomain TFs Hth-Exd-Hox modulate DNA recognition and subsequent target gene expression differently (Noro et al., 2006; Ryoo et al., 1999; Slattery et al., 2011). Our data also show that a similar mechanism acts in muscles to determine the neuronal interaction partner. This suggests that the cooperative binding of TF and co-factors on regulatory sequences of target genes (Jolma et al., 2015; Kribelbauer et al., 2020) modulates synaptic specificity across individual cell identities in functionally related cells.

Together, the combination of single cell genomics (co-expression), chromatin immunoprecipitation (binding of enhancers) and genetics (phenocopies) demonstrate that while the homeo-code specifies unique cellular identities, Ig proteins act as the effectors of synaptic specificity.

3.4 IG PROTEIN EXPRESSION IS A MEDIATOR OF SYNAPTIC SPECIFICITY

Classification of cellular identities using the homeo-code allowed us to systematically investigation the CSP distribution between cells. In line with results of De Wit and Gosh, most CSPs change more gradually between cells, and unlike homeodomain TF expression, the binary expression of Ig proteins alone is insufficient to specify single cells (De Wit & Ghosh, 2016; van Oostrum et al., 2020). Previous studies on single molecules even indicated that already small changes in relative expression levels of CSPs in matching partners can change synaptic specificity (Sweeney et al., 2012; Yogev & Shen, 2014). Here only one Ig molecule class, the DIP genes, were found to be specific for cellular identities. Genetic manipulations on single candidates cause synaptic specificity defects and change synaptic affinity to target cells (ON-Mode). However, the penetrance of these knockdowns is not 100%, suggesting that combinatorial expression with other factors is necessary to robustly change cellular affinities.

Due to the lethality of these knockdowns and the involvement of DIP-Dpr networks in cell survival, these wiring defects might not be observed in animals that survive (Carrillo et al., 2015; Menon et al., 2019; Xu et al., 2018). Furthermore, we found in our data set some *Ig* genes, such as *Dscams*, to be broadly expressed but not present in some defined cells (OFF-Mode). Finally, most other *Ig* genes are less specific for cellular identities and relative expression levels change gradually between cells, as suggested previously (De Wit & Ghosh, 2016; van Oostrum et al., 2020). In sum, our data show that a highly combinatorial Ig-code drives synaptic specificity and connectivity between cells.

3.5 A MODEL FOR SYNAPTIC WIRING IN THE NEUROMUSCULAR SYSTEM

Our data justify a model whereby the position of every cell is imprinted early in embryonic development by patterns of homeodomain TFs (Fig. 7). Expression patterns of these factors become more complex and combinatorial with each cell division until small groups of cells and possibly every single cell is uniquely labelled by a homeo-code. These unique combinations of homeodomain TFs then in turn regulate specific downstream programs of *Ig* gene expression. We have shown that in the target cells (here: muscles), connectivity is functionally specified by a similar TF code, which possibly induces the expression of a complementary Ig receptor expression program. This molecular logic enables every single cell to find its corresponding interaction partner based on complementary adhesive properties mediated by combinations of Ig domain molecules. In sum, this concept explains how a molecular memory of cell body position is translated into invariant cell-cell adhesion events by means of the homeo-Ig-code.

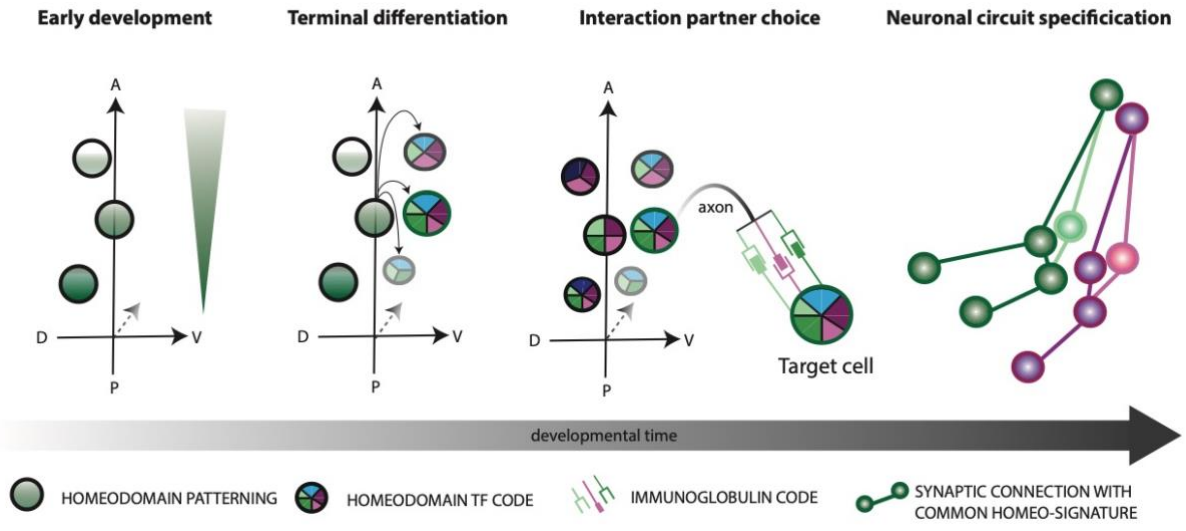


Figure 7: The molecular mechanism for stereotyped synaptic partner selection and neuro-muscular circuit architecture

Left panel: Early developmental programs pattern single cells according to their position in the embryo. Morphogenic gradients establish specific homeodomain patterns along embryonic body axis (i.e. dorsal ventral axis = DV axis, AP axis = AP axis). *Left middle panel:* After several divisions of NBs, neurons terminally differentiate and establish unique identities with distinct homeodomain TF codes (color code). *Right middle panel:* A unique homeodomain TF code specifies Ig domain receptor expression in matching synaptic partners and differential affinities of Ig domain proteins promote selective synaptic target choice. *Right panel:* Common homeo-signatures specify interaction partners within individual neuro-muscular circuits.

3.6 HOMEODO-TF EXPRESSION AND NEURONAL CIRCUIT DEVELOPMENT DURING EVOLUTION

A prerequisite for the function of our peripheric nervous system is the precise matching of neuron-muscle pairs. Therefore, every neuro-muscular circuit is localized in specific regions to fulfil distinct stereotypic movements such as hatching, crawling and feeding. This raises the fundamental question about the evolutionary origin of the underlying molecular mechanisms for synaptic matching.

The nervous system found in vertebrates and invertebrates emerged from a simple nerve net of loosely interconnected neurons that covers the body wall and the digestive tract of metazoan animals (Detlev Arendt, Musser, et al., 2016; Denes et al., 2007; Holland, 2003; Miller, 2009; Weiss et al., 1998). Nerve net neurons of cnidaria and ctenophores directly innervate neighboring muscle fibers to perform simple muscle contractions in response to stimuli (Detlev Arendt, Tosches, et al., 2016; Galliot & Quiquand, 2011; Watanabe et al., 2009), suggesting that simple neuron-muscle partners are already specified early in metazoan evolution. However, no regional specific functions of these units are reported in early metazoans, where *Hox*- and *homeo*- like genes are not functionally involved in A-P axis patterning, even though expression in anterior and posterior parts of early metazoan animals are observed (Chourrout et al., 2006; Steinmetz et al., 2011). Interestingly, in ancestors of the bilaterian clade both the functional role of *Hox* genes in A-P patterning and a complex nerve cord with regional specific locomotion's similar to the nervous system in vertebrates evolved. In most bilaterians, such as *Drosophila*. a brain structure processes sensory inputs and translates them into motor outputs. In addition, a nerve cord extends along the animal body axis to coordinate regional specific locomotion. The peripheral nervous system is similar to the nerve net in early metazoan. It consists of nerve fibers linked between the CNS (brain and VNC and sensory neurons or muscle fibers. Although the nervous system is restructured in the bilaterian taxa compared to early metazoans, a putative co-evolution of the nerve net and regional specific homeodomain TF expression patterns are a possible mechanism driving increased neuronal complexity during evolution.

In this study, I show that matching synaptic partners are specified by distinct homeo-codes, even over long distance or when non-linear mapping rules are involved, but it remains unclear how this concept has been established during metazoan evolution.

In sum, synaptic partners are pre-defined during nerve net evolution of metazoan, while functional regionalization of neuronal circuits evolved during bilaterian evolution, when Hox and homeodomain TFs developed functions in regional patterning.

4. OUTLOOK

Our data suggest that the structure of neuronal circuit is encoded in the gene regulatory logic of Ig domain encoding genes. How their regulation by homeo-TFs is hardwired into their cis-regulatory elements is a topic for future investigation. In particular, it is unclear how specificity is achieved, despite homeodomain TF binding to relatively similar motifs. Future chromatin immunoprecipitation (ChiP) experiments with different homeodomain TF, allows for identification of novel and combinatorial binding motifs on regulatory elements of Ig domain proteins. To this end, ChiP-Seq experiments can be performed in a tissue specific manner in combination with fluorescence activated cell sorting (FACS) that selects and enriches for a tissue type of interest. For example, reporter gene expression for selection can be activated in neurons (elav-Gal4), MNs (Ok371-Gal4) or somatic muscle cells (Tau-Mhc-GFP). After ChiP-Seq experiments of different homeodomain TFs in a tissue specific manner, novel binding motifs can be functional dissected, and these elements can be investigated in enhancer trap lines in vivo or EMSA assays in vitro. Together, this analysis can reveal how specificity is achieved by cooperative binding of homeodomain TFs on Ig targets. Going beyond hypothesis-driven investigation of individual enhancers and TFs, massively parallel reporter assays *in vivo* could be used to further dissect the regulatory logic of CSP enhancers (Fuqua et al., 2020), but require capabilities of creating and handling fly lines at very high throughput, as well as advanced systems for rapidly querying the spatial expression patterns of reporter genes. Together, such data will enable a quantitative understanding of the cis-regulatory codes specifying stereotypic neural wiring.

Next to the cis-regulatory logic, it will be interesting to investigate how combinatorial interactions between homeodomain TFs, as well as CSPs, affect synaptic specificity. In mammalian systems, combinatorial genetic screens (Replogle et al., 2020; Zetsche et al., 2017) are being translated to an *in vivo* setting, also in a neurobiology context (Jin et al., 2020). Currently, these screens are limited to gene expression readouts, and can systematically profile the gene expression consequences of combinatorial homeo-TF knockdowns or overexpression. In the future, advances in connectome sequencing (Huang et al., 2020) will permit for massively parallel genetic screens that query the connectomic consequences of combinatorial TF and/or CSP knockdowns.

Our data also indicates a coordinated expression of homeodomain TFs in single units of functionally connected cells. Whether downstream programs of homeodomain TFs such as Ig

domain encoding genes are following a similar logic is matter of future examinations. In particular, the existence of a matching Ig domain expression code in individual synaptic partners is unknown. Preliminary experiments in this study have shown the correlation of *Hox* genes and *Ig* genes in the neuronal and muscle tissue, indicting a segmental distribution of some *Ig* genes (see results in section Figure 36). A multitude of possible interactions between *Ig* genes have been described, making an inference of binding partners difficult (see supplementary table 4). In addition, it is often unknown if these interactions are adhesive or repulsive, or homophilic or heterophilic (Figure 40 A, B, C and supplementary table 4).

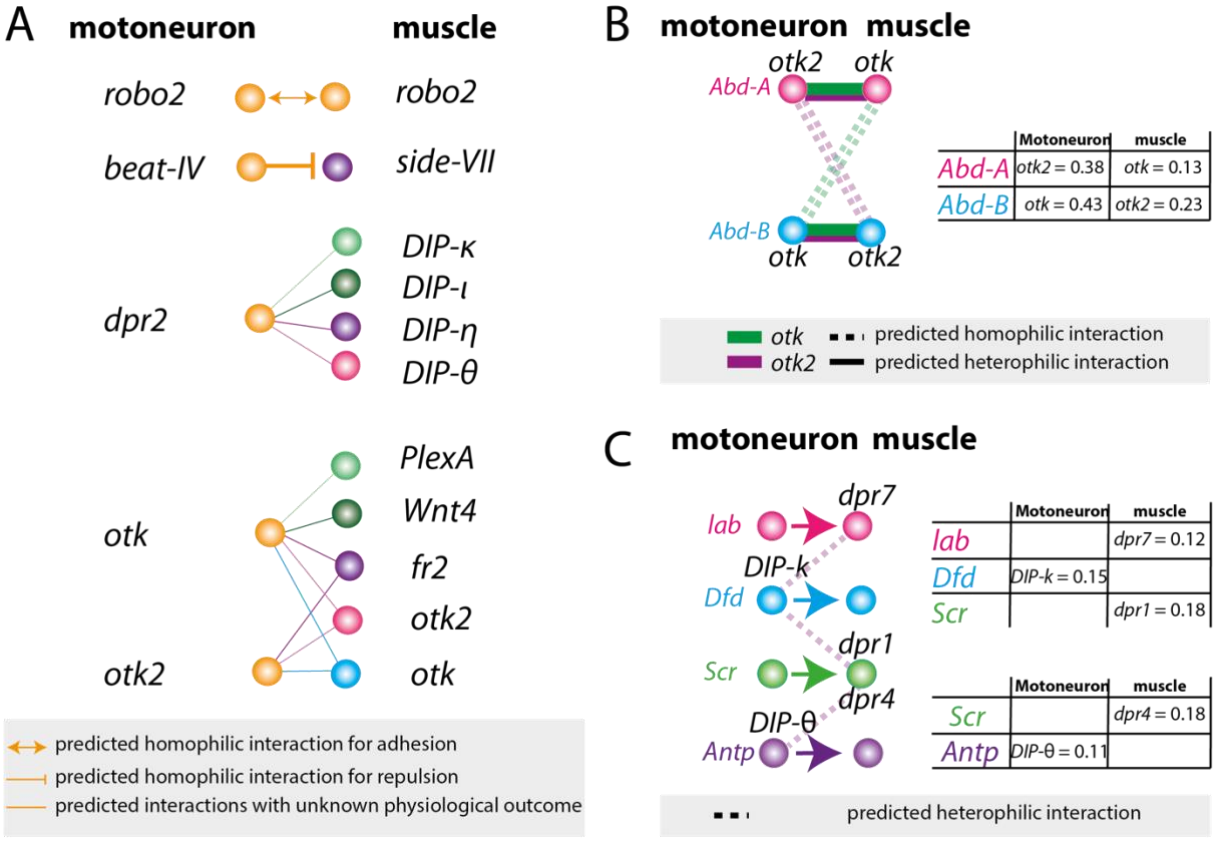


Figure 40: Predictions of cell-cell interactions mediated by Ig domain proteins

(A-C) *Ig* domain encoding genes are selected from the motoneuronal and muscle scRNA data set and compared with in vitro interactome data of Özkan et al. 2013. (A) Examples for binding predictions of selected Ig domain proteins (list of the entire set of predicted interactions, see supplementary table 4). (B) Interaction type of the regional enriched Ig domain encoding gene *otk* and *otk2* in *Abd-A* and *Abd-B* expressing MNs and muscle cells (see Figure 36 and supplementary table 4). The type of interaction can be homophilic or heterophilic, while the cellular outcome (adhesion or repulsion) is unknown. (C) Dpr/DIP domain encoding genes are identified across different regions i. e. *DIP-k* is enriched in *Dfd* expressing MNs, while their predicted interaction partner *dpr7* or *dpr1* are expressed in *lab* or *Scr* expressing muscle cells. Dpr and DIP interactions in the same region were not found to be enriched in the motoneuronal and muscle data set. The cellular outcome of these interactions (adhesion or repulsion) is not described in literature.

Further experiments in this study resolved the Ig domain expression code of individual cell identities in MNs, while the identification of such a code in individual somatic muscle cells could not be resolved due to technical limitations (see discussion above). Thus, in future studies the experimental design needs to be improved; first, by sorting of MNs and muscle cells in a single experiment to avoid temporal biases of gene expression; second, by targeting not only *Hox* genes, but also other discriminative homeodomain TFs in single cell sequencing experiments to increase coverage; third, by refining the dissociation protocol of somatic muscle cells to increase the yield of high-quality somatic muscle cells. Ultimately, a spatial mapping approach of somatic muscle cells needs to be developed to assign individual cell identities to functionally connected neurons and muscles. In addition, we can apply computational approaches to predict close-proximity locations of cells based on co-expression of cell adhesion molecules such as Igs and their cognate binding partner (Baccin et al., 2020). After identification of such a cognate Ig-code in matching synaptic partners, multi-color fluorescence in situ hybridization (FISH) experiments can be used for validation in individual MNs and muscle cells. In sum, exact scrutinization of a homeo-Ig-code in individual matching muscle and neuronal cells in space, will broaden our understanding on the molecular logic of synaptic wiring.

5. MATERIAL AND METHODS

5.1 DROSOPHILA STRAINS AND EXPERIMENTAL CROSSES

We used the *OK371*-GAL4 driver (Mahr & Aberle, 2006) crossed to UAS-*mCD8-RFP* to perform single cell sorting of MNs, and the *Mhc-TAU-GFP* line (E. H. Chen & Olson, 2001) for sorting of somatic muscle cells. Crosses were kept for 1 hours at 25°C to oviposit on apple-juice plates with yeast paste. Subsequently eggs were incubated on apple-juice plates for further 19 hours and then embryos were dissociated for FACS sorting.

For genetic experiments, we used the pan-neuronal *elav*-GAL4 driver (L. Luo et al., 1994) and the muscle specific *Mef2*-GAL4 driver (Ranganayakulu et al., 1998) as main driver. To confirm phenotypes in MNs the *OK6*-GAL4 driver (Mahr & Aberle, 2006) was used in a few examples. In this study we used second generation *TRiP RNAi* lines from Bloomington (Ni et al., 2011), because these *short hairpin RNA* (*shRNA*) are more effective in embryonic stages than long hairpin of first generation of the TRiP project based on own experiences. Therefore, only one UAS-*RNAi* line fulfilling the required criteria has been available for most of the target genes. The UAS-*Ubx-HA* line is described in (Domsch et al., 2019). For innervation rate assays we crossed virgins of the corresponding driver line to males carrying UAS-*RNAi* or overexpression constructs. As Control we performed the experiment under the same conditions with UAS-*mcd8-RFP* lines or UAS-$EGFP^{RNAi}$ lines. The crosses were kept for at least 16 hours at 29°C for knockdown experiments and 20 hours at 25°C for overexpression experiments on apple-juice plates. Early stage 17 embryos were selected after embryo fixation.

For hatching rate assays, the *elav*-GAL4 driver, *Mef2*-GAL4 and the *elav*-GAL4; *Mef2*-GAL4 driver (designed for this study) were crossed to UAS-Dfd^{RNAi} (designed using the second generation of TRiP system) or to Bloomington lines UAS-$DIPgamma^{RNAi}$ and UAS-$DIPkappa^{RNAi}$. UAS-*mcd8-RFP* lines or UAS-$EGFP^{RNAi}$ lines crossed to the above-mentioned driver lines were used as controls. Crosses were set up in duplicates (equal amounts of males and females for every replicate, sample and control, 1:1 ratio) and egg laying was performed at 29°C for 3 hours on apple-juice plates with yeast paste. Subsequently, eggs were washed, counted and unfertilized eggs were removed or counted and transferred on fresh apple juice plates without yeast paste. Eggs were incubated for additional 24 hours and then the hatching rate was quantified.

For characterization of synaptic phenotypes in larvae (L3), we crossed *elav*-GAL4 to UAS-Dfd^{RNAi} or UAS-$DIPkappa^{RNAi}$ and developed progeny for about 6 days at 25°C for larval dissections and fixation.

5.2 PLASMID CONSTRUCTION AND TRANSGENESIS

The UAS-*Dfd*^{RNAi} line was generated following the conventional TRiP protocol (second generation).

5.3 IMMUNOHISTOCHEMISTRY

Immunofluorescence experiments on *Drosophila* late stage embryos and mouth hook muscles and CNS of late L3 larvae was performed according to standard protocols.

The embryo fixation protocol in brief. First embryos were bleached for 2 min to remove the chorion. After washing in water, embryos were transferred to fixing solution (3,7% Formaldehyde in PBS + 100% Heptane) and incubated for 20 min at room temperature on a nutator. Fixing is stopped by removal of Formaldehyde. Then equal amounts of Methanol are added to Heptane and vortexed for about 40 seconds to remove the vitellin membrane. Subsequently the Heptane phase is removed, and embryos are washed in Methanol.

For antibody stainings, embryos were hydrated and washed for 3 times in PBT (with 0.1% Tween 20). The primary antibodies were used at 4°C overnight from Abcam (rat anti-Myosin 1:1000) DSHB (mouse anti-FasII, 1:50; mouse anti-Scr, 1:50; mouse anti-Abd A, 1:50; mouse anti-Antp; rat anti-Elav 1:50), Invitrogen (rabbit anti-GFP, 1:500), produced by Katrin Domsch (guinea pig anti-Dfd 1:500; guinea pig anti-Ubx 1:500; anti-Lab 1:500) or unknown source (anti-Abd B, anti-Hth, rat anti-Vvl). After 3x washing in PBT (with 0.1% Tween 20) embryos were incubated for 2 hours at room temperature with secondary antibodies from Jackson Immunoresearch. Vectashield with DAPI or TSO was used as mounting medium.

For double stainings with antibodies originating from the same animal (mouse anti-FasII + mouse anti-Antp; mouse anti-FasII + mouse anti-Abd A; mouse anti-FasII + mouse anti-Scr), we performed a modified protocol for sequential antibody staining by TSA (according to the manufacture protocol) The CNS and mouth hooks with mouth hook muscles attached were dissected from L3 larvae. The mouth hooks with larval brains were fixed with 4% paraformaldehyde for 20 min at room temperature and subsequently washed 3x with PBT (0.3% Triton-100). Primary and secondary antibodies were used with similar concentrations than for antibody staining's of embryos.

5.4 MICROSCOPY AND IMAGE ANALYSIS

Fixed embryos and larval brains with mouth hook muscles were imaged by Leica TSC SP8 confocal microscope. The 20x Objective was used for imaging of embryos and the 63x objective to visualize the neuromuscular junction of the MHE muscle. Images were processed

by Fiji. Embryos used for quantification of innervation rate assay were processed with a standardized imaging pipeline (programed by Patrick van Nierop). Confocal pictures shown in figures were processed manually to optimize the result.

For AP axis measurements along the VNC, we used FasII staining as reference and measure protein intensities of Lab, Dfd, Scr, Antp, Ubx, Abd A and Abd B by Fiji. We normalized the length of the VNC between different embryos (0=anterior to 2000=posterior), set thresholds for the protein intensities to remove the overall unspecific background caused by antibody staining's and normalized the protein intensities of different source antibodies (0=min intensity to 100=max intensity).

5.5 DISSICIATION OF EMBRYONIC CELLS AND FLOW CYTOMETRY

Collections of 19-20h old *Drosophila* embryos were dechorionated with bleach (5%Chloride) for 2 min. For dissociation embryos were transferred into syringes filled with 1ml of 10x Trypsin-EDTA (0,5% Trypsin). For additional mechanical dissociation embryos were transferred 20x between two syringes (25G needle) and 10x between two syringes with 27G needles. Debris was reduced by filtering dissociated cell via cell strainer. The cells were stained for 5 min with DRAQ5 (1mg/ml) and DAPI (1mg/ml). Subsequently cells were sorted into 96-well plates containing 5 µl of Smart Seq2 lysis buffer at 4°C by a BD FACS Aria III flow cytometer equipped with 405nm, 488nm, 561nm and 633nm laser. Directly after cell sorting plates were shock frozen in nitrogen.

5.6 ScRNA SEQUENCING

A pooled cell population of Ok371>RFP positive motoneuronal cells and a pooled population of Mhc-TAU-GFP positive somatic muscle cells derived from FACS-sorted 19-20h old *Drosophila* embryos was used for scRNA Sequencing. The standard smart-seq2 protocol (Picelli et al., 2014) was modified by targeting low expressed Hox genes during the reverse transcription and PCR stages (*lab, Dfd, Scr, Antp, Ubx, abdA, AbdB*). Therefore, 1µM targeted primer mix of each *Hox gene* (See key resource table) was added to the PCR and RT buffers, respectively. The above described approach targets all expressed isoforms of each Hox gene. Method specific biases were ruled out by quality checks via Bioanalyzer of the whole transcriptome after SmartSeq2 preparation. Libraries of scRNA transcriptomes were created using a home-made Tn5 transposase (Hennig et al., 2018) and seventeen 96 well plates were sequenced with 75bp single-end on an Illumina NextSeq platform.

5.7 RAW DATA ROCESSING; QUALITY CONTROL AND NORMALIZATION

Sequencing reads were demultiplexed and the Poly-A tail trimmed. Read count tables were generated by pseudo-alignment to the cDNA of the *Drosophila* transcriptome (BDGP6 ensemble) using Kallisto (Bray et al., 2016). For quality control, we used standard settings (Velten et al., 2017): Cell were removed unless they contained at least 10 reads for each of at least 500 genes, and genes were removed unless they were expressed in at least 5 cells with 10 reads each. Data was normalized and scaled using the indexplorer pipeline (Velten et al., 2017). Posterior odds ratio scaling (Velten et al., 2017) was used to scale variance according to an estimate of true biological variance, as opposed to technical variance.

5.8 CLUSTERING AND DIMENSIONALITY REDUCTION

Following the indexplorer pipeline, PCA was computed using scaled, normalized expression values of all genes. t-SNE was then computed on the first ten principal components based on a visual inspection of the PCA Elbow plot. Alternatively, and following (H. Li et al., 2017), cells were clustered by hierarchical clustering using only the 20 most variably expressed genes of our scRNA data set; Ward linkage was used on an Euclidean distance metric.

5.9 INFERENCE OF SPATIAL POSITION FROM SINGLE GENE EXPRESSION DATA

To infer spatial position from single cell gene expression data, we first created a reference map of protein expression for seven *Hox* genes (Supplementary Figs 3a, 3b). To that end, immunofluorescence experiments were performed as described above.

Immunofluorescence data for gene g was thereby represented as a function of position $Y_g(x) \in (0,1)$ whereas single cell gene expression was represented as a matrix of read count values across genes and cells, $D_{g,c} \in \mathbb{N}_0$. We then assume that the probability of observing $D_{g,c}$ given that the position of cell c is really x_c depends on the expression of g at that position:

$$p(D_{g,c}|x_c) \sim \begin{cases} Pois\left(D_{g,c}|r_g * Y_g(x)\right) & \text{if } Y_g(x_c) > 0 \\ Pois(D_{g,c}|\lambda) & \text{if } Y_g(x_c) = 0 \end{cases}$$

Here, λ is a constant corresponding to a background number of reads observed in non-expressing cells, and r_g is a gene-wise scaling factor that estimates the average number of RNA-seq reads in a cell maximally expressing the protein. A maximum likelihood estimate of x_c is then obtained by computing

$$\widehat{x_c} = \arg\max_x \prod_g p(D_{c,g}|x)$$

For the analysis shown in the manuscript, λ was set to 0.1 (Kharchenko et al., 2014) and r_g was set to the mean gene expression in expressing cells:

$$r_g = \frac{\sum_c D_{c,g}}{\sum_c D_{c,g} > 0}$$

We used Latin Hypercube Sampling across a wide range of values for λ and r to demonstrate that the position estimate was independent of parameter choice (not shown).

To identify genes with spatially variable expression, B-Spline models with 3 degrees of freedom were fitted through scaled and normalized gene expression values for each gene individually, using inferred position as the only covariate. These models were then compared to null models not containing the position term using the Bayesian Information Criterion, similar to the workflow for selecting genes with variable expression over pseudotime described in (Velten et al., 2017).

For visualization and clustering, expression values for all variably expressed genes were arranged by inferred position and a floating mean was computed by 1D-convolution with an absolute exponential kernel with decay rate 10. Smoothened gene expression values obtained thereby were then compared to immunofluorescence images for model validation (Fig. 1E) or used for clustering of gene into modules with coherent expression patterns over space (Fig. 1f).

5.10 ZiNB-WAVE ANALYSIS

To identify gene whose expression was variable but statistically independent from the AP axis, we made use of the ZiNB-WAVE model (Risso et al., 2019). Unlike a PCA, ZiNB-WAVE separately estimates a matrix of known-covariate factors as well as a matrix of unknown-covariate factors; furthermore, ZiNB-WAVE uses a zero-infalted negative binomial distribution to account for the sparse nature of scRNA-Seq data. We ran ZiNB-WAVE using default settings and the number of genes observed as well as the inferred AP position as known sample-level covariates. While the PCA of the dataset was dominated by effects related to the number of genes observed, viability dye incorporation, and AP position (see Supplementary Fig. 1), the unknown-covariate factors from ZiNB-WAVE arranged genes according to their DV expression.

5.11 ChIP-SEQ REANALYSIS AND iREGULON ANALYSIS

To identify the reletionship between motoneuronal genes expressed in our scRNA-Seq data set and genes bound by the homeodomain TF Dfd, we used a whole embryo Dfd ChIP (Sorge et al., 2012) to investigate if Ig domain proteins are overrepresented among putative Dfd-regulated

genes compared to other genes expressed in MNs. Genes were classified as regulated by Dfd, if Dfd is bound to the target gene, max. 1kb upstream of the promotor region (Sorge et al., 2015). To calculate a p-value for enrichment, we performed a hypergeometric test.

The iRegulon (Janky et al., 2014) analysis was used to identify TF motifs enriched in the vicinity of Ig encoding genes expressed in our motoneuronal scRNA data set. Therefore, we performed an iRegulon analysis on all Ig domain encoding proteins expressed in MNs under standard settings (9,713PWM; 5kb upstream, 50 UTR and first intron with standard cutoff) to identify the 15 highest ranked TF motifs. Then we used iRegulon to investigate which TF families are predicted to bind to these motifs and chose 3 of the 5 motifs predicted to be regulated by homeodomain TFs.

5.12 DATA VISUALIZATION

All plots were generated using the ggplot2 (v. 3.2.1) and pheatmap (v. 1.0.12) packages in R 3.6.2. Boxplots are defined as follows: The middle line corresponds to the median; lower and upper hinges correspond to first and third quartiles. The upper whisker extends from the hinge to the largest value no further than 1.5 * IQR from the hinge (where IQR is the inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 * IQR of the hinge. Data beyond the end of the whiskers are called "outlying" points and are plotted individually.

5.13 CODE AVAILABILITY

Most analyses were done using the indexplorer software for interactive exploration of scRNA-Seq datasets (Velten et al, 2017), which is available from https://git.embl.de/velten/indeXplorer . Custom scripts for inference of spatial position are available from the corresponding authors upon request.

5.14 DATA AVAILIBILITY

RNA-seq data is available for interactive exploration at https://ie-collaborators.shiny.embl.de/?who=allNeurons (for the neuronal dataset) or https://ie-collaborators.shiny.embl.de/?who=allMuscle (for the muscle dataset). Raw data and count tables are available from gene expression omnibus, ID: GSE155578 (for the motoneuronal data set) and ID: GSE155586 (for the muscle data set) reviewer access for both data sets: svgvmiqijvobfwz.

6. RESOURCES

Reagents or Resources	source	identifier
Antibodies		
mouse anti-FasII	DSHB	
rat anti-Myosin	abcam	
guinea pig anti-Dfd	Katrin Domsch	
guinea pig anti-Ubx	Katrin Domsch	
rabbit anti-lab	Katrin Domsch	
mouse anti-Scr	DSHB	
mouse anti-Abd A	DSHB	
rabbit anti-Abd B		
mouse anti-Antp	DSHB	
rabbit anti-Hth		
rabbit anti-GFP	invitrogen	
rat anti-vvl		
rat anti-ELAV	DSHB	Cat# 7E8A10; RRID:AB_528218
Chemicals, Peptides, and Recombinant Proteins		
Paraformaldehyde	Sigma Aldrich	
VectaShield + DAPI	Vector Laboratories	
SMARTScribe™ Reverse Transcriptase	clonetech	Cat. #: 639537
5xSMART First Strand buffer	clonetech	
KAPA HiFi HotStart ReadyMix	fisherscientific	Cat. #: NC0295239
Ampure XP beads	beckman	A63880
DTT		
RNAse Inhibitor		
Tn5	PepCore EMBL	
Deposited Data		
Experimental Models: Organisms/Strains		

UAS-mCD8-GFP (y[1] w[*]; P{w[+mC] = UASmCD8::GFP.L}LL5, P{UAS- mCD8::GFP.L}2)	BDSC	5137
UAS-EGFP RNAi (y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8] = VALIUM20-EGFP.shRNA.3}attP2)	BDSC	41560
Frq1MI00709-GFSTF (y[1]w[*]; Mi{PT-GFSTF.0}Frq1MI00709-GFSTF.0)	BDSC	60284
mirr-GFP (y[1]w[*]; P{mirr-GFP.FPTB}attP40)	BDSC	68183
UAS-mirr RNAi (y[1] sc[*] v[1]; P{TRiP.HMC06139}attP2)	BDSC	65877
UAS-hth RNAi (P{y[1] sc[*] v[1]sev21; TRiP.HMS01112}attP2)	BDSC	34637
UAS-DIP gamma RNAi (y[1] sc[*] v[1]; P{TRiP.HMS06062}attP40)	BDSC	80461
UAS-DIP kappa RNAi (y[1] sc[*] v[1]; P{TRiP.HM04050}attP2	BDSC	31740
UAS-dpr1 (y[1] w[*]; P{UAS-dpr1.N}2/CyO)	BDSC	25081
UAS-Dfd RNAi	Sebastian Sorge (designed for this study)	
UAS-Ubx-HA	Domsch et al. 2020	
UAS-lab	Katrin Domsch	
UAS-mcd8-RFP	Pedro Pinto	
Hb9-Gal4	Broihier and Skeath, 2002	
Mef2-Gal4	Ranganayakulu et al., 1998	
elav-Gal4;Mef2-Gal4	designed for this study	
Ok6-Gal4	Aberle et al.,2002	
Ok371-Gal4	Mahr and Aberle, 2006	

elav-Gal4	Luo et al., 1994	
MHC-TAU-GFP	Chen and Olson,2001	
Oligonucleotides		
oligo-dt	AAGCAGTGGTATCAACGCAGAGTACT30V N	Sigma Albrich
TSO	AAG CAG TGG TAT CAA CGC AGA GTA CAT rGrG+G	
IS PCR primer	AAGCAGTGGTATCAACGCAGAGT	
targeted_Dfd_RT_rv	TCG GAT TGT TGC TGT TGA AG	
targeted_Ubx_RT_rv	CAG AAT TTT GCT CGC ATT CA	
targeted_AbdA_RT_rv	CAT GCG TTG CTC TAT CAA A	
targeted_AbdB_RT_rv	AAT ATA ATG CTC GGG GCA AA	
targeted_Scr_RT_rv	ATT GGG CGA TAC AAA CGA AG	
targeted_Lab_RT_rv	CCC TTC AAC TTT GCT TGC TC	
targeted_Antp_RT_rv	AAC CAT ACC CAG TCC ACC AA	
targeted_Dfd_IS_fw	AAG CAG TGG TAT CAA CGC AGA GTC CCT GGA TGA AGA AGA TCC A	
targeted_Ubx_IS_fw	AAG CAG TGG TAT CAA CGC AGA GTA AGG AGC TGA ACG AAC AGG A	
targeted_AbdA_IS_fw	AAG CAG TGG TAT CAA CGC AGA GTC TGG ACA AGA GCA ATC ACG A	
targeted_AbdB_IS_fw	AAG CAG TGG TAT CAA CGC AGA GTC GGA TTC GAT TTT AGC AAA TG	
targeted_Scr_IS_fw	AAG CAG TGG TAT CAA CGC AGA GTT CGA ATG CAA CTT GTT CTG C	
targeted_Lab_IS_fw	AAG CAG TGG TAT CAA CGC AGA GTC CCT GAT AAT GGC GAA CAG T	
targeted_Antp_IS_fw	AAG CAG TGG TAT CAA CGC AGA GTA GAG GAA CAG CAA AGC GAA A	
targeted_Dfd_IS_rv	AAG CAG TGG TAT CAA CGC AGA GTT CGG ATT GTT GCT GTT GAA G	
targeted_Ubx_IS_rv	AAG CAG TGG TAT CAA CGC AGA GTC AGA ATT TTG CTC GCA TTC A	
targeted_AbdA_IS_rv	AAG CAG TGG TAT CAA CGC AGA GTC ATG CGT TGC TCT ATC AAA	
targeted_AbdB_IS_rv	AAG CAG TGG TAT CAA CGC AGA GTA ATA TAA TGC TCG GGG CAA A	

targeted_Scr_IS_rv	AAG CAG TGG TAT CAA CGC AGA GTA TTG GGC GAT ACA AAC GAA G	
targeted_Lab_IS_rv	AAG CAG TGG TAT CAA CGC AGA GTC CCT TCA ACT TTG CTT GCT C	
targeted_Antp_IS_rv	AAG CAG TGG TAT CAA CGC AGA GTA ACC ATA CCC AGT CCA CCA A	
Software and Algorithms		
Fiji (ImageJ)	Schindelin et al. (2012)	https://imagej.net/Fiji
IndeXplorer	Velten et al. (2017)	
iRegulon	Janky et al. (2014)	

7. SUPPLEMENTARY

Supplementary table 1: Regional specific gene expression patterns along the AP axis

Genes with significant variation along the AP axis were identified and clustered into 10 groups of distinct expression pattern (Methods).

Gene	ID	cluster
abd-A	(FBgn0000014)	7
Ald	(FBgn0000064)	7
Eno	(FBgn0000579)	7
flw	(FBgn0000711)	7
Gapdh1	(FBgn0001091)	7
Gapdh2	(FBgn0001092)	7
Syt1	(FBgn0004242)	7
msi	(FBgn0011666)	7
Syx1A	(FBgn0013343)	7
Pglym78	(FBgn0014869)	7
Lim1	(FBgn0026411)	7
CG7781	(FBgn0032021)	7
Hasp	(FBgn0032797)	7
nrv3	(FBgn0032946)	7
CG1358	(FBgn0033196)	7
CG13287	(FBgn0035643)	7
CG10960	(FBgn0036316)	7
tey	(FBgn0036899)	7
cpx	(FBgn0041605)	7
CG31324	(FBgn0051324)	7
Nlg3	(FBgn0083963)	7
Tpi	(FBgn0086355)	7
Pgk	(FBgn0250906)	7
kcc	(FBgn0261794)	7
Vha68-1	(FBgn0265262)	7
Dscam2	(FBgn0265296)	7
Bx	(FBgn0265598)	7
PyK	(FBgn0267385)	7
Fas1	(FBgn0285925)	7
Abd-B	(FBgn0000015)	9
ImpL3	(FBgn0001258)	9
Wnt4	(FBgn0010453)	9
otk2	(FBgn0267728)	9
D	(FBgn0000411)	1
Ddc	(FBgn0000422)	1

fkh	(FBgn0000659)	1
lab	(FBgn0002522)	1
Pu	(FBgn0003162)	1
heph	(FBgn0011224)	1
Dop1R1	(FBgn0011582)	1
Mef2	(FBgn0011656)	1
mbc	(FBgn0015513)	1
Rfx	(FBgn0020379)	1
Rab27	(FBgn0025382)	1
Hsc70Cb	(FBgn0026418)	1
scro	(FBgn0028993)	1
CG4577	(FBgn0031306)	1
erm	(FBgn0031375)	1
Ttc19	(FBgn0032744)	1
CG13739	(FBgn0033403)	1
Or45a	(FBgn0033404)	1
CG2269	(FBgn0033484)	1
Mapmodulin	(FBgn0034282)	1
twz	(FBgn0034636)	1
unc-13-4A	(FBgn0035756)	1
Dhpr	(FBgn0035964)	1
ssp2	(FBgn0036389)	1
CG13408	(FBgn0038929)	1
Sox102F	(FBgn0039938)	1
Sox21b	(FBgn0042630)	1
Lmx1a	(FBgn0052105)	1
CG32281	(FBgn0052281)	1
CG32683	(FBgn0052683)	1
Ddr	(FBgn0053531)	1
bma	(FBgn0085385)	1
CG42541	(FBgn0260658)	1
Sytalpha	(FBgn0261089)	1
Sytbeta	(FBgn0261090)	1
rdgA	(FBgn0261549)	1
CR43212	(FBgn0262848)	1
CG43373	(FBgn0263131)	1
CG43689	(FBgn0263772)	1
Hmx	(FBgn0264005)	1
Ca-alpha1T	(FBgn0264386)	1
rad	(FBgn0265597)	1
CG45781	(FBgn0267428)	1
Dfd	(FBgn0000439)	2
noc	(FBgn0005771)	2

Akap200	(FBgn0027932)	2
btsz	(FBgn0266756)	2
br	(FBgn0283451)	2
eag	(FBgn0000535)	5
CG17716	(FBgn0000633)	5
tj	(FBgn0000964)	5
Pfk	(FBgn0003071)	5
shi	(FBgn0003392)	5
tsh	(FBgn0003866)	5
Ubx	(FBgn0003944)	5
ush	(FBgn0003963)	5
scrt	(FBgn0004880)	5
comm	(FBgn0010105)	5
bol	(FBgn0011206)	5
HLH3B	(FBgn0011276)	5
Nckx30C	(FBgn0028704)	5
fend	(FBgn0030090)	5
Frq1	(FBgn0030897)	5
REPTOR-BP	(FBgn0032202)	5
sky	(FBgn0032901)	5
dpr13	(FBgn0034286)	5
CG9896	(FBgn0034808)	5
DOR	(FBgn0035542)	5
chrb	(FBgn0036165)	5
Sox21a	(FBgn0036411)	5
CG9821	(FBgn0037636)	5
Octbeta2R	(FBgn0038063)	5
plum	(FBgn0039431)	5
CG33199	(FBgn0053199)	5
Antp	(FBgn0260642)	5
app	(FBgn0260941)	5
Rbp6	(FBgn0260943)	5
hkb	(FBgn0261434)	5
ps	(FBgn0261552)	5
Bsg	(FBgn0261822)	5
Glut4EF	(FBgn0267336)	5
grn	(FBgn0001138)	6
CG34347	(FBgn0085376)	6
Dop1R2	(FBgn0266137)	6
hth	(FBgn0001235)	3
Scr	(FBgn0003339)	3
vvl	(FBgn0086680)	3
CG43222	(FBgn0262858)	3

otk	(FBgn0004839)	8
Uch-L5	(FBgn0011327)	8
CG11306	(FBgn0037108)	8
CG34155	(FBgn0083991)	8
CG34351	(FBgn0085380)	8
Xe7	(FBgn0010772)	10
Frq2	(FBgn0083228)	10
pre-rRNA:CR45856	(FBgn0267516)	10
nej	(FBgn0261617)	4

Supplementary table 2: Homeo-code across 60 clusters

Expression of homeodomain encoding genes across n= 758 single *twit*^{low} MNs. 60 Clusters are identified by hierarchical clustering and ordered according to inferred AP position. Normalized mean expression levels and median expression levels of each gene across single clusters.

Cluster A-P	Gene	cluster	mean	median
1	zfh2	23	12.16881329	18.03684598
1	Lim1	23	17.22073646	17.73893009
1	NK7.1	23	12.6583943	14.98146937
1	lab	23	13.97117156	14.15838144
1	exd	23	11.31473925	13.34674456
1	hth	23	8.695438811	12.30980093
1	onecut	23	8.501889129	11.98682455
2	dve	41	18.21815575	18.60885007
2	ct	41	14.85088506	17.34068616
2	Dr	41	10.86167229	16.32825939
2	oc	41	14.16955877	15.96911852
2	hth	41	11.48789715	15.04557065
2	ap	41	9.523453674	13.70899082
2	lab	41	12.16386608	12.426291
2	onecut	41	7.248925116	6.511161282
3	pb	25	14.46620665	17.91680769
3	Hmx	25	15.98210402	17.13840107
3	hth	25	17.0766247	16.80104964
3	mirr	25	16.66524075	16.39063366
3	otp	25	13.70495122	16.26123295
3	zfh1	25	12.16135482	14.07865437
3	tup	25	8.607628993	12.99216643
3	lab	25	9.771876751	12.57718796
3	eyg	25	6.503635848	8.749429775
4	hth	4	18.28434269	18.27021733
4	toy	4	16.04944056	16.7295848
4	Lim1	4	16.19218266	16.18146407
4	zfh2	4	16.01733239	15.790712
4	vvl	4	14.88286846	15.29090857
4	unc-4	4	14.75803925	14.9229589
4	Dfd	4	10.85244231	13.37004399
4	exd	4	11.19688686	12.6428707
4	Dr	4	9.209226329	12.22074843
4	onecut	4	7.780407095	9.195592253

4	nub	4	7.410803519	5.530939437
4	eyg	4	4.877351288	4.18589048
5	hth	49	19.03536361	18.95337578
5	Lim3	49	17.34970121	17.46253334
5	Ptx1	49	15.53600452	15.7706025
5	zfh1	49	15.23236872	15.30955228
5	Drgx	49	11.3615762	13.60076782
5	exd	49	8.408106682	11.8848262
5	vvl	49	7.47121515	11.37560559
5	lab	49	9.895128308	11.37560559
6	mirr	40	17.70233109	17.97967787
6	vvl	40	17.57335575	17.80669895
6	zfh2	40	15.4458531	16.93554027
6	Lmx1a	40	16.88415285	16.87454963
6	hth	40	13.16756621	15.54124559
6	hbn	40	11.11426587	15.36525593
6	toy	40	12.52981785	14.53028695
6	Rx	40	13.9256519	14.08836562
6	exd	40	9.949303138	13.20475972
6	onecut	40	9.448031304	11.68803855
6	lab	40	8.719156479	11.41246169
7	hth	43	18.08006573	18.20062258
7	mirr	43	14.61076171	18.09234703
7	onecut	43	13.42793757	15.44373157
7	ro	43	9.061633979	14.37303825
7	exd	43	14.36099219	13.79721131
7	oc	43	8.662317664	12.66781694
7	zfh1	43	9.352471469	12.57425315
7	lab	43	9.916319574	12.04644794
7	eyg	43	10.98304418	10.9897637
7	Hmx	43	7.389611177	9.486239962
8	zfh2	36	17.58697389	18.125704
8	mirr	36	11.10680983	17.23556085
8	Lmx1a	36	16.85302258	16.79706017
8	CG4328	36	15.26685196	15.36417088
8	hth	36	11.01991538	13.86780099
8	vvl	36	10.19463084	13.72772061
8	onecut	36	10.05427172	13.37657297
8	NK7.1	36	11.42435197	12.85656568
8	lab	36	6.755352491	7.769607546
8	exd	36	7.161084503	5.940856786
8	Rx	36	6.03296132	4.020551169
9	mirr	57	17.61710856	17.74610486
9	hth	57	17.25675545	17.43005633
9	vvl	57	16.92046072	16.97270228
9	zfh2	57	14.04489211	16.7583809
9	nub	57	14.95270436	16.16320401
9	Dfd	57	16.03146509	16.1143939
9	B-H1	57	13.49068921	15.17474396
9	tup	57	12.37410013	15.17413317
9	pdm2	57	10.35146604	14.74716657
9	B-H2	57	13.961208	14.18904719
9	exd	57	8.934090744	11.49468129
9	onecut	57	6.68750056	6.150035659
9	NK7.1	57	6.57280635	5.812050232
10	hth	39	17.8599514	17.94969416

10	mirr	39	17.38504549	17.4728006
10	zfh2	39	17.38916624	17.43855368
10	pb	39	15.9043831	16.55823802
10	Dfd	39	16.20103226	16.08766439
10	exex	39	15.68474337	15.63105402
10	caup	39	15.23740791	15.07337309
10	vvl	39	14.17182429	14.42105455
10	zfh1	39	13.97475885	14.25153576
10	exd	39	13.8478479	13.65945965
10	ara	39	12.43820062	13.63679269
10	Lim3	39	8.988158715	12.62406665
10	onecut	39	10.01079699	11.23511461
10	lms	39	5.809162956	4.993225321
10	eyg	39	4.53224224	3.704290778
11	dve	18	18.04321182	18.14059353
11	Ptx1	18	16.97121992	16.99619074
11	tup	18	12.0871364	15.55402131
11	onecut	18	11.36010369	14.64875599
11	exd	18	13.40390885	13.23807096
11	zfh1	18	13.69206582	12.69468768
11	eyg	18	8.745310026	10.90053008
11	vvl	18	8.940961933	8.245115872
11	hth	18	7.457365516	6.090167689
11	lab	18	5.743787139	5.250011613
11	NK7.1	18	5.200389586	5.150767559
11	Vsx1	18	4.870986348	4.651338908
11	Ubx	18	4.870986348	4.651338908
12	nub	16	16.62178177	17.85606497
12	zfh2	16	17.53392849	17.55898241
12	pdm2	16	14.50146282	16.35460791
12	toy	16	13.52249131	15.70367757
12	exd	16	10.77808203	14.10072795
12	hth	16	11.77921555	13.82996797
12	Lim3	16	7.306401112	7.89344487
12	oc	16	7.743768175	6.898646244
12	onecut	16	6.229856678	4.998907614
12	lms	16	6.02651039	4.945909969
12	zfh1	16	7.087081321	4.787405192
12	NK7.1	16	6.643736261	4.04412825
13	hth	6	18.05473562	18.05375583
13	mirr	6	17.8279925	17.70721388
13	zfh2	6	17.12205919	17.18434433
13	Lim1	6	16.1572025	16.58165955
13	unc-4	6	13.17608695	15.01542072
13	toy	6	14.52799664	14.90353135
13	NK7.1	6	8.908604697	14.08915134
13	slou	6	8.441220446	13.07279865
13	Dfd	6	10.79524087	13.07279865
13	exd	6	6.794449925	8.278902515
14	hth	50	18.37650061	18.30401865
14	Lim1	50	16.98848122	16.69057411
14	ct	50	9.982817514	16.14466715
14	zfh1	50	16.07541396	15.94378087
15	dve	45	18.20206148	17.92624358
15	zfh2	45	14.78647896	17.73653914
15	oc	45	14.71189263	15.11086378

15	onecut	45	15.14919674	14.33345374
15	exd	45	12.52503074	14.20943455
15	zfh1	45	10.81075204	11.46667463
15	hth	45	9.69489348	10.47244935
15	Lim3	45	8.129527088	9.436578621
15	ap	45	7.386071401	6.256981367
15	NK7.1	45	7.185036319	5.266365411
15	Antp	45	5.471781508	4.33533532
16	zfh2	60	17.51707064	18.10362197
16	Lim1	60	17.59336244	17.48787151
16	hth	60	14.3691801	16.31334763
16	toy	60	11.0664523	15.32602502
16	lab	60	11.08865052	12.07081705
16	NK7.1	60	7.75727729	8.153645481
17	hth	24	18.36107649	18.24780055
17	vvl	24	16.81375012	16.91693345
17	mirr	24	15.10757424	16.69454297
17	Scr	24	14.7399652	16.65630454
17	zfh2	24	10.73741269	14.48526586
17	Lim1	24	9.526013677	12.47348402
18	hth	20	18.95006348	19.31572084
18	Lim1	20	16.86628841	16.94360707
18	exd	20	9.617611093	11.93914166
18	onecut	20	8.821289471	11.24929459
18	Dr	20	8.095813098	10.8827382
19	zfh2	19	12.02048054	18.58008037
19	hth	19	15.78497711	17.0885956
19	lab	19	11.35411292	13.53751476
19	exd	19	7.563784811	9.288205126
19	onecut	19	6.320462056	4.460960782
20	CG32532	56	16.35688819	17.03888158
20	vvl	56	17.09480668	16.89365821
20	zfh1	56	11.60974794	16.38285345
20	NK7.1	56	13.39030493	16.11237804
20	B-H1	56	14.3479413	14.7853752
20	toy	56	12.8826807	14.71683293
20	zfh2	56	10.91459347	14.36135283
20	onecut	56	8.172433172	7.84659553
20	hth	56	7.94668207	7.250290496
20	Antp	56	7.058904396	5.683030604
20	B-H2	56	6.867296475	3.753749554
20	exd	56	5.608267743	3.257702961
21	hth	59	17.3632336	17.46740076
21	mirr	59	17.38254558	17.35742142
21	zfh2	59	17.49181074	17.11838881
21	Scr	59	16.33074221	16.68365476
21	Lim3	59	13.33008132	16.16639567
21	ara	59	9.78133496	15.95708975
21	caup	59	15.67475106	15.87631012
21	zfh1	59	13.21311084	15.37882033
21	onecut	59	12.17968701	15.12861471
21	vvl	59	14.07753829	14.91350394
21	eyg	59	7.494384795	8.981683914
22	hth	32	17.4873644	18.36020736
22	zfh2	32	15.71214298	17.36962256
22	vvl	32	12.10313745	16.32994484

22	Lim1	32	12.50875098	14.86718497
22	Antp	32	13.29185993	14.55101867
22	zfh1	32	9.051640786	13.04039008
22	onecut	32	8.723095127	12.59195226
22	exd	32	8.210103082	11.07425683
23	hth	9	16.36962552	17.16582149
23	ct	9	16.69649971	16.84748285
23	zfh2	9	16.99927656	16.69952933
23	vvl	9	12.54310638	16.39986477
23	Scr	9	13.63705734	16.25536677
23	onecut	9	9.274277612	11.18987684
24	mirr	53	17.64404581	17.69076823
24	zfh2	53	17.38537564	17.27655936
24	vvl	53	16.92762478	17.01061371
24	Dfd	53	13.21643413	15.42403653
24	Ptx1	53	8.8626921	14.14623021
24	hth	53	11.21854607	13.65891038
24	Dr	53	8.88679931	13.2444797
24	NK7.1	53	9.189938984	12.50835009
24	zfh1	53	8.27554142	9.970833909
25	Lim1	29	17.88858303	17.81056084
25	Drgx	29	17.68191325	17.74538909
25	exd	29	12.35904517	14.97663502
25	zfh1	29	9.188790071	11.1651164
25	ct	29	9.181100766	8.668922284
25	toy	29	8.545374123	7.46427209
25	onecut	29	7.775788609	7.427400831
25	hth	29	7.142153771	6.427473861
25	lab	29	6.551940124	5.606564945
25	pros	29	5.614267281	4.446815266
26	zfh2	28	11.29106214	18.43174347
26	otp	28	14.62075565	18.10375467
26	toy	28	10.70440389	17.16750755
26	NK7.1	28	15.87893983	17.16017551
26	onecut	28	12.84403555	15.78725852
26	exd	28	11.63662263	15.31269419
26	eyg	28	7.927494487	10.10078414
26	lab	28	6.58294946	5.370676144
27	hth	26	18.18823653	18.60910128
27	Antp	26	16.08146019	16.86565255
27	mirr	26	16.33415302	16.72755719
27	zfh2	26	16.38739058	16.08362483
27	Lim1	26	13.02420707	15.6099023
27	ct	26	14.94465893	15.20030681
27	vvl	26	13.88930775	13.59574898
27	NK7.1	26	14.34173889	13.58186794
27	oc	26	13.64772867	13.47919514
27	onecut	26	10.98938338	12.96032496
27	exd	26	8.545256377	10.02967105
27	Lim3	26	6.521212351	8.354425632
27	dve	26	8.265191478	7.75355086
28	hth	55	18.18081382	18.48840699
28	zfh2	55	15.82866984	16.03144644
28	Lim3	55	15.70486046	15.98756659
28	lab	55	8.541388356	12.06175221
28	zfh1	55	10.65441318	11.50374793

28	Scr	55	7.617453076	7.248833526
28	NK7.1	55	7.759487888	6.018547063
29	hth	44	17.84037331	18.44667254
29	zfh2	44	16.74595074	17.35367437
29	otp	44	16.34201019	16.09667573
29	onecut	44	15.52199608	15.56135324
29	vvl	44	9.279987106	14.93161473
29	Antp	44	15.06481955	14.7616955
29	Lmx1a	44	11.37427937	13.59587343
29	exd	44	8.004480006	12.51193966
30	Antp	35	13.60480177	17.32491978
30	zfh2	35	16.83999004	16.81710579
30	NK7.1	35	14.36685443	15.56933364
30	hth	35	8.45464446	10.94078087
31	hth	14	16.77150592	17.75289498
31	mirr	14	12.86465156	17.30311335
31	zfh2	14	14.41481948	16.83284697
31	vvl	14	15.96042937	15.97230822
31	Lim1	14	15.82354467	15.92042252
31	dve	14	14.02729572	15.55936527
31	oc	14	11.37284636	13.01428312
31	Ubx	14	8.882632103	12.25186309
31	Antp	14	8.276140881	11.30460516
31	exd	14	7.649917058	11.19679653
31	lab	14	5.45128799	7.744785015
32	hth	15	15.10487648	17.38618963
32	Lim3	15	13.90363291	15.72256897
32	exd	15	10.80451252	13.23775375
32	Abd-B	15	7.611491244	4.980736377
32	ct	15	7.002297618	4.598390904
33	Ubx	5	17.52143754	18.09244238
33	zfh2	5	17.70361321	17.86108035
33	Dr	5	13.99294236	16.30922299
33	Lim1	5	15.70701084	15.84751859
33	Antp	5	15.8050256	15.81001719
33	zfh1	5	11.06265881	13.82941633
33	hth	5	10.77999646	13.43064858
33	exd	5	10.21443119	13.24849146
33	ey	5	9.430454513	13.05738054
33	onecut	5	7.497571667	9.278392015
34	Ubx	8	16.74413092	18.22614325
34	mirr	8	17.08740184	17.27085407
34	Dr	8	13.84614477	16.32368685
34	zfh2	8	12.57910221	15.65959062
34	Lim1	8	13.22394313	15.39803863
34	vvl	8	12.92002765	15.35118451
34	hth	8	10.43062258	13.83209517
34	Antp	8	8.240946996	12.75435785
34	exd	8	9.354740009	11.20242947
35	Antp	52	17.19721241	17.84651774
35	Ubx	52	13.67541317	17.37329206
35	mirr	52	17.05138872	17.30642588
35	vvl	52	16.30471435	16.17216957
35	zfh2	52	13.04583248	15.67539736
35	ara	52	12.13379323	15.6396813
35	exex	52	11.93925234	15.55399687

35	Lim3	52	13.32066899	14.4530417
35	exd	52	8.902937002	12.75062567
35	onecut	52	8.533196581	10.98303842
35	lab	52	7.251747292	9.75209031
35	caup	52	7.443940918	7.667511053
36	Antp	13	13.88180577	18.02952463
36	unpg	13	15.08263475	16.63280775
36	mirr	13	15.25064208	16.62557674
36	Ubx	13	13.34389462	16.37407627
36	zfh2	13	16.17723881	16.17158596
36	zfh1	13	13.81909954	15.71254083
36	hth	13	13.83697842	15.24856546
36	vvl	13	14.03543456	15.16342356
36	abd-A	13	8.338615785	13.74679818
36	onecut	13	8.268511667	11.78121578
36	Abd-B	13	8.457527713	9.316824317
37	dve	54	18.12965906	18.38399401
37	hth	54	16.54011636	17.26211606
37	ct	54	14.52565243	15.928546
37	Ubx	54	11.08364685	15.68564207
37	zfh1	54	14.29560264	15.50305762
37	Antp	54	14.15469361	15.4908524
37	Lim1	54	13.88560878	15.27609214
37	oc	54	13.51585448	14.44491828
37	NK7.1	54	9.783926977	13.17410571
37	exd	54	8.705851891	11.87195468
37	nub	54	8.011122836	10.77859116
37	abd-A	54	8.310732917	7.813107654
37	onecut	54	6.044796372	4.192172256
38	Ubx	10	14.98649929	17.59989976
38	toy	10	17.10725464	16.87132972
38	zfh2	10	15.73372727	16.4370826
38	abd-A	10	12.17619593	15.99176984
38	Dr	10	15.82031316	15.99093416
38	Antp	10	15.37717713	15.7704745
38	vvl	10	12.15155156	15.51409846
38	hth	10	11.16702166	14.40043519
38	Lim1	10	11.9050657	14.25861932
39	pb	46	18.20093516	18.17606585
39	Ubx	46	16.86418078	17.33027128
39	Lim1	46	15.52187879	16.55910898
39	Dr	46	15.25282196	16.48454785
39	abd-A	46	12.33082646	16.41135843
39	zfh2	46	16.18549214	16.30987607
39	hth	46	10.21675376	15.03172311
39	ey	46	12.86284077	14.82177494
39	NK7.1	46	9.110911639	13.64238149
39	exd	46	9.254441027	11.2588609
39	eyg	46	4.519157206	3.98123585
39	onecut	46	5.959790687	3.690199164
40	hth	7	16.34065162	16.93681672
40	Antp	7	15.13758158	16.71721908
40	Lim1	7	16.62428883	16.46399311
40	Ubx	7	12.94878596	16.27459458
40	zfh1	7	13.96509479	16.10271941
40	Dr	7	15.32249971	15.9056488

40	ey	7	15.11342453	15.59963916
40	onecut	7	10.74763788	15.34404831
40	vnd	7	12.785854	14.08031924
40	abd-A	7	9.132124469	12.47239084
40	exd	7	7.931253795	10.82907403
41	zfh2	48	14.89853114	17.74084304
41	Ubx	48	13.4188331	17.04366215
41	abd-A	48	13.70414986	15.74329134
41	zfh1	48	11.91835661	15.41300024
41	hth	48	13.1901287	15.38761961
41	Antp	48	12.3873822	13.48709341
42	zfh2	11	17.36659535	17.33057357
42	Ubx	11	16.5698788	16.99878442
42	mirr	11	16.47102703	16.70885724
42	Antp	11	12.93953325	16.53602235
42	zfh1	11	14.9140331	16.23612522
42	Lim1	11	14.1959748	15.34216177
42	hth	11	11.71853877	15.25564841
42	oc	11	11.885938	14.76743167
42	vvl	11	10.64008795	14.61598419
42	abd-A	11	11.545431	14.52763043
43	dve	33	17.25529063	17.65812831
43	mirr	33	15.72918243	17.25128868
43	Ubx	33	17.03239045	17.04787347
43	abd-A	33	13.33745423	16.19433023
43	Antp	33	13.54683922	16.12372906
43	zfh2	33	12.87343941	15.96641827
43	Lim1	33	15.90005657	15.86643257
43	ct	33	13.71834488	14.9873643
43	oc	33	14.19203742	14.43360036
43	NK7.1	33	10.63015474	14.0715824
43	onecut	33	9.696874239	14.00746874
43	hth	33	7.035213746	6.241805155
44	Lim1	1	16.82968774	17.47068295
44	Ubx	1	14.35230022	16.72115409
44	ct	1	11.28230415	15.43728395
44	abd-A	1	14.9453173	14.7616955
44	mirr	1	13.60427179	14.46726835
45	pb	38	18.11110888	18.11110888
45	unpg	38	18.09367156	18.09367156
45	Antp	38	16.98002167	16.98002167
45	Lim3	38	16.02888746	16.02888746
45	abd-A	38	15.49923541	15.49923541
45	zfh2	38	15.18800659	15.18800659
45	Ubx	38	14.68880026	14.68880026
45	zfh1	38	14.67890454	14.67890454
45	ct	38	14.60211843	14.60211843
45	Vsx1	38	11.92899925	11.92899925
45	lab	38	10.89841361	10.89841361
45	Abd-B	38	7.795586678	7.795586678
45	dve	38	5.657792936	5.657792936
45	lms	38	3.827279328	3.827279328
45	pros	38	3.17166551	3.17166551
46	Ubx	27	17.4815775	17.58891757
46	mirr	27	17.16667934	17.31828178
46	ct	27	14.7336573	16.7495472

46	Lim1	27	14.79946778	16.01019721
46	vvl	27	14.92871761	15.88603686
46	abd-A	27	12.64195135	15.17320262
46	Antp	27	14.41306835	15.17186299
46	oc	27	13.39395301	13.9726867
46	NK7.1	27	7.844178254	11.11579614
46	onecut	27	9.822629203	10.91734873
46	exd	27	7.831098725	10.44682382
46	hth	27	7.814976999	9.386718692
47	hth	17	17.4100535	17.71774945
47	zfh2	17	15.58519476	17.33691335
47	Ubx	17	16.67692571	16.8278995
47	Lim1	17	16.315157	16.53471753
47	abd-A	17	16.09450949	16.11757769
47	Antp	17	11.72039395	14.99236572
47	oc	17	8.380083985	12.77966205
47	onecut	17	7.460131082	5.604685224
48	zfh2	30	16.66101302	17.73137852
48	Ubx	30	16.01230857	17.41946486
48	Lim1	30	16.57867612	16.97434209
48	zfh1	30	12.47546932	14.80803258
48	hth	30	12.42623298	14.69800317
48	Abd-B	30	10.01434153	14.02958126
49	mirr	47	15.64392105	17.03254376
49	Ubx	47	15.69660541	17.00758613
49	hth	47	16.26177809	16.46077902
49	abd-A	47	15.24454503	16.3775554
49	zfh1	47	15.67169451	16.10373983
49	Lim1	47	12.22409521	15.64265259
49	ct	47	11.04829511	13.05735494
49	Antp	47	9.170747134	12.8921171
49	exd	47	7.579477995	10.93377409
49	onecut	47	8.397522208	9.546735647
49	NK7.1	47	7.01584096	5.918848274
49	vvl	47	7.073621959	4.516352325
50	vvl	3	17.650424	17.43888825
50	ap	3	16.88442283	16.81151663
50	zfh2	3	15.9238964	16.45554917
50	slou	3	15.76722854	16.20058731
50	abd-A	3	10.67027998	15.46051859
50	onecut	3	15.39642881	14.79952044
50	exd	3	14.49490367	14.54387538
50	Ubx	3	13.51413553	14.32430859
50	nub	3	8.932197874	13.03167913
51	abd-A	12	16.90142462	17.51694109
51	vvl	12	14.59118207	17.47986517
51	Ptx1	12	16.2484687	17.43559099
51	Abd-B	12	11.45511212	16.51203471
51	Lim3	12	13.57344962	15.60705527
51	NK7.1	12	10.89410019	15.07764813
51	onecut	12	7.395272892	7.141012766
51	hth	12	7.137178093	6.236327622
51	exd	12	6.113480242	4.688377698
52	zfh2	34	17.97964504	18.14274314
52	abd-A	34	14.32381486	16.72870399
52	slou	34	16.13473229	16.70905204

52	Lim1	34	16.65746943	16.51957537
52	hth	34	15.24916578	16.51224265
52	unc-4	34	15.99760759	16.40550426
52	Ubx	34	10.52557693	14.34608036
52	zfh1	34	13.47011117	14.27033096
52	Antp	34	8.459075912	10.3709655
52	onecut	34	6.688758497	6.125840377
52	lab	34	4.93962275	4.514473002
53	Ubx	42	16.65044134	17.39252226
53	Antp	42	15.46064008	17.35459862
53	zfh2	42	15.85995856	16.68365476
53	abd-A	42	16.431275	16.60135566
53	Lim1	42	16.1026124	16.02469961
53	ey	42	13.44897552	16.02031236
53	Dr	42	12.75727709	15.58996781
53	zfh1	42	12.77638826	15.43574609
53	Abd-B	42	8.282134127	12.30249757
53	vnd	42	7.802110741	11.72059448
53	hth	42	7.319953741	9.981711416
54	vvl	31	16.82234194	17.34508888
54	Dr	31	16.87491584	17.19075342
54	Ubx	31	12.35228167	17.05711003
54	zfh2	31	16.87196147	16.95624197
54	mirr	31	16.84749755	16.78325861
54	Abd-B	31	15.88717504	16.76745774
54	Lim1	31	16.07104879	16.24308706
54	abd-A	31	15.46157109	16.06523394
54	Antp	31	14.28864927	14.70537096
55	zfh2	2	17.40128151	17.85722426
55	Abd-B	2	16.38530485	17.4746216
55	mirr	2	14.34802808	16.4992928
55	abd-A	2	13.10111539	16.12052887
55	Lim3	2	14.19471248	15.80418561
55	vvl	2	10.17671273	14.99466493
55	hth	2	13.86255736	14.91341514
55	NK7.1	2	8.443342365	10.80954465
56	Abd-B	51	16.01462415	17.58732293
56	Lim1	51	17.03894852	17.28814623
56	zfh2	51	16.52253426	16.94888717
56	Dr	51	15.12949007	16.64778367
56	abd-A	51	13.5084363	14.47794817
56	ey	51	11.53855795	13.64046173
56	pb	51	8.599856045	7.869551094
56	onecut	51	7.976972942	7.52740085
57	mirr	22	15.75718013	17.09430838
57	ct	22	16.77208565	17.07546946
57	Ubx	22	16.42487661	16.67304299
57	Antp	22	12.92172951	16.47470172
57	Abd-B	22	15.87888406	16.25429264
57	abd-A	22	13.15784297	15.75150139
57	vvl	22	11.48821921	15.64851368
57	Lim1	22	14.38520269	15.56042681
57	oc	22	14.5839187	14.95551647
57	hth	22	9.691817493	9.39261997
57	NK7.1	22	7.609218288	7.007841475
58	zfh2	37	18.26289169	18.22692077

58	Abd-B	37	16.39767436	17.37845062
58	vvl	37	14.52027044	17.25759697
58	Lim1	37	11.38729246	15.90380894
58	onecut	37	12.77730492	15.45355342
58	oc	37	11.00906401	15.43974793
58	abd-A	37	11.30254465	15.26994983
59	mirr	21	14.87930462	17.22196134
59	zfh2	21	16.46523965	16.91461526
59	Lim1	21	16.71845447	16.75891901
59	abd-A	21	15.98268104	16.25066163
59	Abd-B	21	12.54183583	16.191079
59	slou	21	13.69181598	16.00404046
59	Ptx1	21	13.48185006	15.55647591
59	toy	21	11.21110108	14.84313271
59	unc-4	21	14.47326772	14.21739879
59	Antp	21	7.526191601	12.51327895
59	eyg	21	7.739514531	10.43509179
59	hth	21	9.168793698	8.678884404
59	onecut	21	6.813465818	8.472112625
60	abd-A	58	17.99643352	18.19739513
60	Abd-B	58	14.20676295	17.3410994
60	Lim1	58	15.76744164	15.87229978
60	onecut	58	10.42657928	14.4439201
60	Dr	58	9.958921026	13.55069713
60	zfh1	58	11.71438705	13.50625084
60	Antp	58	9.336968346	11.55307006
60	hth	58	7.165361809	6.11571445
60	vvl	58	7.223808073	5.602129844

Supplementary table 3: Ig-code across 60 clusters

Expression of Ig encoding genes across n= 758 single *twit*^{low} MNs. 60 Clusters are identified by hierarchical clustering and ordered according to inferred AP position (identical to homeo-code clusters in supplementary table 2). Normalized mean expression levels and median expression levels of each gene across single clusters.

Cluster A-P	Gene	cluster	mean	median
1	CG17716	23	12.49931763	15.41922769
1	Dscam2	23	14.18583324	15.18269806
1	Ptp99A	23	12.80387067	14.38059352
1	klg	23	9.083645397	13.91119812
1	kek2	23	11.60935776	13.88836296
1	CG42313	23	10.56292403	13.84913158
1	CG34114	23	8.8970402	13.73612349
1	dpr1	23	8.095713274	13.45277316
1	dpr6	23	7.722904375	13.37017821
1	beat-VII	23	8.528919022	13.27382022
1	beat-IIa	23	10.26708815	13.13422191
1	dpr8	23	11.71861267	13.09006714
1	kek1	23	8.780218497	12.97794554
1	Fas2	23	10.69608711	12.70582807
1	side	23	8.594958136	12.52525398
1	hbs	23	10.50814682	12.44222272
1	CG34371	23	10.13000248	12.26436459
1	dpr11	23	7.528870174	11.7478212
1	ed	23	8.006629219	11.27751383

1	dpr9	23	7.726461222	11.18953024
1	dpr2	23	7.512570919	10.27254137
1	Dscam3	23	8.714506889	10.09644204
1	fred	23	7.037635304	9.088594534
1	kirre	23	6.141637229	9.088594534
1	Dscam4	23	6.627475454	7.111917183
2	kek2	41	11.17643302	15.19092541
2	side	41	12.29235612	14.81458003
2	Ptp99A	41	15.10118817	14.73546134
2	beat-IIa	41	14.48458156	14.28057994
2	kek1	41	11.07324245	14.2652495
2	dpr8	41	10.84228964	14.11475767
2	robo3	41	12.08184008	13.77724425
2	klg	41	9.076830875	13.7527439
2	CG34353	41	12.020843	13.43407748
2	CG17716	41	10.51927282	13.17788136
2	dpr13	41	10.08738269	12.85344044
2	beat-VI	41	10.65582342	12.71817812
2	Dscam4	41	13.24642903	12.50424033
2	CG34114	41	8.393459997	12.35663741
2	dpr1	41	8.704687351	11.57328135
2	kirre	41	8.005819044	11.33003148
2	beat-VII	41	7.566572699	8.662640233
2	Dscam2	41	7.494624558	6.5516479
2	dpr6	41	6.324568681	5.389798833
2	robo2	41	7.341097456	5.299057285
2	nolo	41	5.639924144	4.352247919
2	dpr9	41	6.012550738	4.11022583
2	dpr18	41	5.653206034	2.958871398
3	klg	25	13.65990787	16.99475317
3	Dscam2	25	12.98998306	16.27198839
3	CG42313	25	12.41604914	15.55840936
3	side	25	11.94684811	14.89641769
3	kek1	25	9.134510949	14.71041221
3	CG17716	25	11.95109143	14.17107097
3	Dscam4	25	13.549605	14.03054654
3	Ptp99A	25	11.60818351	13.74963374
3	CG34114	25	8.405937175	13.33463115
3	beat-IIa	25	10.6173709	13.31275393
3	ed	25	10.96328152	13.02980943
3	beat-VI	25	8.302471421	12.81134626
3	hig	25	9.929422624	12.56373701
3	nolo	25	7.651272662	12.26439543
3	dpr17	25	8.23902625	12.11636468
3	dpr20	25	7.970855803	12.11242492
3	kek3	25	8.294608302	12.02618783
3	dpr13	25	8.810777529	11.72807498
3	dpr3	25	7.71371204	11.59567378
3	dpr2	25	6.804359175	10.38768143
3	dpr9	25	7.282319734	9.529090579
4	Ptp99A	4	15.35698852	16.09632647
4	dpr8	4	10.07826537	15.19570238
4	CG17716	4	13.08149303	15.09376099
4	kek2	4	12.08993703	14.1984678
4	dpr13	4	10.61753734	13.93577841
4	kek1	4	11.13430873	13.8855283

4	robo3	4	10.51727555	13.7653428
4	CG12484	4	9.627496664	13.75642116
4	klg	4	9.360079399	13.57921124
4	dpr6	4	9.391144853	13.44684584
4	beat-IIa	4	8.648610092	12.57406967
4	CG34371	4	8.539732511	12.48059563
4	robo2	4	9.572599389	12.03759385
4	Dscam2	4	8.677812823	12.03524532
4	beat-VI	4	9.102927226	12.02629417
4	kek3	4	7.373463788	11.42748029
4	dpr9	4	10.80098861	11.08012122
4	beat-IV	4	8.161801869	11.07159899
4	kirre	4	7.416575666	9.871345598
4	CG34114	4	7.621337802	9.542015856
4	beat-VII	4	7.360071273	9.469230098
4	DIP-gamma	4	7.45953144	7.17610054
4	dpr11	4	7.486415378	6.944949586
4	side	4	7.479862446	6.842892926
4	ed	4	6.414835941	5.928472451
4	dpr10	4	6.109210875	4.91210046
4	otk	4	6.00144724	4.911013819
4	fred	4	6.878928439	4.595916727
4	hbs	4	5.351464816	4.120124668
5	DIP-beta	49	13.13104207	16.2373261
5	Fas2	49	14.44820784	15.49240572
5	beat-Ic	49	12.40893638	15.30374853
5	kek2	49	14.95359121	15.28237541
5	dpr13	49	15.00334212	15.24552128
5	dpr9	49	11.6154	14.97907998
5	CG17716	49	14.39482364	14.9482795
5	Dscam2	49	11.91670886	14.75034684
5	dpr6	49	14.55011167	14.67573676
5	ed	49	11.80584745	14.62543496
5	CG12484	49	11.10752507	14.50419599
5	nolo	49	11.57927665	14.18293721
5	Ptp99A	49	13.92056611	13.49373356
5	CG14372	49	13.43703171	13.34229908
5	dpr8	49	10.71777772	13.03223769
5	Dscam3	49	10.67782773	12.97976869
5	dpr1	49	8.671327062	12.97921399
5	Dscam4	49	13.82269979	12.66305411
5	kirre	49	12.66697655	12.60166323
5	CG42313	49	10.81358906	12.48267201
5	beat-Ib	49	7.958276051	12.17199306
5	beat-Va	49	9.622014309	11.89783554
5	fred	49	9.495470672	11.75705977
5	dpr2	49	7.830446966	11.66327653
5	DIP-theta	49	7.884180873	11.48292406
5	beat-Vc	49	8.127028943	10.539148
5	dpr11	49	6.220925937	8.221092941
5	beat-IIa	49	6.951261386	8.083198707
5	dpr20	49	6.24272115	7.225919663
6	beat-VI	40	14.86220724	15.30861372
6	Ptp99A	40	15.56614512	15.15249858
6	beat-IIa	40	12.3395086	14.84415478
6	CG17716	40	10.49931237	14.63173235

6	kek1	40	14.99448935	14.50021358
6	dpr1	40	10.80411934	14.13492412
6	beat-IIIb	40	8.858614283	14.11072034
6	Dscam4	40	12.27774257	14.06933951
6	side	40	9.610818686	13.93529703
6	Fas2	40	8.268175325	13.91528221
6	kek2	40	13.32732027	13.65780423
6	dpr9	40	11.42462938	13.4827336
6	CG31814	40	8.237814516	13.46060202
6	beat-VII	40	9.634154781	13.02334924
6	CG34353	40	8.923527932	11.023869
6	dpr17	40	7.254553988	10.31913902
6	kirre	40	6.802610179	10.06949986
6	dpr8	40	7.679988248	8.905982439
6	nolo	40	6.579522103	8.447490397
6	kek3	40	5.732917828	7.752932268
6	CG12484	40	5.446200125	6.45983532
6	CG34114	40	7.24437159	6.202885033
7	side	43	12.64993008	15.99132476
7	CG12484	43	15.46292081	15.823069
7	kek2	43	11.64174353	14.60872237
7	dpr13	43	11.75328767	14.58606358
7	kek1	43	10.62970775	14.19325398
7	Ptp99A	43	10.91981627	13.77912915
7	CG17716	43	12.83431709	13.70354519
7	dpr9	43	12.37257624	13.68565009
7	dpr8	43	8.992795834	13.53365916
7	Dscam4	43	12.89797756	12.86686929
7	CG34371	43	11.15503692	12.77885529
7	beat-Ib	43	7.982037749	12.28550092
7	klg	43	8.922421016	12.18706982
7	nolo	43	7.591460301	10.97757471
7	dpr3	43	7.453366954	9.890916556
7	kek3	43	6.753559972	9.155868823
7	beat-IIa	43	6.435833789	8.843570962
7	dpr6	43	7.307852122	8.478409403
8	side	36	13.14589662	14.6376637
8	CG17716	36	12.89713089	14.62544641
8	kek1	36	12.50451049	14.27800459
8	dpr9	36	12.8318661	13.96170282
8	beat-VI	36	11.12443227	13.95840664
8	CG12484	36	9.369878339	13.93766177
8	Ptp99A	36	12.6547209	13.78413147
8	Dscam4	36	10.80053479	13.4335545
8	kek2	36	10.63814932	13.41168395
8	dpr8	36	10.03777059	12.59096592
8	kek3	36	8.851130052	11.92039211
8	kek5	36	9.023198955	11.86261662
8	kirre	36	11.49846419	11.48061721
8	ed	36	7.550472987	9.273970526
8	dpr1	36	7.668116747	7.007150414
8	CG42313	36	7.396929131	6.837145544
8	hig	36	7.035241306	6.664625691
8	hbs	36	6.615441699	6.186879956
8	beat-Ic	36	7.579632318	6.161731614
8	fred	36	6.579916364	6.009337437

8	beat-VII	36	6.116789714	5.936996189
8	nolo	36	6.910487713	5.758159699
8	CG34114	36	6.891858555	5.712606787
8	beat-IIa	36	6.399872035	4.941581747
8	dpr17	36	5.257747523	4.295232732
8	dpr18	36	5.844218545	2.952949656
9	klg	57	16.3115735	16.23320446
9	dpr1	57	16.10594945	15.96031219
9	CG17716	57	15.85432639	15.92458772
9	CG12484	57	14.88700743	15.42180937
9	kek2	57	14.87564049	15.1668447
9	beat-Ic	57	14.39685368	15.12873215
9	Ptp99A	57	12.31502529	14.19465284
9	robo2	57	12.09406841	14.01575783
9	kek1	57	14.18001403	13.97269015
9	dpr13	57	9.55546941	13.778177
9	beat-VI	57	12.15664556	13.75596741
9	robo3	57	9.3612805	13.54937244
9	CG34114	57	13.21044534	13.45495304
9	CG42313	57	9.869348334	13.18476902
9	beat-Ib	57	12.69118182	13.12926219
9	beat-IV	57	10.95666731	13.12086746
9	DIP-gamma	57	10.88217353	13.08679592
9	kirre	57	12.42562399	12.97086182
9	dpr10	57	10.43441923	12.88594094
9	otk	57	8.891946143	12.45961232
9	dpr8	57	9.708466645	11.72090413
9	dpr18	57	9.343942943	10.48437808
9	nolo	57	9.239710496	9.596668966
9	DIP-zeta	57	6.749054004	8.502625539
9	beat-VII	57	6.485564282	8.159900334
9	CG34371	57	6.771570979	6.327605076
9	side	57	6.785235213	6.210166498
9	dpr20	57	5.914629606	4.120811167
9	Dscam3	57	5.40187352	3.985807702
9	beat-IIa	57	5.476788971	3.368915015
10	robo2	39	10.85916935	15.9784594
10	beat-Ic	39	13.19218648	15.77680635
10	Ptp99A	39	15.80170878	15.77640021
10	side	39	15.27196936	15.35952807
10	CG17716	39	14.82128475	15.02319933
10	klg	39	14.65348307	14.61572914
10	robo3	39	12.3629758	14.56567284
10	kek1	39	14.26293002	14.10192101
10	beat-IIIc	39	11.48807191	14.02438403
10	CG42313	39	12.85259346	13.55114893
10	beat-IIa	39	12.79064817	13.27445643
10	kek5	39	13.58791639	13.20687167
10	dpr3	39	13.03692494	12.99185296
10	kek3	39	12.77099825	12.82342148
10	Fas2	39	9.317149828	12.80843091
10	beat-IV	39	8.99115654	12.73919399
10	beat-Ia	39	9.217724187	12.50426456
10	dpr13	39	9.676360033	12.49040117
10	dpr10	39	10.43877093	12.18543156
10	kek2	39	8.606102472	12.15900613

10	otk	39	8.529983676	12.03149733
10	beat-IIIa	39	8.270427533	11.95353938
10	CG14372	39	8.525575582	11.26333279
10	dpr2	39	8.201633969	11.23952075
10	CG34353	39	7.880823136	9.321108983
10	Dscam2	39	7.125746316	8.274164092
10	dpr6	39	6.931769468	6.451129615
10	dpr9	39	6.806855844	5.848991918
10	nolo	39	5.745391539	5.460343198
10	Dscam4	39	5.73336607	5.349325983
10	dpr17	39	4.662770789	3.716568639
10	hbs	39	4.915050688	3.515252684
11	Dscam2	18	16.76986671	17.49437362
11	Ptp99A	18	15.67458999	16.37027846
11	kek1	18	15.02057872	15.66472289
11	dpr8	18	14.9825824	15.00965137
11	dpr1	18	15.14148494	14.76634717
11	CG34353	18	14.21630522	14.21421653
11	beat-IIIb	18	14.15626782	14.16426545
11	beat-IIIc	18	10.93171028	13.85755653
11	Dscam4	18	12.85020572	13.57608105
11	ed	18	12.67690904	13.53553343
11	klg	18	10.40759483	13.50130896
11	side	18	13.67143299	13.49979857
11	fred	18	10.08685876	12.67490766
11	dpr6	18	9.901287186	12.40873607
11	CG31814	18	9.731034619	12.30177252
11	hbs	18	11.13985396	12.19683205
11	beat-VI	18	9.783945401	12.11195667
11	kirre	18	8.917484135	11.19233338
11	beat-IIIa	18	8.749973983	10.53488462
11	kek3	18	8.794666053	10.18284739
11	kek2	18	8.334640262	10.0296169
11	beat-Ib	18	7.201527891	6.996278621
11	beat-Ic	18	7.076462089	6.752256611
11	kek5	18	7.22943457	6.640895624
11	CG34114	18	6.693878226	6.54037395
11	CG17716	18	6.829552125	6.437738594
11	CG34371	18	6.445497451	6.368430824
11	beat-VII	18	6.66883348	6.105386879
11	Dscam3	18	5.999413652	5.403941478
11	nolo	18	5.424949969	5.161389003
11	dpr20	18	4.580967155	3.505476216
12	CG34353	16	11.31621602	15.32966356
12	ed	16	9.376936828	14.54389017
12	kek1	16	12.11554653	14.31071813
12	CG17716	16	12.60720302	14.3065891
12	side	16	9.073004757	14.05870079
12	beat-Ic	16	9.676285384	14.05151899
12	dpr9	16	12.19114216	13.9863351
12	Ptp99A	16	11.72679718	13.80737326
12	robo3	16	10.17784626	13.76907152
12	CG42313	16	8.841553684	13.76446782
12	Dscam2	16	10.62050871	13.52740816
12	dpr8	16	9.631536457	13.12635722
12	robo2	16	9.112006017	12.88751497

12	kek2	16	8.139009597	11.42396636
12	CG34114	16	9.155781558	11.26643266
12	beat-VI	16	9.520493372	11.12690993
12	beat-Ib	16	8.04367334	10.89610821
12	kirre	16	7.040883515	9.690063733
12	beat-IIa	16	7.367118778	9.276547941
12	klg	16	8.073164647	7.074461875
12	kek3	16	6.521626528	5.083050509
12	beat-IV	16	6.501194797	5.082020376
12	Dscam4	16	5.349621182	4.328580941
12	dpr2	16	5.845970365	3.59413898
12	dpr13	16	6.026669135	3.05561138
13	dpr13	6	16.19646887	16.27308883
13	CG31814	6	15.47706873	15.94495133
13	dpr9	6	12.01120735	15.36518038
13	Dscam2	6	9.355932756	14.92613681
13	kek1	6	9.22808463	14.82211721
13	CG34353	6	12.0665529	14.65141352
13	CG17716	6	12.03585856	14.56504619
13	kek2	6	14.64180916	14.49005405
13	beat-Vc	6	8.990433989	14.2751623
13	DIP-gamma	6	8.632421847	14.14799496
13	CG42313	6	11.55175568	14.01702935
13	CG34114	6	8.783661293	13.92082004
13	Ptp99A	6	13.32002374	13.63480906
13	side	6	8.716785422	13.58851218
13	dpr20	6	11.20473163	13.58099263
13	dpr17	6	11.20508249	13.47304472
13	fred	6	11.27960902	12.82618606
13	klg	6	10.71812668	12.81386616
13	kek3	6	8.221459491	12.72466976
13	beat-VII	6	10.27694189	12.66832237
13	dpr6	6	7.938951532	12.66832237
13	beat-IIIb	6	8.671796315	12.66578496
13	beat-Va	6	10.42886302	12.64793332
13	kirre	6	10.13415545	12.58654255
13	dpr8	6	11.25457812	12.55087182
13	robo3	6	9.766348288	12.07703413
13	beat-Ic	6	8.406360694	11.23327041
13	beat-IIa	6	7.420746429	10.71498867
13	Dscam4	6	6.227793274	9.493740198
13	CG12484	6	6.374776899	9.233094784
13	CG14372	6	7.256607375	8.909777913
13	beat-IV	6	7.547473707	8.085919915
14	kek1	50	15.01221955	15.42538544
14	Dscam2	50	10.92215978	15.00776667
14	CG12484	50	8.790217945	14.86417318
14	side	50	12.47381584	14.72744989
14	dpr9	50	10.697422	14.60712434
14	Ptp99A	50	12.54761485	14.46589744
14	CG17716	50	9.021123992	14.16318426
14	beat-IIa	50	12.4945364	14.13059866
14	CG42313	50	11.58913039	13.70525192
14	dpr2	50	9.831781011	13.52230026
14	kek2	50	12.21992665	13.45266555
14	Dscam4	50	11.70695405	13.39822272

14	beat-IIIb	50	9.873079058	13.26780101
14	kek3	50	7.974413997	13.06137264
14	dpr10	50	7.769916975	13.04359508
14	beat-Ic	50	7.776652085	12.43969883
14	dpr8	50	7.556687192	12.22561558
14	beat-IIIa	50	7.655321334	12.1639995
14	kek5	50	7.247878428	11.4624381
14	CG14372	50	7.237553509	9.284926491
14	dpr5	50	6.279581117	8.481824608
15	CG12484	45	13.00404627	15.57216494
15	side	45	14.89490352	15.11555613
15	kek1	45	11.70907035	14.2654388
15	dpr9	45	9.85800425	14.01861746
15	klg	45	9.77493127	13.99603704
15	beat-IIa	45	11.05560139	13.9573652
15	CG34353	45	13.49631987	13.74578465
15	Dscam2	45	10.87902006	13.66207545
15	CG17716	45	14.05759694	13.57808876
15	CG34114	45	13.62890895	13.27947361
15	robo2	45	9.000234561	13.16191534
15	dpr20	45	8.883139436	13.01805428
15	ed	45	10.97032601	12.95280416
15	Ptp99A	45	10.42260584	12.94492621
15	dpr11	45	8.835695584	12.50017884
15	kek5	45	8.584954531	12.05110599
15	kirre	45	11.10037006	11.53945608
15	kek2	45	9.595704179	11.47998905
15	hbs	45	8.174818861	10.57261882
15	Dscam4	45	9.364245672	10.38140863
15	dpr3	45	7.473292166	7.088632354
15	dpr13	45	7.276007502	7.05636479
15	dpr6	45	7.36305926	7.032951903
15	CG42313	45	7.374388701	6.873529558
15	dpr8	45	7.431051259	6.793832433
15	beat-VI	45	7.19597654	6.62198335
15	beat-Ic	45	7.090454559	6.058550057
15	kek3	45	6.16988241	6.042160716
15	CG14372	45	5.463103634	3.440569497
15	dpr2	45	5.301441478	3.117048408
16	Ptp99A	60	15.74273472	15.93159165
16	CG42313	60	10.05805133	13.9547081
16	kek2	60	11.42804495	13.86818826
16	CG34114	60	10.92079133	13.85195707
16	dpr9	60	9.460551325	13.50400101
16	dpr8	60	8.018452459	13.27399648
16	beat-IIIb	60	9.075060452	13.12664991
16	Dscam2	60	9.4505291	13.09622846
16	kek1	60	7.924928043	13.07962466
16	Dscam3	60	10.90675058	13.03919715
16	side	60	8.866390047	12.94609917
16	beat-VII	60	9.536166605	12.89589622
16	dpr10	60	8.876632698	12.7308005
16	dpr1	60	8.045656127	12.59729806
16	kirre	60	11.29211043	12.30269448
16	fred	60	7.404334162	11.84569594
16	dpr13	60	8.061416645	11.73723436

16	CG17716	60	8.90317969	11.68669128
16	CG34353	60	7.478275096	11.62444527
16	dpr6	60	7.829599316	10.90106509
16	beat-VI	60	8.050455634	10.62262865
16	Dscam4	60	8.672744307	10.15472474
16	klg	60	7.364754173	10.0381236
17	dpr13	24	13.44577053	15.50026038
17	side	24	15.04497523	15.42946994
17	CG17716	24	14.89226777	15.39133569
17	robo2	24	13.31969708	14.98909843
17	dpr1	24	11.67371498	14.80692316
17	CG34114	24	9.947934235	14.70381671
17	kek1	24	12.64135767	14.46550952
17	kek2	24	13.08282476	13.51486471
17	Ptp99A	24	12.68261464	13.04096883
17	CG34371	24	7.704350738	12.71542837
17	beat-IIa	24	9.282638829	12.63717431
17	dpr9	24	11.4241542	12.12688019
17	robo3	24	8.628103539	11.76143315
17	beat-VI	24	7.917554598	11.46595585
17	dpr10	24	7.410687648	10.90073704
17	CG12484	24	8.049304943	10.70029177
17	Dscam2	24	8.931619954	10.56699279
17	beat-Ic	24	6.938185623	10.4724698
17	CG42313	24	7.375149688	9.88801492
17	beat-IV	24	6.276487076	7.010219829
18	Ptp99A	20	14.50695692	15.26132884
18	Dscam2	20	13.13649077	14.86931492
18	dpr6	20	9.099590408	14.25171005
18	kek1	20	11.17554291	14.07723675
18	beat-Ic	20	9.542144946	13.9549746
18	CG17716	20	9.017492116	13.31403905
18	klg	20	9.577568715	12.92645957
18	CG42313	20	8.659332548	12.89964999
18	CG12484	20	9.081472925	12.70458685
18	CG14372	20	9.686562715	12.65736544
18	beat-IIa	20	8.5744477	12.13531225
18	dpr10	20	8.151697962	11.64412529
18	Dscam4	20	7.251930811	11.57934517
18	side	20	8.465274822	11.53326386
18	dpr11	20	7.547567823	11.19812137
18	dpr1	20	8.28347866	10.78598066
18	robo3	20	9.220939495	10.76744479
18	dpr9	20	7.186668859	10.64803382
18	ed	20	7.630356487	10.45408263
18	CG31814	20	7.611902723	10.26024107
18	kek2	20	7.857819551	10.17102914
18	kek3	20	7.168244007	9.681475885
18	hbs	20	8.770077124	9.548207574
18	dpr8	20	7.76179676	9.512080956
18	CG34371	20	7.039900699	8.171769894
18	robo2	20	7.084821807	6.256289244
18	Dscam3	20	7.184362198	6.12389884
18	dpr13	20	7.223936562	5.739445222
18	beat-Vc	20	6.671657315	5.170813022
19	Ptp99A	19	13.63264468	15.77955603

19	side	19	12.2552586	14.90803031
19	kek1	19	11.0950506	14.2131969
19	dpr13	19	8.738456515	13.03202847
19	Dscam2	19	8.991725481	12.94073731
19	CG17716	19	9.122451474	12.32555892
19	kek2	19	8.491534655	12.20661795
19	Dscam4	19	8.581981797	11.33021566
19	dpr9	19	8.324225926	11.23581586
19	beat-IIa	19	7.995993222	11.21489129
19	kirre	19	8.850130454	10.9327582
19	Dscam3	19	8.0168676	10.34429906
19	klg	19	7.553169514	8.73199198
19	dpr8	19	7.846452531	8.556441023
19	hbs	19	6.301467026	7.997186278
19	CG12484	19	6.897822592	5.548088688
19	CG34371	19	6.513186648	4.153841693
19	kek3	19	6.249886892	4.062085094
19	fred	19	6.413759715	3.814472215
20	Fas2	56	16.09529964	16.20319898
20	Dscam2	56	12.76885247	14.99807203
20	dpr8	56	13.78715857	14.97230165
20	dpr6	56	10.16563133	14.54094925
20	Ptp99A	56	12.09772433	14.38057988
20	beat-VI	56	10.18992276	14.10840525
20	beat-IIIb	56	10.07068851	13.97156597
20	CG14372	56	12.78632708	13.78629329
20	kek2	56	11.9039979	13.69970998
20	fred	56	12.00318413	13.59646964
20	CG34114	56	11.07469172	13.30795105
20	robo3	56	9.192753323	13.18372732
20	side	56	9.104584463	13.15127411
20	dpr20	56	8.999398219	12.92147826
20	ed	56	10.60802316	12.89407442
20	CG12484	56	8.867010489	12.41219517
20	dpr10	56	8.738687939	12.40195283
20	CG17716	56	9.844423593	12.31348557
20	dpr1	56	10.10092162	11.89610522
20	CG34353	56	8.961270463	11.67585194
20	dpr11	56	9.743175798	11.32748563
20	hbs	56	8.392835263	11.2278273
20	kirre	56	6.783442689	7.833987341
20	otk	56	7.271909778	7.048726147
20	dpr9	56	7.334100347	6.969960257
20	kek3	56	6.314035111	5.849627268
20	kek1	56	6.599961929	4.826295698
20	DIP-beta	56	4.940480601	4.482236728
21	robo2	59	16.42284507	16.7813192
21	klg	59	9.879858456	15.95145693
21	CG17716	59	15.23295451	15.4700549
21	kek1	59	15.2005455	15.26316343
21	robo3	59	11.92076138	14.90819114
21	kek2	59	14.57575425	14.86213191
21	side	59	14.63002328	14.84597203
21	CG34371	59	11.48716567	14.56696693
21	dpr9	59	14.4745191	14.35065864
21	Ptp99A	59	14.52536483	13.76916614

21	dpr1	59	11.31889865	13.63482778
21	dpr6	59	8.42343135	13.62177265
21	beat-Vc	59	12.98839994	13.3865815
21	beat-Ia	59	8.026197256	13.24983754
21	beat-IIIb	59	10.63893573	13.19242836
21	dpr13	59	13.14714391	13.14825394
21	dpr11	59	8.183079916	13.05126814
21	dpr8	59	10.58888216	12.98223483
21	beat-IIIa	59	10.38536635	12.86285999
21	beat-IIIc	59	10.21444059	12.17033037
21	DIP-gamma	59	7.465966067	12.16792573
21	beat-VI	59	9.113782741	11.6146161
21	dpr10	59	9.521362838	11.11763553
21	kek3	59	9.285873151	11.06519216
21	CG31814	59	6.435209249	10.6150761
21	Dscam4	59	5.833609463	7.900643664
22	Ptp99A	32	13.36298928	14.96691857
22	dpr1	32	10.10644669	14.87261027
22	Fas2	32	11.78624537	14.48494836
22	beat-VI	32	11.55064508	14.30278593
22	dpr9	32	11.74830753	14.07427967
22	side	32	10.07995328	13.84391325
22	CG17716	32	10.57409906	13.24426972
22	klg	32	8.512624926	13.01171727
22	kek1	32	8.922360503	12.97151453
22	beat-VII	32	10.03404094	12.76686865
22	CG34353	32	7.609816271	12.72077281
22	dpr6	32	8.906642862	12.69860173
22	kek3	32	7.757928413	12.33832267
22	Dscam2	32	7.895827963	12.23743341
22	kek2	32	7.708700904	11.86128295
22	ed	32	9.771533676	11.35159298
22	robo3	32	8.077656698	11.32803999
22	otk	32	7.046468649	11.31374993
22	dpr8	32	9.282038064	11.00382024
22	beat-Ic	32	7.358761293	10.63205809
22	CG34114	32	6.457697628	9.913937469
22	CG14372	32	7.020688672	9.729920502
22	kirre	32	6.163678518	9.715328099
22	Dscam4	32	6.5606307	6.559067825
23	Dscam4	9	11.31346623	15.43594931
23	kek1	9	12.47794506	15.24562404
23	dpr13	9	14.53210849	15.13827622
23	side	9	14.92358556	14.87519277
23	CG17716	9	12.54989744	14.82174562
23	CG34114	9	12.4815162	14.75655239
23	dpr1	9	10.68491906	14.47328389
23	Fas2	9	10.56926752	14.28356837
23	dpr11	9	11.80285272	14.16538298
23	beat-IIa	9	8.199865596	13.66073581
23	Ptp99A	9	11.76869366	13.54971341
23	Dscam2	9	11.83124103	13.48222015
23	dpr8	9	10.10405117	13.04169259
23	beat-IIIc	9	7.975483242	12.9892315
23	kirre	9	10.87705284	12.69381595
23	nolo	9	10.92763149	12.62744527

23	kek3	9	9.293960458	12.32294239
23	beat-Ic	9	9.232396513	12.28395597
23	beat-Ib	9	6.688770948	9.869043068
23	dpr9	9	7.54231417	8.630861218
23	kek2	9	5.758838874	8.238949814
23	ed	9	5.954814983	6.114811459
23	dpr3	9	6.876586004	6.114811459
24	kek1	53	13.18270004	15.21385248
24	dpr8	53	12.47122656	14.60774519
24	beat-VI	53	13.80636286	14.44163111
24	CG17716	53	8.305130392	14.36217091
24	Ptp99A	53	14.12997453	14.09218122
24	kek2	53	14.10543379	13.95949096
24	Dscam2	53	12.26345876	13.93025279
24	CG31814	53	12.08803673	13.64495843
24	fred	53	9.655861117	13.48147322
24	dpr13	53	10.3415582	13.45027503
24	dpr9	53	11.79940942	13.44169594
24	Dscam4	53	11.94642413	13.17824354
24	ed	53	7.973587613	13.00861575
24	dpr10	53	9.63754779	12.64507105
24	robo2	53	8.910724582	12.55730071
24	robo3	53	9.625061477	11.84963305
24	beat-IIa	53	6.972359837	9.93163838
24	CG14372	53	6.361415389	8.247946823
24	kirre	53	6.041289895	6.768069973
25	side	29	13.07193186	15.91562716
25	CG17716	29	14.89130528	15.63525987
25	Ptp99A	29	14.95928569	15.28796654
25	beat-Ic	29	12.40738748	14.79414527
25	klg	29	13.55940641	14.73623567
25	dpr6	29	14.28197594	14.40053896
25	beat-VI	29	9.996454735	13.82821327
25	beat-IIIb	29	11.18585144	13.46885474
25	CG34353	29	9.526494733	13.4594759
25	dpr9	29	10.76705889	13.20954643
25	beat-Ib	29	10.42910327	13.19675989
25	kek2	29	11.59728935	12.90935289
25	dpr8	29	10.329889	12.81059065
25	ed	29	8.04355305	11.77753111
25	robo3	29	8.838109723	11.6012682
25	beat-VII	29	10.74313752	10.854062
25	Dscam4	29	8.455242184	9.041999313
25	hbs	29	6.948907098	8.052991931
25	otk	29	6.765637074	6.672764842
25	CG34371	29	7.059060243	6.621264688
25	beat-Vc	29	7.415694563	6.466618414
25	fred	29	7.106009277	6.338283428
25	dpr3	29	6.432638825	6.206615847
25	sdk	29	6.670774028	6.180296664
25	beat-IIa	29	6.23076143	6.104495781
25	dpr11	29	6.139317997	5.852994876
25	CG34114	29	6.379254151	5.10313123
25	dpr20	29	6.419444581	4.843438151
25	beat-IIIa	29	5.53050236	4.503391523
25	dpr10	29	5.836694217	3.846058966

26	CG12484	28	15.31898706	16.10507687
26	Ptp99A	28	15.97610442	15.59475792
26	kek1	28	11.80360842	14.44181268
26	kek2	28	10.17050767	14.1377411
26	Dscam2	28	8.750678248	13.89392031
26	Fas2	28	10.40870572	13.66401059
26	dpr9	28	14.04704577	13.50311412
26	Dscam3	28	10.71906705	13.33593418
26	kirre	28	11.0661762	13.00351752
26	Dscam4	28	11.90444826	12.3282228
26	dpr18	28	9.823942951	12.20718036
26	side	28	10.29365694	12.1356178
26	dpr8	28	12.40336396	12.11360249
26	robo3	28	7.795669315	11.8662248
26	dpr17	28	7.416689343	11.12951681
26	kek5	28	9.685720993	10.91365281
26	klg	28	8.810802411	10.7671307
26	CG34114	28	7.186448863	10.61095534
26	beat-Ib	28	7.502977986	10.3290642
26	ed	28	7.563712015	8.961843894
26	dpr13	28	6.797705059	7.097047095
27	klg	26	13.03930261	15.61813761
27	Dscam2	26	13.1611409	15.29374734
27	beat-IIa	26	13.32949837	14.89845901
27	CG42313	26	10.33721649	14.17371427
27	CG12484	26	9.702801698	13.6542353
27	Ptp99A	26	12.59960808	12.90134012
27	CG17716	26	9.687477751	12.77158208
27	kek5	26	8.48175124	12.04051186
27	Dscam4	26	8.291646533	11.85467642
27	kirre	26	8.428400636	11.7418659
27	fred	26	8.564182274	11.60555811
27	side	26	8.308348414	10.03327662
27	ed	26	7.48677706	8.481737299
27	DIP-iota	26	8.04495245	7.848398964
27	Fas2	26	7.205362863	7.02729627
27	dpr13	26	7.914522565	6.895279541
27	dpr9	26	7.281376737	6.825929533
27	beat-VI	26	7.188717925	6.375468157
27	dpr1	26	7.189070066	5.976286187
27	dpr18	26	6.381807469	5.83157332
27	kek1	26	7.093270754	5.83157332
27	dpr11	26	5.99536873	5.176050065
27	hbs	26	6.013977783	5.115471644
27	DIP-gamma	26	6.65370331	4.8524248
27	nolo	26	5.873181039	4.616071632
27	dpr20	26	6.125941374	4.502744963
27	hig	26	5.828511481	4.377036169
27	CG14372	26	4.668208194	4.248388045
27	beat-Vc	26	6.579186138	3.878704967
28	side	55	15.49244998	15.61566238
28	kek1	55	11.36879067	15.03441051
28	Ptp99A	55	9.550606457	14.93306547
28	nolo	55	9.287489069	14.33898688
28	dpr13	55	10.38349181	14.15789098
28	dpr9	55	14.26242502	13.94541428

28	dpr11	55	12.10574195	13.90710692
28	kek2	55	9.015096452	13.70140536
28	beat-IIa	55	10.86183974	13.69425865
28	beat-VI	55	9.807374657	13.60830674
28	CG17716	55	8.325520941	10.3983887
28	Fas2	55	7.919705193	7.248746672
28	kek3	55	7.219617448	6.150035659
28	beat-Ia	55	5.926743269	5.20551406
28	beat-Ic	55	5.890133075	2.966035938
29	kek2	44	15.46959838	15.56981823
29	klg	44	12.73803449	15.55771831
29	beat-VI	44	9.74506474	15.52799403
29	CG34114	44	11.02934486	14.8479407
29	Fas2	44	11.95582268	14.72889224
29	CG34353	44	14.44325614	14.65695507
29	beat-VII	44	11.56760658	14.45821682
29	dpr9	44	11.72889919	14.38488712
29	beat-Ia	44	10.8335196	14.29735612
29	dpr13	44	11.51021664	14.23583508
29	dpr1	44	8.721493508	13.83503998
29	dpr8	44	8.954517535	13.67937701
29	dpr17	44	8.259802359	13.32050911
29	beat-IIa	44	11.14545096	13.2258818
29	CG17716	44	8.187005544	12.7445723
29	DIP-gamma	44	7.905792932	12.56402809
29	Ptp99A	44	9.744164134	11.82722132
29	Dscam4	44	9.217498466	11.65124641
29	kek5	44	7.256066489	10.19091508
29	dpr6	44	6.992159845	8.931009474
29	dpr18	44	4.899989961	7.351946464
29	dpr11	44	5.64792459	7.351946464
29	CG12484	44	7.126646803	7.351946464
30	kek1	35	10.82955663	14.62531075
30	Ptp99A	35	12.69246951	14.52577915
30	side	35	10.83623677	14.46509035
30	beat-VI	35	10.68814836	14.32677672
30	dpr13	35	10.49360577	13.90864825
30	klg	35	10.31115516	13.78404191
30	CG12484	35	10.17577571	13.69439713
30	CG17716	35	9.963527215	13.04046241
30	Dscam4	35	10.71406846	12.33864277
30	dpr9	35	9.739396887	12.18667063
30	dpr10	35	7.451698213	12.06659345
30	kek2	35	8.06585065	12.05250518
30	kirre	35	8.862757985	11.86629657
30	beat-IIIa	35	6.928632553	11.22265773
30	dpr1	35	7.72319429	10.88904902
30	dpr11	35	7.046038933	10.52740492
30	beat-IIa	35	6.494711077	9.930233
30	Dscam2	35	8.150024691	9.522867689
30	CG14372	35	7.646056565	9.194252861
30	CG34114	35	6.825702741	9.019219278
30	CG31814	35	6.830992161	8.787208921
30	ed	35	6.973918324	7.614207565
31	CG17716	14	12.73048107	15.25697308
31	dpr8	14	11.77501308	14.95164802

31	klg	14	11.37094762	14.7405654
31	beat-VI	14	10.73701868	14.58878869
31	kek1	14	13.37091868	14.56939183
31	beat-IIa	14	11.5482179	14.55327283
31	dpr9	14	12.55691325	14.2404812
31	dpr6	14	9.268890466	13.92705364
31	Ptp99A	14	10.68394441	13.56417751
31	kek2	14	9.738266301	13.52575277
31	dpr1	14	9.839588145	13.51806362
31	Dscam2	14	11.30234318	13.39113448
31	side	14	9.476500789	13.12203534
31	dpr11	14	8.258050995	13.00454109
31	dpr13	14	9.065573078	12.48711335
31	dpr2	14	7.862872358	12.07781315
31	dpr10	14	9.252421905	12.04738667
31	beat-IIIb	14	8.236485257	12.04738667
31	kek3	14	8.106158935	12.0237183
31	CG34114	14	6.966095948	11.80644054
31	kirre	14	7.035739058	11.20061123
31	dpr3	14	6.975083885	10.99324831
31	Fas2	14	8.302731834	10.91222135
31	beat-Ic	14	7.033862176	10.46200066
31	Dscam4	14	7.161624097	9.519203139
31	CG34371	14	6.3644821	6.770851346
32	CG17716	15	11.30824948	14.33170273
32	CG42313	15	10.3206856	14.24042473
32	dpr1	15	10.23518267	14.21660634
32	CG34114	15	10.17002964	13.95016117
32	Dscam2	15	9.996290533	13.77550708
32	kek1	15	12.50265068	13.77161333
32	Ptp99A	15	11.10029335	13.33142378
32	dpr9	15	10.97866279	13.29830237
32	kek2	15	12.18372629	13.28183347
32	klg	15	9.988970884	13.01796749
32	dpr13	15	9.761886874	12.70564496
32	Dscam4	15	12.16833245	12.04691811
32	kirre	15	9.398574238	11.90418112
32	dpr6	15	9.348735112	11.41100385
32	CG34353	15	9.687671925	11.18608963
32	hig	15	8.22967122	10.28986327
32	hbs	15	8.297080783	10.06888861
32	beat-Ia	15	7.620952948	8.606566517
32	dpr2	15	7.005124691	6.440867791
32	dpr10	15	6.75338897	6.353623312
32	beat-Va	15	6.746991588	6.110876503
32	kek5	15	6.751081652	5.609403737
32	CG14372	15	6.460844551	5.520274938
32	robo3	15	6.57824742	5.331198671
32	dpr3	15	6.21354954	4.831643619
32	beat-IIIb	15	6.561738033	4.769523218
32	beat-IIIc	15	6.633032743	4.769523218
32	Fas2	15	6.084078708	4.769523218
32	ed	15	5.141254358	2.908872142
33	CG17716	5	16.00631136	16.00684778
33	dpr13	5	11.60296642	14.94577511
33	dpr8	5	9.343124782	13.97918277

33	kek1	5	12.14916902	13.85964387
33	kek2	5	10.79659154	13.3673646
33	kek5	5	11.08183059	13.34887931
33	dpr11	5	8.848324326	12.88936673
33	kek3	5	8.597468046	12.61538482
33	otk	5	9.335907548	12.57622038
33	Dscam2	5	8.470797668	12.35386916
33	Dscam4	5	8.591512869	12.04587133
33	Ptp99A	5	9.141483742	11.90697416
33	dpr9	5	7.781165267	11.62727407
33	beat-VI	5	8.541337786	11.16071925
33	beat-IIa	5	8.839857221	11.03310778
33	CG42313	5	8.431856947	10.72602997
33	robo3	5	8.407769883	10.56820145
33	beat-Vc	5	7.058981216	7.90800142
33	klg	5	7.096208567	6.776553103
33	DIP-theta	5	7.360637442	6.344646695
33	Fas2	5	6.290369771	4.903560382
34	dpr13	8	13.34792997	15.9440824
34	Fas2	8	11.0841707	15.69400505
34	CG17716	8	14.51550573	15.62284445
34	side	8	13.56907258	14.62109719
34	Ptp99A	8	12.12910043	14.15753054
34	beat-VI	8	10.53363544	14.08224665
34	CG34353	8	9.230074736	14.07222309
34	Dscam2	8	9.623047466	14.02034613
34	robo2	8	10.76083784	13.97533784
34	klg	8	10.09088463	13.4761262
34	beat-IIa	8	9.213416278	13.0431638
34	robo3	8	9.776571096	12.90894451
34	Dscam3	8	8.392817845	12.71564351
34	otk2	8	9.863580468	12.20488808
34	dpr1	8	8.407467148	12.17517261
34	dpr9	8	7.635216417	12.17346457
34	dpr10	8	8.33609542	12.08930758
34	CG34114	8	7.147363946	12.05880448
34	kek2	8	9.859836197	12.02435658
34	hig	8	7.617468303	11.91049262
34	Dscam4	8	12.21424016	11.77928378
34	kirre	8	7.641210877	10.47014386
34	CG12484	8	7.609596942	9.529888063
34	ed	8	7.947028882	9.404534589
34	fred	8	7.787097815	9.387794019
34	dpr8	8	6.822464358	8.869240822
34	kek1	8	6.706308261	6.663888892
35	kek1	52	15.8844452	15.9180003
35	CG12484	52	12.09732606	15.55672002
35	robo2	52	10.55229347	15.21546507
35	klg	52	12.28360148	15.16790866
35	robo3	52	14.68677338	15.0847268
35	CG17716	52	11.24157082	14.53299823
35	kek3	52	10.80718241	14.3159667
35	Ptp99A	52	14.3673509	14.18558431
35	beat-IIIb	52	13.95348817	14.17357263
35	kek2	52	14.12594789	14.05927494
35	dpr13	52	11.04341124	13.54988748

35	kek5	52	11.51096713	13.20020317
35	beat-IIIc	52	11.13444842	12.98716338
35	otk	52	7.825165094	12.56707474
35	Dscam4	52	10.13416769	12.45732859
35	dpr9	52	7.415472862	12.36078886
35	side	52	9.378781823	12.32462834
35	beat-IIIa	52	8.764808211	12.18944773
35	kirre	52	9.248000815	12.01464466
35	CG14372	52	8.524870088	11.98580532
35	DIP-beta	52	7.657468319	11.63142213
35	beat-IIa	52	6.488619003	10.12502319
35	beat-IV	52	6.862116901	9.115190907
35	dpr10	52	7.225647621	8.117790093
35	dpr18	52	6.782435397	7.11783003
36	CG17716	13	14.29161102	15.85605008
36	kek1	13	15.57646057	15.73902452
36	dpr13	13	13.05189746	15.65102754
36	dpr5	13	9.749376122	14.11096055
36	CG12484	13	9.035323261	14.07850157
36	otk	13	11.7763057	13.73522205
36	dpr9	13	11.20005304	13.61829635
36	beat-IIIb	13	10.66817531	13.5299195
36	otk2	13	7.729853334	13.05863539
36	side	13	8.125938526	12.93818254
36	klg	13	10.11384539	12.89016314
36	beat-Vc	13	9.099664737	12.46014941
36	CG42313	13	8.016025876	12.36413412
36	beat-IIIa	13	10.84658958	12.23451424
36	dpr8	13	9.071781792	12.16777612
36	Dscam2	13	7.628629636	12.04192802
36	Ptp99A	13	8.544530232	11.93516093
36	CG34114	13	9.976457831	11.57257431
36	kek2	13	8.958782316	10.68025572
36	beat-Va	13	7.327760073	10.64855244
36	CG34353	13	7.066607482	9.943876149
36	dpr11	13	7.017379396	6.771934468
36	Fas2	13	6.453534201	6.771934468
37	Dscam2	54	17.2416045	17.26757372
37	dpr8	54	14.45567824	15.38236397
37	Ptp99A	54	14.97418386	15.06575429
37	CG17716	54	12.35945256	14.908516
37	dpr9	54	14.70683145	14.84004689
37	dpr1	54	14.87187624	14.55066702
37	klg	54	11.62232066	14.06191714
37	side	54	12.63839548	14.0438102
37	kek2	54	10.01928991	13.99504867
37	dpr13	54	11.91293904	13.87729201
37	kek1	54	11.14727951	13.82076085
37	dpr11	54	12.41591228	13.70743498
37	dpr5	54	12.98770129	12.88735122
37	CG12484	54	10.43154287	12.63652647
37	dpr3	54	9.777579184	12.60099757
37	beat-VI	54	9.662532204	12.47449201
37	dpr20	54	9.88944521	12.25055254
37	DIP-delta	54	10.51070084	12.22362479
37	beat-IIa	54	10.39918277	12.19697751

37	dpr18	54	8.669701933	12.05789377
37	dpr2	54	9.337027171	11.27307347
37	kirre	54	7.65036317	10.87066312
37	CG34114	54	7.121969038	9.803054147
37	Dscam4	54	8.079603707	8.775191611
37	beat-Va	54	6.452786018	5.445882191
38	kek1	10	14.29846576	16.00812431
38	robo2	10	8.738989179	15.05882085
38	dpr1	10	11.74966711	14.92400545
38	dpr13	10	10.90982484	14.81365096
38	CG17716	10	10.88741501	14.44376147
38	side	10	10.51438357	14.41518688
38	Dscam2	10	12.443438	14.34505165
38	DIP-beta	10	11.55000189	14.34460407
38	Ptp99A	10	13.57461308	14.15905793
38	kek3	10	11.12004839	13.98914279
38	CG42313	10	11.22514958	13.93008391
38	kek2	10	9.012423575	13.79582431
38	beat-IV	10	9.189895849	13.37746023
38	robo3	10	13.10405191	11.85531204
38	Dscam4	10	8.007722333	10.24232555
38	dpr8	10	8.485939212	9.872637531
38	CG14372	10	7.364264757	9.640825154
38	dpr10	10	6.823317255	9.225930642
38	dpr9	10	7.546938819	8.974799075
38	ed	10	6.290464425	8.874175625
38	kek5	10	6.524815209	8.290749581
38	fred	10	7.162943867	8.290749581
38	nolo	10	8.152090822	8.075452199
38	hbs	10	7.752143877	7.977663646
38	CG12484	10	7.709724788	7.250638848
39	dpr1	46	12.34450106	16.55094896
39	CG17716	46	14.98089833	15.63266111
39	Ptp99A	46	15.2815157	15.40992851
39	beat-VI	46	11.63085576	15.03115983
39	CG34114	46	9.647982066	15.02409046
39	beat-Ic	46	12.84562291	15.02281914
39	kek1	46	13.22697576	15.0196388
39	dpr9	46	14.1467895	14.40759248
39	dpr8	46	13.99904784	14.311139
39	hbs	46	10.66150438	14.05630705
39	side	46	9.724090612	13.60395445
39	Dscam2	46	8.936673821	13.48604574
39	beat-Ib	46	11.87044225	13.46174885
39	otk	46	10.44725424	13.41259174
39	Dscam4	46	8.986912497	13.23980859
39	beat-IIa	46	11.99686401	13.17179362
39	dpr10	46	9.874700993	12.91282051
39	Fas2	46	9.701910031	12.79945797
39	DIP-beta	46	11.24532123	12.6341215
39	dpr13	46	8.738029488	12.57449171
39	beat-VII	46	9.619031735	12.45516775
39	DIP-eta	46	7.436713665	11.33981183
39	ed	46	7.788105258	11.2863795
39	nolo	46	7.820336435	11.19338411
39	hig	46	7.782939083	10.9025532

39	DIP-delta	46	10.1526027	10.90136635
39	beat-IV	46	7.163800829	8.594968325
39	beat-Ia	46	7.429416618	7.069782355
39	Dscam3	46	6.615564779	6.307193024
39	dpr6	46	7.436797071	6.130203203
39	dpr20	46	6.41431272	5.488006284
39	CG34353	46	6.227731406	4.440353301
39	CG31814	46	5.212523146	3.093933984
39	kek2	46	5.103543514	2.926182144
40	Ptp99A	7	16.15617758	16.10616253
40	dpr11	7	13.89954724	15.39485154
40	dpr13	7	15.69737267	15.36353491
40	CG42313	7	13.72243281	15.34284122
40	CG17716	7	13.37409831	14.80722545
40	Dscam3	7	12.36790996	14.72420784
40	robo3	7	14.67435838	14.65403123
40	Dscam2	7	12.02990672	14.20644907
40	kek5	7	12.04383446	13.65408719
40	dpr1	7	9.390613276	13.32267276
40	dpr6	7	10.57466952	12.86567269
40	kek2	7	7.672509589	11.59843128
40	beat-IIa	7	8.848786379	11.46991586
40	dpr8	7	8.421551559	11.4065793
40	dpr10	7	9.044140187	11.36150226
40	kek1	7	8.733825342	9.885970222
40	beat-Va	7	6.159046659	8.625749137
41	dpr13	48	12.94390386	15.33108905
41	CG42313	48	11.52866252	14.84589658
41	CG17716	48	12.61853645	14.78626185
41	Dscam2	48	9.203036347	14.65505693
41	side	48	8.824277565	13.97961985
41	kek1	48	10.54180403	13.66848511
41	Ptp99A	48	10.33470488	13.32148705
41	klg	48	9.131256483	13.02588183
41	nolo	48	8.738475263	12.16701073
41	dpr9	48	7.986301378	11.60560211
41	beat-IIa	48	9.147700906	10.16795158
41	beat-VI	48	7.984094298	8.801022259
41	kirre	48	6.365987465	8.311777438
41	kek3	48	6.814349855	7.902468142
41	Dscam4	48	6.515856697	7.355209507
42	Dscam2	11	15.46098805	16.72087752
42	CG42313	11	14.7289176	15.81580467
42	kek1	11	14.38984015	15.69968134
42	CG12484	11	14.68633195	15.37199042
42	CG17716	11	14.96512966	15.09269912
42	side	11	14.06614344	14.91371636
42	beat-Ic	11	13.74591088	14.74302189
42	kek2	11	13.45062077	14.45337244
42	kek3	11	13.41806417	14.36987145
42	Ptp99A	11	13.56789571	14.3533815
42	dpr6	11	13.44593298	13.88119358
42	beat-Ia	11	12.09256554	13.59247543
42	kirre	11	12.30942393	13.31039129
42	beat-Ib	11	9.169068913	13.21132899
42	dpr9	11	12.320479	13.19176351

42	beat-IIa	11	12.80913665	13.05539324
42	Fas2	11	13.01295593	12.88966024
42	dpr1	11	8.192050694	12.50695033
42	dpr2	11	8.946800698	12.450027
42	beat-VI	11	10.302428	12.42254405
42	dpr13	11	9.943609038	12.31404012
42	dpr10	11	9.382513637	12.01138656
42	dpr8	11	11.05162481	11.96633724
42	DIP-beta	11	8.059489212	10.93889348
42	otk2	11	7.323920288	10.8266775
42	beat-Va	11	7.872394441	10.50480969
42	beat-Vc	11	8.170546365	10.48468898
42	CG14372	11	5.869789479	9.369363822
42	hbs	11	8.127005537	9.339006856
42	otk	11	7.013835817	6.228030315
43	Dscam2	33	15.59000687	17.27532268
43	klg	33	14.71718791	16.06239954
43	beat-VI	33	13.72349353	15.70903367
43	side	33	13.06818754	14.87603974
43	dpr13	33	12.16508906	14.83715513
43	kek1	33	14.95572469	14.79651881
43	dpr1	33	11.36793614	14.73673218
43	Ptp99A	33	13.54429899	14.35330947
43	kek2	33	11.84103579	14.05795609
43	beat-IV	33	9.859572689	13.53597375
43	kirre	33	11.48469999	13.25200084
43	CG17716	33	9.493263854	13.12914999
43	DIP-delta	33	11.79950811	13.02763225
43	dpr8	33	9.272439435	13.02727949
43	CG12484	33	9.559234442	12.90433017
43	dpr9	33	11.9124096	12.45532106
43	CG34353	33	8.841600358	12.1078955
43	Dscam4	33	9.885466923	11.51246376
43	beat-VII	33	7.685177859	11.40622046
43	ed	33	8.546002607	10.80603606
43	beat-Ic	33	7.279559282	6.760665393
43	beat-IIa	33	6.387189354	5.123503136
43	kek5	33	6.553072239	5.002042856
43	CG42313	33	6.344235853	4.198106793
44	Ptp99A	1	10.33293595	15.32062144
44	Dscam2	1	12.68027985	15.12519114
44	dpr1	1	10.97813483	14.96515063
44	CG42313	1	12.78801182	14.67476309
44	side	1	8.680091817	14.3805597
44	beat-IIa	1	12.40920989	14.36126971
44	dpr9	1	14.61965196	14.04620291
44	CG17716	1	11.45840325	13.85168329
44	klg	1	13.85895275	13.68510715
44	beat-VII	1	7.875676233	12.46286076
44	robo3	1	12.17613939	11.78731855
44	CG14372	1	7.252946319	11.46311629
44	fred	1	7.399504312	11.25180079
44	kirre	1	8.314764013	10.666607
44	kek1	1	7.085170289	8.982410556
44	beat-VI	1	6.48727992	7.831028258
44	dpr6	1	7.623308795	7.831028258

44	dpr11	1	5.809456657	7.831028258
44	CG12484	1	7.262170693	7.831028258
44	beat-IV	1	7.311606402	6.258681994
45	CG17716	38	16.32374132	16.32374132
45	Dscam2	38	15.57333115	15.57333115
45	CG12484	38	15.50956801	15.50956801
45	dpr20	38	14.88685683	14.88685683
45	dpr2	38	14.63617944	14.63617944
45	kek1	38	14.11857051	14.11857051
45	side	38	14.10269416	14.10269416
45	dpr13	38	13.93222675	13.93222675
45	otk2	38	13.81779103	13.81779103
45	otk	38	13.72607962	13.72607962
45	robo3	38	13.52329889	13.52329889
45	dpr9	38	13.37641686	13.37641686
45	Fas2	38	12.67862073	12.67862073
45	kirre	38	12.56878651	12.56878651
45	kek2	38	12.1419549	12.1419549
45	CG34353	38	11.38508117	11.38508117
45	klg	38	10.93217075	10.93217075
45	Dscam4	38	10.49005193	10.49005193
45	dpr11	38	8.459064265	8.459064265
45	dpr1	38	8.096651647	8.096651647
45	beat-Vc	38	7.551025637	7.551025637
45	beat-Ia	38	7.321571671	7.321571671
45	beat-VII	38	7.18699538	7.18699538
45	beat-Ib	38	7.136605107	7.136605107
45	CG14372	38	7.130641159	7.130641159
45	dpr17	38	7.072107705	7.072107705
45	Ptp99A	38	7.016107681	7.016107681
45	kek3	38	6.972074714	6.972074714
45	robo2	38	6.807012181	6.807012181
45	CG42313	38	6.625458769	6.625458769
45	dpr6	38	6.182659056	6.182659056
45	CG34114	38	5.918063339	5.918063339
45	CG34371	38	5.286822296	5.286822296
45	dpr3	38	5.05697998	5.05697998
45	hig	38	4.626567889	4.626567889
45	beat-Va	38	4.100016957	4.100016957
45	hbs	38	3.023589086	3.023589086
45	DIP-gamma	38	2.534415865	2.534415865
46	side	27	10.41095045	15.42417292
46	dpr8	27	13.43672835	14.85850285
46	klg	27	13.36690734	14.7375526
46	Dscam2	27	12.26347816	14.69923637
46	CG42313	27	12.33180181	14.45917772
46	CG17716	27	9.309857128	14.1821038
46	beat-IIa	27	12.67574531	14.06218954
46	CG34353	27	10.31297121	13.97058245
46	beat-VII	27	10.28049574	13.54288445
46	dpr13	27	10.05385893	13.034359
46	DIP-delta	27	10.38141016	12.70651131
46	kek5	27	9.911567204	12.24454151
46	dpr9	27	9.383123636	12.05425872
46	hbs	27	8.503472706	11.94639671
46	Ptp99A	27	9.739286475	11.74059563

46	Dscam4	27	8.446585518	11.55407467
46	nolo	27	6.026990297	7.62549033
46	DIP-theta	27	7.728486817	6.408868461
46	beat-IIIa	27	6.805174809	4.611159484
47	Dscam2	17	13.90273936	16.1473275
47	beat-Ic	17	14.49938963	15.28922555
47	kek1	17	14.21788972	15.16681447
47	CG17716	17	12.24382978	15.00542901
47	dpr13	17	12.0176691	14.83095437
47	CG34114	17	9.996134837	13.74048056
47	DIP-gamma	17	8.682342002	13.42124866
47	beat-VI	17	10.00626813	13.16663721
47	kek3	17	8.154768762	12.7769983
47	dpr10	17	11.84146833	12.77606887
47	DIP-eta	17	9.319893904	12.63343453
47	beat-Ib	17	9.044371348	12.42915517
47	klg	17	8.493993699	11.90266897
47	otk	17	8.233420021	11.64030598
47	dpr9	17	8.156969479	11.42702485
47	Ptp99A	17	7.678626362	10.83201993
47	Dscam4	17	7.640728805	10.13917478
47	kek2	17	7.086599083	7.891031093
47	DIP-beta	17	7.251185805	6.439790079
47	dpr8	17	7.364864003	6.024404966
47	dpr6	17	7.393552517	5.253206234
47	hbs	17	6.020117327	4.572700055
47	CG42313	17	6.23402724	2.698363485
48	Dscam2	30	12.10077102	16.3456881
48	dpr9	30	11.13966055	15.11640381
48	side	30	9.621163611	14.2319154
48	CG17716	30	13.55977301	14.22428482
48	dpr13	30	9.373981675	14.22323936
48	dpr1	30	10.70812442	13.96889599
48	beat-IIa	30	12.45360745	13.93374951
48	Ptp99A	30	10.29503742	13.82778155
48	kek1	30	8.04178895	12.91248169
48	dpr17	30	7.435474889	12.19310886
48	Dscam4	30	8.207805918	12.18296078
48	beat-VII	30	7.71758421	11.76540149
48	dpr6	30	9.453781934	11.31100041
48	beat-VI	30	8.585216869	10.87164283
48	kirre	30	8.922344897	10.85721926
48	Fas2	30	8.903293432	10.71932034
48	dpr8	30	7.791105764	10.45686197
48	kek3	30	7.029975427	10.14439737
48	CG34114	30	6.364610844	9.312702913
48	klg	30	7.88654249	8.13863016
48	kek2	30	6.741719946	6.623223304
49	klg	47	13.02030645	15.56750059
49	beat-VI	47	12.94623006	15.42516487
49	Dscam2	47	12.05034164	14.43898614
49	beat-IIa	47	11.64487088	13.9998978
49	CG17716	47	8.805273082	13.97773574
49	dpr8	47	11.42863296	13.96687995
49	kirre	47	12.23893177	13.40744806
49	Ptp99A	47	9.766883642	13.30838312

49	beat-Ic	47	8.615535641	12.72777386
49	Dscam4	47	10.59788782	12.16154247
49	robo3	47	8.418961738	12.14085098
49	dpr6	47	8.406737439	12.1184939
49	ed	47	10.59773992	11.85895617
49	dpr18	47	7.300054648	11.31525598
49	CG14372	47	7.315200996	10.33072371
49	beat-VII	47	8.185260859	10.26280956
49	dpr9	47	7.220513101	9.746130688
49	dpr1	47	8.176384342	7.827541016
49	side	47	8.006584414	7.558005925
49	otk	47	6.910403245	5.890964429
49	dpr10	47	6.766270035	5.59817273
49	Fas2	47	6.847873819	5.468127647
49	beat-IIIc	47	6.707291913	4.857059359
49	kek5	47	5.88519102	4.015665274
50	klg	3	15.67811029	15.3521304
50	beat-VI	3	10.64594921	15.13689513
50	dpr13	3	14.4425521	15.0683542
50	Dscam2	3	14.9476061	14.96118592
50	side	3	14.99478948	14.87166453
50	Fas2	3	13.99309028	14.27594601
50	DIP-gamma	3	9.59450895	14.16959957
50	beat-Ic	3	9.792854689	14.08527149
50	CG34353	3	14.2081476	13.78592457
50	kek3	3	9.270291228	13.50988503
50	beat-VII	3	9.511591878	12.8092651
50	dpr8	3	9.455619261	12.80407354
50	dpr9	3	8.55759696	12.48174168
50	dpr10	3	8.395238214	12.37644954
50	kek2	3	8.678194027	12.37644954
50	kek1	3	9.240591213	12.33290571
50	CG17716	3	13.05918721	12.26878808
50	Dscam4	3	11.76695855	12.1037599
50	hig	3	12.87985848	12.03835247
50	dpr1	3	9.286987145	11.92516986
50	CG14372	3	8.169982039	11.76993379
50	robo3	3	8.101863348	11.20199404
50	kirre	3	10.75442786	11.13163533
50	beat-IIIb	3	7.704877119	10.71681216
50	Ptp99A	3	11.54530182	10.68926859
50	dpr18	3	7.843692529	10.42640893
50	beat-Va	3	7.514413398	9.618255831
50	dpr3	3	7.948635558	9.106052523
50	CG12484	3	8.408583217	9.106052523
50	beat-Vc	3	7.286788914	7.815482762
50	kek5	3	7.234436577	6.328157662
51	klg	12	16.41537574	16.23512919
51	side	12	15.28035789	15.5061238
51	kek1	12	10.53812486	15.03179053
51	kirre	12	11.71114375	14.14349244
51	beat-IIIb	12	9.549275831	13.97570614
51	dpr10	12	13.54723836	13.91367511
51	CG34114	12	9.441214509	13.47813627
51	dpr8	12	11.60384104	13.35952838
51	dpr9	12	13.68179873	13.33289015

51	hbs	12	10.76433362	13.21763373
51	otk	12	9.070352133	13.17082857
51	Dscam2	12	9.297665905	13.01139892
51	beat-IIa	12	13.16929	12.88856888
51	CG31814	12	9.320014565	12.53325086
51	dpr17	12	9.100603688	12.05530687
51	Ptp99A	12	8.477578766	11.94115914
51	CG17716	12	8.740569282	11.68336853
51	fred	12	7.966359476	10.73244853
51	Dscam4	12	10.78742636	10.36464016
51	beat-VI	12	8.088111838	8.498592046
51	Dscam3	12	7.57354268	7.271284738
51	CG34353	12	7.577461034	7.19048745
51	Fas2	12	7.284262342	7.107202125
51	kek2	12	7.242679043	7.088651216
51	dpr6	12	7.407413446	7.062518704
51	dpr1	12	6.9304679	6.224322475
51	beat-Ic	12	6.734639927	5.797139106
51	ed	12	6.36521462	4.890824183
51	robo3	12	5.692494708	4.280719877
51	CG34371	12	5.299576135	2.304874924
52	dpr9	34	16.18921705	16.46038277
52	Dscam2	34	15.39590589	16.10429952
52	side	34	16.06047257	16.05688417
52	beat-IIa	34	13.61164193	14.88498725
52	CG17716	34	14.39584475	14.60038257
52	CG34114	34	12.13501412	14.37395891
52	CG42313	34	10.06511482	14.28881026
52	CG12484	34	11.28571226	13.76459842
52	dpr2	34	13.81562371	13.47403593
52	beat-Ic	34	9.0359106	13.02508735
52	dpr18	34	8.893848081	12.8994835
52	dpr6	34	8.875883575	12.31320055
52	Dscam4	34	10.37535083	12.20307918
52	fred	34	12.47436023	12.16651789
52	dpr10	34	9.014517219	11.91341249
52	beat-VII	34	7.984151683	10.88422409
52	kek2	34	8.281314264	10.48653114
52	ed	34	7.450587637	10.20848593
52	Ptp99A	34	7.355781509	9.041744682
52	kek1	34	7.133774068	6.79042936
52	robo3	34	6.912047212	6.620957895
52	CG31814	34	6.219648416	5.306033067
52	beat-Ia	34	5.527192908	4.808374328
52	otk	34	6.000585692	4.610622908
52	dpr8	34	5.402662413	3.119033695
52	kek5	34	4.2906083	2.498638
53	CG17716	42	15.45310565	15.89795041
53	dpr13	42	12.26701908	15.05274745
53	kek2	42	11.45835601	14.1307491
53	dpr11	42	9.224815766	14.09806438
53	beat-IIIb	42	10.82264295	13.98776362
53	kek3	42	10.74733726	13.96564365
53	kek1	42	10.58357099	13.62806631
53	Dscam4	42	10.41569007	13.55482052
53	otk2	42	7.987801257	13.05200935

53	Dscam2	42	9.302613108	12.88736487
53	beat-IIa	42	10.79173251	12.73899366
53	dpr10	42	8.350553145	12.28428335
53	dpr8	42	9.287948894	12.13099035
53	Ptp99A	42	9.280329285	12.08223559
53	kek5	42	7.030698015	10.41955273
53	Dscam3	42	7.342747537	9.913474582
53	beat-VI	42	7.995341465	9.567149457
53	otk	42	6.524979533	9.323950094
54	CG17716	31	16.43292715	16.39657563
54	Ptp99A	31	15.0920085	16.1406091
54	dpr13	31	15.88633466	16.09233065
54	Fas2	31	14.9154386	15.27502687
54	robo2	31	12.91706032	14.84619616
54	dpr6	31	12.7780493	14.3356018
54	otk	31	9.987542915	14.0817081
54	DIP-delta	31	9.731835668	14.0367604
54	hig	31	11.54263195	14.00377629
54	side	31	13.88295221	13.94837634
54	kek2	31	13.61542831	13.81443268
54	dpr1	31	8.054782551	13.44989807
54	CG34371	31	13.50619679	13.31371124
54	CG34114	31	7.669262706	12.61992335
54	Dscam3	31	9.439219882	12.55692106
54	dpr10	31	11.17221494	12.450027
54	beat-VI	31	7.688577071	12.36792076
54	otk2	31	7.492220905	12.36257114
54	DIP-theta	31	7.308815264	12.11036246
54	Dscam4	31	11.84474375	11.55607747
54	robo3	31	12.11302228	11.55607747
54	beat-IV	31	8.733676958	10.91891371
54	CG34353	31	8.410557346	10.85385437
54	dpr8	31	6.537144729	8.665415072
54	dpr18	31	6.18454819	7.754937764
55	CG17716	2	13.95456262	15.81876025
55	beat-IIIc	2	10.57451913	15.42602399
55	Ptp99A	2	14.33028651	14.80998935
55	kek3	2	12.95084173	14.78866159
55	kek2	2	13.61742579	14.55609562
55	kek1	2	12.54334456	14.14899351
55	beat-IIIb	2	11.21665336	14.10479936
55	otk2	2	11.38397126	13.98800822
55	klg	2	10.04713853	13.78517451
55	CG12484	2	9.563743548	13.68764073
55	otk	2	8.171502197	13.61940529
55	dpr9	2	10.78162626	13.37905512
55	CG42313	2	10.48586309	13.24784502
55	beat-Ic	2	8.911772686	12.9004335
55	Dscam2	2	10.12126622	12.43778132
55	robo3	2	9.544655449	12.22527796
55	Dscam4	2	9.218331613	12.09322009
55	CG34114	2	7.63298812	11.67440292
55	kirre	2	7.261796879	11.66748867
55	dpr10	2	7.230503709	10.71325964
55	beat-VI	2	8.110911576	10.44945671
55	beat-IIa	2	8.320666843	10.03509879

55	side	2	7.026424571	9.865009996
55	dpr8	2	7.205830946	9.058527498
55	ed	2	7.257983078	8.866556239
55	dpr6	2	6.93588342	7.478966354
55	CG34371	2	5.465361732	5.136677607
56	dpr1	51	14.18831726	15.64120574
56	kek1	51	15.28961676	15.41398803
56	CG17716	51	13.45942198	15.2468826
56	otk	51	10.49257711	14.66583704
56	beat-Ic	51	8.94903345	14.29375647
56	dpr8	51	9.071822281	14.0304057
56	beat-VI	51	10.56240739	13.74827324
56	kek5	51	9.885542275	13.52001383
56	otk2	51	11.53845723	12.79126004
56	Ptp99A	51	10.20176016	12.33665083
56	hig	51	9.866684926	12.30984424
56	ed	51	10.28999357	11.80313863
56	Dscam4	51	9.236248981	11.7938203
56	beat-IIa	51	8.950155275	11.51060204
56	hbs	51	8.510549147	10.05283702
56	beat-VII	51	7.943879695	9.647265913
56	CG34114	51	7.938712712	7.141081808
56	DIP-gamma	51	7.469059374	6.536100653
56	CG34353	51	7.063948422	6.419718414
56	side	51	7.100567662	6.349882158
56	dpr13	51	7.2283392	6.23641113
56	Dscam2	51	6.705152732	5.939048724
56	CG42313	51	7.121382413	5.768581366
56	dpr20	51	6.538587591	5.263237455
56	fred	51	6.407012981	5.216715356
56	kek2	51	6.2630727	4.608587561
57	Dscam2	22	13.08889929	16.13055675
57	CG17716	22	9.014851505	15.08230705
57	CG42313	22	10.78126973	14.64923663
57	side	22	10.53933535	14.55876329
57	klg	22	12.6510931	14.54490419
57	beat-IIa	22	12.46656652	14.15982782
57	dpr8	22	8.926030919	13.32312629
57	CG34353	22	9.276197769	13.15990664
57	kek2	22	8.555003468	12.42319987
57	kek1	22	7.318705751	12.35441527
57	Ptp99A	22	8.090994812	11.95350791
57	beat-VII	22	8.797799293	11.7815058
57	kirre	22	8.293578333	11.5626502
57	DIP-delta	22	7.895043993	11.03290037
57	Dscam4	22	7.729420216	9.90088508
57	dpr9	22	6.679856821	9.80776569
57	dpr6	22	6.445486451	6.271653503
57	beat-VI	22	6.192011401	6.007537145
57	otk	22	6.099305664	5.133710455
58	dpr13	37	13.9971798	16.26622287
58	Dscam2	37	14.81587163	15.98284439
58	CG42313	37	10.78606544	15.03994019
58	beat-Ib	37	12.22633846	14.35515306
58	DIP-eta	37	10.42649941	14.35515306
58	ed	37	12.02329767	14.19197754

58	CG34114	37	10.18295081	14.03904737
58	kek3	37	13.89507797	14.00780155
58	beat-Ic	37	10.68508139	13.71614278
58	Ptp99A	37	13.55965025	13.58285007
58	DIP-beta	37	8.204147942	13.52351301
58	dpr1	37	8.521448654	13.50607526
58	kek2	37	9.23789728	13.28515439
58	kek1	37	9.37281913	12.85777893
58	CG34371	37	11.91994084	12.50585148
58	kirre	37	10.48169551	11.90807729
58	otk2	37	8.689940222	11.49535683
58	beat-VI	37	8.403352712	10.05487459
58	dpr8	37	6.879083587	7.554801825
58	beat-IIa	37	5.974850142	6.712075171
59	side	21	15.59825122	16.17238577
59	beat-VI	21	13.65819037	16.079908
59	dpr13	21	13.95358287	15.83720668
59	Dscam2	21	13.28542851	15.5735049
59	Fas2	21	13.29862504	15.37749224
59	CG17716	21	12.46226233	14.48767296
59	CG34353	21	11.15751644	14.01818478
59	kek1	21	10.13008377	13.91968903
59	dpr6	21	11.98801059	13.81491434
59	kek2	21	13.80783205	13.57802077
59	otk2	21	7.996190582	13.23992338
59	robo3	21	11.47892146	13.22254498
59	robo2	21	8.289540681	13.06209807
59	kirre	21	12.93235171	12.92052741
59	otk	21	9.176518582	12.87309264
59	dpr20	21	8.233202806	12.77999605
59	DIP-gamma	21	7.951532151	12.50771777
59	CG34371	21	9.093430087	12.05907794
59	Dscam3	21	7.475700415	11.96650554
59	dpr9	21	7.818432781	11.80160543
59	dpr10	21	8.354329508	11.4719893
59	dpr18	21	7.705322475	10.96686599
59	beat-Vc	21	7.363784785	10.77432388
59	Ptp99A	21	8.292079647	9.975899854
59	beat-Va	21	5.892755873	9.414703003
59	beat-IIa	21	6.580692056	7.971902836
59	beat-VII	21	6.78916515	6.506397028
59	dpr17	21	5.399122497	5.597005223
60	Ptp99A	58	14.90078358	15.57589936
60	CG42313	58	12.63978416	14.84110018
60	dpr13	58	12.5881634	14.82977799
60	CG34353	58	10.15282202	14.50388086
60	CG17716	58	14.83554715	14.49417533
60	dpr11	58	12.21604846	14.21709823
60	kek1	58	9.616446643	13.78675045
60	Fas2	58	9.982691236	13.6873884
60	Dscam2	58	11.21947078	13.43933475
60	dpr8	58	9.347183342	13.09481536
60	robo3	58	9.120498216	12.99583482
60	kek2	58	10.50452166	12.8375671
60	dpr9	58	10.7187288	12.58560551
60	kirre	58	12.13867772	12.52645679

60	dpr6	58	10.74627744	12.36104097
60	beat-Vc	58	10.5215775	12.23418374
60	kek5	58	8.258655509	11.16479631
60	Dscam4	58	7.881828721	10.82580177
60	beat-IIa	58	8.314498141	10.22397508
60	robo2	58	7.903459509	7.720021042
60	dpr1	58	8.131700008	7.565225068
60	side	58	7.630776611	7.498505206
60	CG12484	58	7.477455865	7.014731794
60	beat-IIIb	58	6.878081866	6.824608763
60	otk2	58	6.516669913	6.215838655
60	fred	58	6.56324552	6.127416804
60	CG34114	58	6.164515789	5.526791722
60	ed	58	6.199183225	5.526791722
60	hbs	58	6.262267844	5.421268101
60	Dscam3	58	6.032289245	4.609015577
60	CG34371	58	5.689046889	4.337679232

Supplementary table 4: Interactome Ig domain protein in MNs and muscle cells
Table represents predicted *Ig* protein interactions of candidates expressed in the motoneuronal and muscle scRNA data set.

motoneuron	muscle	**Physiological output of interaction**	**Interaction assay**	**Reference**
dpr2	DIP-κ	?	SPR	(*Cosmanescu et al., 2018*)
dpr2	DIP-ι	?	SPR	(*Cosmanescu et al., 2018*)
dpr2	DIP-η	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr2	DIP-θ	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
beat-IV	side-VII	Repulsion?	ELISA, SPR	(*Ozkan et al., 2013*, *Ozkan et al., 2013*)
DIP-κ	dpr7	?	SPR	(*Cosmanescu et al., 2018*)
DIP-κ	dpr1	?	SPR	(*Cosmanescu et al., 2018*)
robo2	robo2	Adhesion?	MS, PD, FA	(*Evans and Bashaw, 2010*, *Simpson et al., 2000*)
robo2	robo1		WB, CoIP, PD, FA	(*Evans et al., 2015*, *Simpson et al., 2000*)
robo2	slit		SPR, SPA, ELISA	(*Evans and Bashaw, 2010*, *Howitt et al., 2004*, *Simpson et al., 2000*)
dpr13	DIP-ε		ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)

kek5	kek5	?	CoIP, WB	(*MacLaren et al., 2004*)
kek5	kek6	?	CoIP, WB	(*MacLaren et al., 2004*)
kek5	kek2	?	CoIP, WB	(*MacLaren et al., 2004*)
kek5	kek1	?	CoIP, WB	(*MacLaren et al., 2004*)
DIP-θ	dpr4	?	ELISA, SPR, CoIP	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*)
DIP-δ	dpr9	?	SPR	(*Cosmanescu et al., 2018*)
Lar	Lar		CoIP, WB	(*Hofmeyer and Treisman, 2009*)
Lar	InR		SPR, CoIP, WB	(*Kaur et al., 2019*, *Madan et al., 2011*)
DIP-ε	dpr18	?	SPR	(*Cosmanescu et al., 2018*)
Hasp	hig		CoIP, WB	(*Nakayama et al., 2016*)
DIP-η	dpr1	?	ELISA, SPR, X-Ray	(*Cheng et al., 2019*, *Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
DIP-η	dpr4	?	ELISA, SPR, X-Ray	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*)
otk	otk		FM, CoIP, WB	(*Linnemannstöns et al., 2014*, *Pulido et al., 1992*)
otk	otk2		CoIP, WB	(*Linnemannstöns et al., 2014*)
otk	fz2		CoIP, WB	(*Linnemannstöns et al., 2014*)
otk	PlexA		CoIP, WB	(*Winberg et al., 2001*)
otk2	otk2		CoIP, WB	(*Linnemannstöns et al., 2014*)
otk2	otk		CoIP, WB	(*Linnemannstöns et al., 2014*)
otk2	fz2		CoIP, WB	(*Linnemannstöns et al., 2014*)
Sema2a	PlexB		MP, GI, AT, LM	(*Rozbesky et al., 2019*, *Wu et al., 2011*, *Ayoob et al., 2006*)
klg	klg		FM, ELISA	(*Ozkan et al., 2013*, *Butler et al., 1997*)
klg	cDIP		ELISA	(*Ozkan et al., 2013*)
dpr6	DIP-α		SPR, X-Ray, ELISA	(*Sergeeva et al., 2020*, *Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*, *Ozkan et al., 2013*)

unc-5	sns		ELISA	(*Ozkan et al., 2013*)
DIP-ζ	dpr14	?	SPR	(*Cosmanescu et al., 2018*)
DIP-ζ	dpr19	?	SPR, ELISA	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
DIP-ζ	dpr13	?	SPR, ELISA	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
rst	rst		FM, ELISA, BA	(*Ozkan et al., 2013*, *Schneider et al., 1995*)
rst	sns		ELISA, BA, FM	(*Ozkan et al., 2013*, *Galletta et al., 2004*)
rst	hbs		ELISA	(*Ozkan et al., 2013*)
Cont	Nrg		Co-IP, WB	(*Banerjee et al., 2006*, *Faivre-Sarrailh et al., 2004*)
Ptp69D	Ptp69D		Co-IP, WB	(*Garrity et al., 1999*)
Ptp69D	robo3		Co-IP, WB	(*Oliva et al., 2016*)
Ptp69D	Dscam1		Co-IP, WB	(*Dascenco et al., 2015*)
DIP-α	dpr6	?	SPR, X-Ray, ELISA	(*Sergeeva et al., 2020*, *Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*, *Ozkan et al., 2013*)
DIP-α	dpr10	Stabiles cell-adhesion	ELISA, X-Ray, SPR	(*Sergeeva et al., 2020*, *Cheng et al., 2019*, *Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
beat-IIa	side-IV		ELISA	(*Ozkan et al., 2013*)
beat-IIa	side		ELISA	(*Ozkan et al., 2013*)
dpr18	DIP-ζ		SPR	(*Cosmanescu et al., 2018*)
dpr18	DIP-ε		SPR	(*Cosmanescu et al., 2018*)
tutl	tutl		FM, BA	*Ferguson et al., 2009*)
dpr4	DIP-η	?	ELISA, SPR, X-Ray	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*)
dpr4	DIP-θ	?	ELISA, SPR, CoIP	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*)
dpr17	DIP-γ	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
fipi	Fas2		ELISA	(*Ozkan et al., 2013*)

hbs	CG15098		EX	(*Guruharsha et al., 2011*)
hbs	rst		ELISA	(*Ozkan et al., 2013*)
hbs	kirre		ELISA, BA, FM, Co-IP, WB	(*Ozkan et al., 2013*, *Bao et al., 2010*, *Shelton et al., 2009*, *Dworak et al., 2001*)
hbs	sns		Co-IP, WB	(*Shelton et al., 2009*)
dpr15	DIP-γ	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*)
dpr15	DIP-β	?	SPR	(*Cosmanescu et al., 2018*)
dpr9	DIP-β	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr9	DIP-λ	?	SPR	(*Cosmanescu et al., 2018*)
dpr9	DIP-ε	?	SPR	(*Cosmanescu et al., 2018*)
dpr9	DIP-δ	?	SPR	(*Cosmanescu et al., 2018*)
dpr9	DIP-ζ	?	SPR	(*Cosmanescu et al., 2018*)
dpr5	DIP-ι	?	SPR	(*Cosmanescu et al., 2018*)
dpr5	DIP-η	?	SPR	(*Cosmanescu et al., 2018*)
dpr5	DIP-θ	?	SPR, ELISA	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
DIP-γ	dpr17	?	SPR, ELISA	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
beat-Vb	side-VI	Repulsion?	AT, ELISA, SPR	(*Li et al., 2017*, *Ozkan et al., 2013*)
dpr8	dpr8	?	SPR	(*Cosmanescu et al., 2018*)
dpr8	DIP-β	?	SPR, ELISA	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr8	DIP-λ	?	SPR	(*Cosmanescu et al., 2018*)
side-VII	CG17839	Repulsion?	ELISA	(*Ozkan et al., 2013*)
side-VII	beat-IV	Repulsion?	SPR, ELISA	(*Ozkan et al., 2013*, *Ozkan et al., 2013*)
dpr20	DIP-ε	?	SPR, ELISA	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)

dpr20	DIP-ζ	?	SPR, ELISA	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr14	DIP-ζ	?	SPR	(*Cosmanescu et al., 2018*)
dpr14	DIP-ε	?	SPR, ELISA	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr11	DIP-β	?	SPR, ELISA	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr11	DIP-γ	?	SPR, ELISA, X-ray	(*Sergeeva et al., 2020*, *Cheng et al., 2019*, *Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr11	cDIP	?	ELISA	(*Carrillo et al., 2015*, *Ozkan et al., 2013*)
beat-Ic	side-III	Repulsion?	SPR	(*Li et al., 2017*)
beat-Va	side-VI	Repulsion?	ELISA, SPR, AT	(*Li et al., 2017*, *Ozkan et al., 2013*)
kek2	kek2	?	CoIP, WB	(*MacLaren et al., 2004*)
kek2	kek5	?	CoIP, WB	(*MacLaren et al., 2004*)
kek2	kek6	?	CoIP, WB	(*MacLaren et al., 2004*)
kek2	kek1	?	CoIP, WB	(*MacLaren et al., 2004*)
dpr3	DIP-η	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr3	DIP-θ	?	ELISA	(*Carrillo et al., 2015*, *Ozkan et al., 2013*)
DIP-ι	dpr1	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*)
vn	boi		ELISA, SPR	(*Ozkan et al., 2013*, *Ozkan et al., 2013*)
vn	ihog		ELISA	(*Ozkan et al., 2013*)
dpr12	dpr12	?	CS, MW	(*Cosmanescu et al., 2018*)
dpr12	DIP-δ	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr12	cDIP	?	ELISA	(*Carrillo et al., 2015*, *Ozkan et al., 2013*)
FBgn0083949	beat-Ic	Adhesion	ELISA, SPR	(*Ozkan et al., 2013*, *Ozkan et al., 2013*)

side-III	side-II	Repulsion?	SPR	(*Li et al., 2017*)
side-VI	hig	Repulsion?	ELISA	(*Ozkan et al., 2013*)
side-VI	beat-Va	Repulsion?	ELISA, SPR, AT	(*Li et al., 2017*, *Ozkan et al., 2013*)
side-VI	beat-Vc	Repulsion?	ELISA, SPR	(*Li et al., 2017*, *Ozkan et al., 2013*)
side-VI	beat-Vb	Repulsion?	ELISA, SPR, AT	(*Li et al., 2017*, *Ozkan et al., 2013*)
Fas3	Fas3		ELISA	(*Ozkan et al., 2013*)
DIP-ε	dpr19	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
DIP-ε	dpr13	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
DIP-ε	dpr6	?	ELISA	(*Carrillo et al., 2015*, *Ozkan et al., 2013*)
DIP-ε	dpr17	?	ELISA	(*Carrillo et al., 2015*, *Ozkan et al., 2013*)
Lac	Lac		FM, BA, ELISA	(*Ozkan et al., 2013*, *Strigini et al., 2006*, *Llimargas et al., 2004*)
Lac	CG6959		ELISA	(*Ozkan et al., 2013*)
dpr19	DIP-ζ	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr19	DIP-ε	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
robo3	robo3		ELISA	(*Ozkan et al., 2013*)
robo3	robo1		Co-IP, WB	(*Evans et al., 2015*)
robo3	sli		Co-IP, WB, ELISA	(*Oliva et al., 2016*, *Howitt et al., 2004*, *Simpson et al., 2000*)
robo3	Ptp69D		Co-IP, WB	(*Oliva et al., 2016*)
kek1	kek1	?	CoIP, WB	(*Alvarado et al., 2004*, *MacLaren et al., 2004*)
kek1	kek6	?	CoIP, WB	(*MacLaren et al., 2004*)
kek1	kek2	?	CoIP, WB	(*MacLaren et al., 2004*)
kek1	kek5	?	CoIP, WB	(*MacLaren et al., 2004*)
ImpL2	ImpL2		X-Ray, ELISA	(*Roed et al., 2018*, *Ozkan et al., 2013*)
ImpL2	CG11656		EX	(*Guruharsha et al., 2011*)

dome	dome		WB, BGA	(*Fisher et al., 2018*, *Fisher et al., 2016*, *Brown et al., 2003*)
beat-Ia	side	Repulsion?	Co-IP, WB, FM, SPR, ELISA	(*Ozkan et al., 2013*, *Ozkan et al., 2013*, *Siebert et al., 2009*)
beat-IIb	side	Repulsion?	ELISA	(*Ozkan et al., 2013*)
beat-IIb	side-IV	Repulsion?	ELISA	(*Ozkan et al., 2013*)
beat-Vc	side-VI	Repulsion?	SPR, ELISA	(*Li et al., 2017*, *Ozkan et al., 2013*)
beat-VI	side-II	Repulsion?	SPR	(*Li et al., 2017*)
CG13506	CG7702		ELISA	(*Ozkan et al., 2013*)
CG17839	CG17839		ELISA	(*Ozkan et al., 2013*)
CG17839	beat-IV	Repulsion?	SPR, ELISA	(*Ozkan et al., 2013*, *Ozkan et al., 2013*)
side-II	side-III	Repulsion?	ELISA	(*Ozkan et al., 2013*)
side-II	beat-VI	Repulsion?	SPR	(*Li et al., 2017*)
Dscam4	Dscam4	Repulsion?	ELISA	(*Ozkan et al., 2013*)
dpr10	DIP-α	Innervation (visual system)	In vivo: Genetics using T2A-Gal4 lines (interaction) and physiological outcome by genetics (phenocopy mutant)	Ashley et al., 2019
dpr1	DIP-ι	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*)
dpr1	DIP-η	?	ELISA, SPR, X-Ray	(*Cheng et al., 2019*, *Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr1	DIP-θ	?	ELISA	(*Carrillo et al., 2015*, *Ozkan et al., 2013*)
Dscam1	Dscam1	Repulsion?	CoIP, WB, X-ray, AT, ELISA	(*Li et al., 2016*, *Ozkan et al., 2013*, *Wu et al., 2012*, *Meijers et al., 2007*, *Millard et al., 2007*, *Wojtowicz et al., 2004*)
Dscam1	sli	Repulsion?	CoIP, WB, AT	(*Alavi et al., 2016*, *Dascenco et al., 2015*)
Dscam1	NetB	Repulsion?	CoIP, WB, AT	(*Alavi et al., 2016*, *Andrews et al., 2008*)
Dscam1	Ptp69D	Repulsion?	CoIP, WB	(*Dascenco et al., 2015*)
Dscam1	robo1	Repulsion?	CoIP, WB	(*Alavi et al., 2016*)
ed	Nrg		CoIP, FM, WB	(*Islam et al., 2003*)
ed	fred		ELISA	(*Ozkan et al., 2013*)
ed	tutl		ELISA	(*Ozkan et al., 2013*)

ed	ed		PD, FA, ELISA, CoIP, WB	(*Ozkan et al., 2013*, *Yue et al., 2012*, *Lin et al., 2007*)
kirre	Src64B		Co-IP, WB	(*Tutor et al., 2014*
kirre	hbs		ELISA, BA, FC, FM, Co-IP, WB	(*Ozkan et al., 2013*, *Bao et al., 2010*, *Shelton et al., 2009*, *Dworak et al., 2001*)
kirre	sns		ELISA, BA, FM, Co-IP, WB	(*Ozkan et al., 2013*, *Bao et al., 2010*, *Shelton et al., 2009*, *Galletta et al., 2004*, *Dworak et al., 2001*)
hig	Hasp		Co-IP, WB	(*Nakayama et al., 2016*)
hig	side-VI		ELISA	(*Ozkan et al., 2013*)
Dscam3	Dscam3	Repulsion?	Co-IP, WB, ELISA	(*Ozkan et al., 2013*, *Millard et al., 2007*)
side-IV	beat-IIa	Repulsion?	ELISA	(*Ozkan et al., 2013*)
side-IV	beat-IIb	Repulsion?	ELISA	(*Ozkan et al., 2013*)
dpr16	DIP-γ	?	SPR, ELISA	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr16	DIP-ε	?	ELISA	(*Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr16	DIP-ζ	?	ELISA	(*Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr7	DIP-η	?	ELISA	(*Carrillo et al., 2015*, *Ozkan et al., 2013*)
dpr7	DIP-θ	?	ELISA	(*Carrillo et al., 2015*, *Ozkan et al., 2013*)
fra	fra		Co-IP, WB	(*Garbe et al., 2007*)
fra	NetB		ELISA	(*Ozkan et al., 2013*)
fra	Trim9		TH	(*Morikawa et al., 2011*)
CG33543	Fas2		ELISA	(*Ozkan et al., 2013*)
DIP-β	dpr9	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
DIP-β	dpr6	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
DIP-β	dpr8	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)
DIP-β	dpr21	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*)

nrm	nrm		BA, FM	(*Kania et al., 1993*)
nrm	CG5597		ELISA	(*Ozkan et al., 2013*)
sns	sns		Co-IP, WB	(*Shelton et al., 2009*)
sns	hbs		Co-IP, WB	(*Shelton et al., 2009*)
sns	kirre		Co-IP, WB, FM, ELISA	(*Ozkan et al., 2013*, *Bao et al., 2010*, *Shelton et al., 2009*, *Galletta et al., 2004*, *Dworak et al., 2001*)
sns	rst		CO-IP, FM, BA	(*Ozkan et al., 2013*, *Galletta et al., 2004*)
sns	unc-5		ELISA, SPR	(*Ozkan et al., 2013*, *Ozkan et al., 2013*)
Fas2	Fas2	Adhesion	BA, FM, ELISA	(*Ozkan et al., 2013*, *Kania et al., 1993*)
Fas2	CG33543		ELISA	(*Ozkan et al., 2013*)
Fas2	fipi		ELISA	(*Ozkan et al., 2013*)
robo1	robo1		MS, Co-IP, WB, AR	(*Evans et al., 2015*, *Evans and Bashaw, 2010*, *Simpson et al., 2000*)
robo1	Dscam1		Co-IP, WB	(*Alavi et al., 2016*)
robo1	robo2		WB, Co-IP, AR, PD	(*Evans et al., 2015*, *Simpson et al., 2000*)
robo1	robo3		Co-IP, WB	(*Evans et al., 2015*)
robo1	sli		Co-IP, WB, AT, FM, MS, SPR, AR	(*Brown et al., 2018*, *Bhat, 2017*, *Manavalan et al., 2017*, *Alavi et al., 2016*, *Reichert et al., 2016*, *Brown et al., 2015*, *Evans et al., 2015*, *Harpaz et al., 2013*, *Evans and Bashaw, 2010*, *Fukuhara et al., 2008*, *Hussain et al., 2006*, *Howitt et al., 2004*, *Battye et al., 2001*, *Brose et al., 1999*)
side	beat-Ia	Repulsion?	Co-IP, WB, AT, FM, MS, SPR, AR, ELISA	(*Ozkan et al., 2013*, *Ozkan et al., 2013*, *Siebert et al., 2009*)
side	beat-IIb	Repulsion?	ELISA	(*Ozkan et al., 2013*)
side	beat-IIa	Repulsion?	ELISA	(*Ozkan et al., 2013*)
kek6	kek6	?	CoIP, WB	(*MacLaren et al., 2004*)
kek6	Vap33	?	CoIP, WB	*Ulian-Benitez et al., 2017*)
kek6	kek1	?	CoIP, WB	(*MacLaren et al., 2004*)

kek6	kek2	?	CoIP, WB	(*MacLaren et al., 2004*)
kek6	kek5	?	CoIP, WB	(*MacLaren et al., 2004*)
Ptp99A	Ptp99A		ELISA	(*Ozkan et al., 2013*)
Ptp99A	CG11110		TH	(*Bugga et al., 2009*)
Ptp99A	InR		SPR	(*Madan et al., 2011*)
dpr21	dpr21	?	CS, MW	(*Cosmanescu et al., 2018*)
dpr21	DIP-β	?	ELISA, SRP	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*)
dpr21	cDIP	?	ELISA	(*Carrillo et al., 2015*)
CG5597	nrm		ELISA	(*Ozkan et al., 2013*)
bdl	tutl		Co-IP, WB	(*Cameron et al., 2013*)
dpr14	DIP-ζ	?	SPR	(*Cosmanescu et al., 2018*)
Ama	Ama		EM, MS, ELISA, Co-SM	(*Ozkan et al., 2013*, *Zeev-Ben-Mordehai et al., 2009*)
dpr13	DIP-ζ	?	ELISA, SPR	(*Cosmanescu et al., 2018*, *Carrillo et al., 2015*, *Ozkan et al., 2013*)

8. ABBRIVIATIONS

aCC	anterior located corner cells
AP	anterior posterior axis
As-c	achaete-scute
BMP	bone morphogenetic protein
CNS	central nervous system
Co-IP	Co-immunoprecipitation
CPS	cephalopharyngeal
CSP	cell surface protein
DNA	deoxyribonucleic acid
DV	dorsal ventral axis
ELISA	enzyme-linked immunosorbent assay
FA	autoradiography
GFP	green fluorescent protein
GMC	ganglion mother cells
Ig	Immunoglobulin
ISN	intersegmental nerve
LR	labial retractor muscle
MB	mushroom body
MHD	mouth hook depressor
MHE	mouth hook elevator
MN	motoneurons
MS	molecular sieving
NB	neuroblast
PCA	principle component analysis
PCR	polymerase chain reaction
PD	pull down
RFP	red fluorescence protein
RNA	ribonucleic acid
RNAi	RNA interference
RT	reverse transcription
scRNA-Seq	single cell RNA sequencing
SN	segmental nerve

SOG	subesophageal ganglion
SPA	solid phase assay
SPR	surface plasmon resonance
TF	transcription factors
TN	transverse nerve
t-SNE	t-distributed stochastic neighbor embedding
UAS	upstream activating sequence
VNC	ventral nerve cord
WB	western blot

9. REFERENCES

Achim, K., Pettit, J. B., Saraiva, L. R., Gavriouchkina, D., Larsson, T., Arendt, D., & Marioni, J. C. (2015). High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nature Biotechnology*, *33*(5), 503–509. https://doi.org/10.1038/nbt.3209

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., … Venter, J. C. (2000). The genome sequence of Drosophila melanogaster. *Science (New York, N.Y.)*, *287*(5461), 2185–2195. https://doi.org/10.1126/science.287.5461.2185

Akam, M. (1987). The molecular basis for metameric pattern in the Drosophila embryo. *Development*, *101*(1), 1–22.

Allan, D. W., & Thor, S. (2015a). Transcriptional selectors, masters, and combinatorial codes: Regulatory principles of neural subtype specification. *Wiley Interdisciplinary Reviews: Developmental Biology*, *4*(5), 505–528. https://doi.org/10.1002/wdev.191

Allan, D. W., & Thor, S. (2015b). Transcriptional selectors, masters, and combinatorial codes: Regulatory principles of neural subtype specification. *Wiley Interdisciplinary Reviews: Developmental Biology*, *4*(5), 505–528. https://doi.org/10.1002/wdev.191

Allen, A. M., Neville, M. C., Birtles, S., Croset, V., Treiber, C. D., Waddell, S., Goodwin, S. F., & Mann, R. S. (2020). A single-cell transcriptomic atlas of the adult Drosophila ventral nerve cord. *ELife*, *9*, 1–32. https://doi.org/10.7554/eLife.54074

Angelini, D. R., & Kaufman, T. C. (2005). Comparative developmental genetics and the evolution of arthropod body plans. *Annual Review of Genetics*, *39*, 95–119. https://doi.org/10.1146/annurev.genet.39.073003.112310

Arendt, D, & Nübler-Jung, K. (1996). Common ground plans in early brain development in mice and flies. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, *18*(3), 255–259. https://doi.org/10.1002/bies.950180314

Arendt, Detlev, Musser, J. M., Baker, C. V. H., Bergman, A., Cepko, C., Erwin, D. H., Pavlicev, M., Schlosser, G., Widder, S., Laubichler, M. D., & Wagner, G. P. (2016). The origin and evolution of cell types. *Nature Reviews. Genetics*, *17*(12), 744–757. https://doi.org/10.1038/nrg.2016.127

Arendt, Detlev, Tosches, M. A., & Marlow, H. (2016). From nerve net to nerve ring, nerve

cord and brain-evolution of the nervous system. *Nature Reviews Neuroscience*, *17*(1), 61–72. https://doi.org/10.1038/nrn.2015.15

Baccin, C., Al-Sabah, J., Velten, L., Helbling, P. M., Grünschläger, F., Hernández-Malmierca, P., Nombela-Arrieta, C., Steinmetz, L. M., Trumpp, A., & Haas, S. (2020). Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. *Nature Cell Biology*, *22*(1), 38–48. https://doi.org/10.1038/s41556-019-0439-6

Baek, M., Menon, V., Jessell, T. M., Hantman, A. W., & Dasen, J. S. (2019). Molecular Logic of Spinocerebellar Tract Neuron Diversity and Connectivity. *Cell Reports*, *27*(9), 2620-2635.e4. https://doi.org/10.1016/j.celrep.2019.04.113

Bageritz, J., Willnow, P., Valentini, E., Leible, S., Boutros, M., & Teleman, A. A. (2019). Gene expression atlas of a developing tissue by single cell expression correlation analysis. *Nature Methods*, *16*(8), 750–756. https://doi.org/10.1038/s41592-019-0492-x

Bantignies, F., Roure, V., Comet, I., Leblanc, B., Schuettengruber, B., Bonnet, J., Tixier, V., Mas, A., & Cavalli, G. (2011). Polycomb-Dependent Regulatory Contacts between Distant Hox Loci in Drosophila. *Cell*, *144*(2), 214–226. https://doi.org/https://doi.org/10.1016/j.cell.2010.12.026

Bhat, K. M. (1999). Segment polarity genes in neuroblast formation and identity specification during Drosophila neurogenesis. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, *21*(6), 472–485. https://doi.org/10.1002/(SICI)1521-1878(199906)21:6<472::AID-BIES4>3.0.CO;2-W

Bossing, T., Udolph, G., Doe, C. Q., & Technau, G. M. (1996). The embryonic central nervous system lineages of Drosophila melanogaster. I. Neuroblast lineages derived from the ventral half of the neuroectoderm. *Developmental Biology*, *179*(1), 41–64. https://doi.org/10.1006/dbio.1996.0240

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525–527. https://doi.org/10.1038/nbt.3519

Brennecke, P., Anders, S., Kim, J. K., Kołodziejczyk, A. A., Zhang, X., Proserpio, V., Baying, B., Benes, V., Teichmann, S. A., Marioni, J. C., & Heisler, M. G. (2013). Accounting for technical noise in single-cell RNA-seq experiments. *Nature Methods*, *10*(11), 1093–1095. https://doi.org/10.1038/nmeth.2645

BRIDGES, C. B. (1935). SALIVARY CHROMOSOME MAPS: With a Key to the Banding of the Chromosomes of Drosophila Melanogaster. *Journal of Heredity*, *26*(2), 60–64. https://doi.org/10.1093/oxfordjournals.jhered.a104022

Broihier, H. T., Kuzin, A., Zhu, Y., Odenwald, W., & Skeath, J. B. (2004). Drosophila homeodomain protein Nkx6 coordinates motoneuron subtype identity and axonogenesis. *Development (Cambridge, England)*, *131*(21), 5233–5242. https://doi.org/10.1242/dev.01394

Broihier, H. T., & Skeath, J. B. (2002). Drosophila homeodomain protein dHb9 directs neuronal fate via crossrepressive and cell-nonautonomous mechanisms. *Neuron*, *35*(1), 39–50. https://doi.org/10.1016/s0896-6273(02)00743-2

Brown, J. B., Boley, N., Eisman, R., May, G. E., Stoiber, M. H., Duff, M. O., Booth, B. W., Wen, J., Park, S., Suzuki, A. M., Wan, K. H., Yu, C., Zhang, D., Carlson, J. W., Cherbas, L., Eads, B. D., Miller, D., Mockaitis, K., Roberts, J., … Celniker, S. E. (2014). Diversity and dynamics of the Drosophila transcriptome. *Nature*, *512*(7515), 393–399. https://doi.org/10.1038/nature12962

Budnik, V., Zhong, Y., & Wu, C. F. (1990). Morphological plasticity of motor axons in Drosophila mutants with altered excitability. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *10*(11), 3754–3768. https://doi.org/10.1523/JNEUROSCI.10-11-03754.1990

Buelow, H. E., Tjoe, N., & Hobert, O. (2005). HOX genes define position-specific projection patterns of ventral nerve cord motor neurons. In *International Worm Meeting*.

Campos-Ortega, J. A., & Hartenstein, V. (1997). The Embryonic Development of Drosophila melanogaster. *The Embryonic Development of Drosophila Melanogaster*. https://doi.org/10.1007/978-3-662-22489-2

Carrillo, R. A., Özkan, E., Menon, K. P., Nagarkar-Jaiswal, S., Lee, P. T., Jeon, M., Birnbaum, M. E., Bellen, H. J., Garcia, K. C., & Zinn, K. (2015). Control of Synaptic Connectivity by a Network of Drosophila IgSF Cell Surface Proteins. *Cell*, *163*(7), 1770–1782. https://doi.org/10.1016/j.cell.2015.11.022

Chan, S. K., & Mann, R. S. (1996). A structural model for a homeotic protein-extradenticle-DNA complex accounts for the choice of HOX protein in the heterodimer. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(11), 5223–5228. https://doi.org/10.1073/pnas.93.11.5223

Chen, E. H., & Olson, E. N. (2001). Antisocial, an Intracellular Adaptor Protein, Is Required for Myoblast Fusion in Drosophila. *Developmental Cell*, *1*(5), 705–715. https://doi.org/10.1016/S1534-5807(01)00084-3

Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., & Zhuang, X. (2015). Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*, *348*(6233), 1360–1363. https://doi.org/10.1126/science.aaa6090

Cheng, S., Ashley, J., Kurleto, J. D., Lobb-Rabe, M., Park, Y. J., Carrillo, R. A., & Özkan, E. (2019). Molecular basis of synaptic specificity by immunoglobulin superfamily receptors in Drosophila. *ELife*, *8*, 1–27. https://doi.org/10.7554/eLife.41028

Cheutin, T., & Cavalli, G. (2018). Loss of PRC1 induces higher-order opening of Hox loci independently of transcription during Drosophila embryogenesis. *Nature Communications*, *9*(1), 3898. https://doi.org/10.1038/s41467-018-05945-4

Chourrout, D., Delsuc, F., Chourrout, P., Edvardsen, R. B., Rentzsch, F., Renfer, E., Jensen, M. F., Zhu, B., de Jong, P., Steele, R. E., & Technau, U. (2006). Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. *Nature*, *442*(7103), 684–687. https://doi.org/10.1038/nature04863

Courgeon, M., & Desplan, C. (2019). Coordination between stochastic and deterministic specification in the Drosophila visual system. *Science (New York, N.Y.)*, *366*(6463). https://doi.org/10.1126/science.aay6727

Couto, A., Alenius, M., & Dickson, B. J. (2005). Molecular, anatomical, and functional organization of the Drosophila olfactory system. *Current Biology : CB*, *15*(17), 1535–1547. https://doi.org/10.1016/j.cub.2005.07.034

Couton, L., Mauss, A. S., Yunusov, T., Diegelmann, S., Evers, J. F., & Landgraf, M. (2015). Development of connectivity in a motoneuronal network in Drosophila larvae. *Current Biology*, *25*(5), 568–576. https://doi.org/10.1016/j.cub.2014.12.056

D'Elia, K. P., & Dasen, J. S. (2018). Development, functional organization, and evolution of vertebrate axial motor circuits. In *Neural Development*. https://doi.org/10.1186/s13064-018-0108-7

Dasen, J. S. (2018). Evolution of Locomotor Rhythms. *Trends in Neurosciences*, *41*(10), 648–651. https://doi.org/10.1016/j.tins.2018.07.013

Dasen, J. S., Tice, B. C., Brenner-Morton, S., & Jessell, T. M. (2005). A Hox regulatory network establishes motor neuron pool identity and target-muscle connectivity. *Cell*.

https://doi.org/10.1016/j.cell.2005.09.009

Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, Ł., Aibar, S., Makhzami, S., Christiaens, V., Bravo González-Blas, C., Poovathingal, S., Hulselmans, G., Spanier, K. I., Moerman, T., Vanspauwen, B., Geurs, S., Voet, T., Lammertyn, J., Thienpont, B., … Aerts, S. (2018). A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. *Cell*. https://doi.org/10.1016/j.cell.2018.05.057

De Wit, J., & Ghosh, A. (2016). Specification of synaptic connectivity by cell surface interactions. *Nature Reviews Neuroscience*, *17*(1), 22–35. https://doi.org/10.1038/nrn.2015.3

Deneris, E. S., & Hobert, O. (2014). Maintenance of postmitotic neuronal cell identity. *Nature Neuroscience*, *17*(7), 899–907. https://doi.org/10.1038/nn.3731

Denes, A. S., Jékely, G., Steinmetz, P. R. H., Raible, F., Snyman, H., Prud'homme, B., Ferrier, D. E. K., Balavoine, G., & Arendt, D. (2007). Molecular architecture of annelid nerve cord supports common origin of nervous system centralization in bilateria. *Cell*, *129*(2), 277–288. https://doi.org/10.1016/j.cell.2007.02.040

Doe, C. Q. (1992). Molecular markers for identified neuroblasts and ganglion mother cells in the Drosophila central nervous system. *Development (Cambridge, England)*, *116*(4), 855–863.

Domsch, K., Carnesecchi, J., Disela, V., Friedrich, J., Trost, N., Ermakova, O., Polychronidou, M., & Lohmann, I. (2019). The hox transcription factor ubx stabilizes lineage commitment by suppressing cellular plasticity in drosophila. *ELife*. https://doi.org/10.7554/eLife.42675

Drachman, D. A. (2005). Do we have brain to spare? *Neurology*, *64*(12), 2004 LP – 2005. https://doi.org/10.1212/01.WNL.0000166914.38327.BB

Ecker, J. R., Geschwind, D. H., Kriegstein, A. R., Ngai, J., Osten, P., Polioudakis, D., Regev, A., Sestan, N., Wickersham, I. R., & Zeng, H. (2017). The BRAIN Initiative Cell Census Consortium: Lessons Learned toward Generating a Comprehensive Brain Cell Atlas. *Neuron*, *96*(3), 542–557. https://doi.org/10.1016/j.neuron.2017.10.007

Eng, C.-H. L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.-C., & Cai, L. (2019). Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature*, *568*(7751), 235–239. https://doi.org/10.1038/s41586-019-1049-y

Enriquez, J., Venkatasubramanian, L., Baek, M., Peterson, M., Aghayeva, U., & Mann, R. S. (2015). Specification of individual adult motor neuron morphologies by combinatorial transcription factor codes. *Neuron*, *86*(4), 955–970. https://doi.org/10.1016/j.neuron.2015.04.011

Friedrich, J., Sorge, S., Bujupi, F., Eichenlaub, M. P., Schulz, N. G., Wittbrodt, J., & Lohmann, I. (2016). Hox Function Is Required for the Development and Maintenance of the Drosophila Feeding Motor Unit. *Cell Reports*, *14*(4), 850–860. https://doi.org/10.1016/j.celrep.2015.12.077

Fuqua, T., Jordan, J., van Breugel, M. E., Halavatyi, A., Tischer, C., Polidoro, P., Abe, N., Tsai, A., Mann, R. S., Stern, D. L., & Crocker, J. (2020). Dense and pleiotropic regulatory information in a developmental enhancer. *Nature*, *587*(7833), 235–239. https://doi.org/10.1038/s41586-020-2816-5

Galliot, B., & Quiquand, M. (2011). A two-step process in the emergence of neurogenesis. *The European Journal of Neuroscience*, *34*(6), 847–862. https://doi.org/10.1111/j.1460-9568.2011.07829.x

Garces, A., & Thor, S. (2006). Specification of Droscophila aCC motoneuron identity by a genetic cascade involving even-skipped, grain and zfh1. *Development*, *133*(8), 1445–1455. https://doi.org/10.1242/dev.02321

García-Bellido, A. (1975). Genetic control of wing disc development in Drosophila. *Ciba Foundation Symposium*, *0*(29), 161–182. https://doi.org/10.1002/9780470720110.ch8

Garrett, A. M., Khalil, A., Walton, D. O., & Burgess, R. W. (2018). DSCAM promotes self-avoidance in the developing mouse retina by masking the functions of cadherin superfamily members. *Proceedings of the National Academy of Sciences*, *115*(43), E10216 LP-E10224. https://doi.org/10.1073/pnas.1809430115

Gebelein, B., McKay, D. J., & Mann, R. S. (2004). Direct integration of Hox and segmentation gene inputs during Drosophila development. *Nature*, *431*(7009), 653–659. https://doi.org/10.1038/nature02946

Goodman, C. S., Bastiani, M. J., Doe, C. Q., du Lac, S., Helfand, S. L., Kuwada, J. Y., & Thomas, J. B. (1984). Cell recognition during neuronal development. *Science (New York, N.Y.)*, *225*(4668), 1271–1279. https://doi.org/10.1126/science.6474176

Gould, A. P., Brookman, J. J., Strutt, D. I., & White, R. A. H. (1990). Targets of homeotic gene control in Drosophila. *Nature*, *348*(6299), 308–312.

https://doi.org/10.1038/348308a0

Hagemann-Jensen, M., Ziegenhain, C., Chen, P., Ramsköld, D., Hendriks, G.-J., Larsson, A. J. M., Faridani, O. R., & Sandberg, R. (2020). Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nature Biotechnology*, *38*(6), 708–714. https://doi.org/10.1038/s41587-020-0497-0

Hales, K. G., Korey, C. A., Larracuente, A. M., & Roberts, D. M. (2015). Genetics on the fly: A primer on the drosophila model system. *Genetics*, *201*(3), 815–842. https://doi.org/10.1534/genetics.115.183392

Harding, K., Wedeen, C., McGinnis, W., & Levine, M. (1985). Spatially regulated expression of homeotic genes in Drosophila. *Science (New York, N.Y.)*, *229*(4719), 1236–1242. https://doi.org/10.1126/science.3898362

Hartenstein, V., & Wodarz, A. (2013). Initial neurogenesis in Drosophila. *Wiley Interdisciplinary Reviews. Developmental Biology*, *2*(6), 823. https://doi.org/10.1002/wdev.117

Hennig, B. P., Velten, L., Racke, I., Tu, C. S., Thoms, M., Rybin, V., Besir, H., Remans, K., & Steinmetz, L. M. (2018). Large-Scale Low-Cost NGS Library Preparation Using a Robust Tn5 Purification and Tagmentation Protocol. *G3 (Bethesda, Md.)*, *8*(1), 79–89. https://doi.org/10.1534/g3.117.300257

Hessinger, C., Technau, G. M., & Rogulja-Ortmann, A. (2017). The Drosophila Hox gene Ultrabithorax acts in both muscles and motoneurons to orchestrate formation of specific neuromuscular connections. *Development (Cambridge)*, *144*(1), 139–150. https://doi.org/10.1242/dev.143875

Hoang, B., & Chiba, A. (2001). Single-cell analysis of Drosophila larval neuromuscular synapses. *Developmental Biology*, *229*(1), 55–70. https://doi.org/10.1006/dbio.2000.9983

Holland, N. D. (2003). Early central nervous system evolution: an era of skin brains? *Nature Reviews. Neuroscience*, *4*(8), 617–627. https://doi.org/10.1038/nrn1175

Huang, L., Kebschull, J. M., Fürth, D., Musall, S., Kaufman, M. T., Churchland, A. K., & Zador, A. M. (2020). BRICseq Bridges Brain-wide Interregional Connectivity to Neural Activity and Gene Expression in Single Animals. *Cell*, *182*(1), 177-188.e27. https://doi.org/https://doi.org/10.1016/j.cell.2020.05.029

Hückesfeld, S., Schoofs, A., Schlegel, P., Miroschnikow, A., & Pankratz, M. J. (2015).

Localization of Motor Neurons and Central Pattern Generators for Motor Patterns Underlying Feeding Behavior in Drosophila Larvae. *PloS One*, *10*(8), e0135011–e0135011. https://doi.org/10.1371/journal.pone.0135011

Hueber, S. D., Bezdan, D., Henz, S. R., Blank, M., Wu, H., & Lohmann, I. (2007). Comparative analysis of Hox downstream genes in Drosophila. *Development (Cambridge, England)*, *134*(2), 381–392. https://doi.org/10.1242/dev.02746

Hung, R.-J., Hu, Y., Kirchner, R., Liu, Y., Xu, C., Comjean, A., Tattikota, S. G., Li, F., Song, W., Ho Sui, S., & Perrimon, N. (2020). A cell atlas of the adult Drosophila midgut. *Proceedings of the National Academy of Sciences*, *117*(3), 1514 LP – 1523. https://doi.org/10.1073/pnas.1916820117

Janky, R., Verfaillie, A., Imrichová, H., Van de Sande, B., Standaert, L., Christiaens, V., Hulselmans, G., Herten, K., Naval Sanchez, M., Potier, D., Svetlichnyy, D., Kalender Atak, Z., Fiers, M., Marine, J.-C., & Aerts, S. (2014). iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS Computational Biology*, *10*(7), e1003731. https://doi.org/10.1371/journal.pcbi.1003731

Jin, X., Simmons, S. K., Guo, A., Shetty, A. S., Ko, M., Nguyen, L., Jokhi, V., Robinson, E., Oyler, P., Curry, N., Deangeli, G., Lodato, S., Levin, J. Z., Regev, A., Zhang, F., & Arlotta, P. (2020). In vivo Perturb-Seq reveals neuronal and glial abnormalities associated with autism risk genes. *Science*, *370*(6520), eaaz6063. https://doi.org/10.1126/science.aaz6063

Jolma, A., Yin, Y., Nitta, K. R., Dave, K., Popov, A., Taipale, M., Enge, M., Kivioja, T., Morgunova, E., & Taipale, J. (2015). DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature*, *527*(7578), 384–388. https://doi.org/10.1038/nature15518

Kania, A., & Jessell, T. M. (2003). Topographic Motor Projections in the Limb Imposed by LIM Homeodomain Protein Regulation of Ephrin-A:EphA Interactions. *Neuron*, *38*(4), 581–596. https://doi.org/10.1016/S0896-6273(03)00292-7

Kim, M. D., Wen, Y., & Jan, Y.-N. (2009). Patterning and organization of motor neuron dendrites in the Drosophila larva. *Developmental Biology*, *336*(2), 213–221. https://doi.org/10.1016/j.ydbio.2009.09.041

Kim, N. C., & Marqués, G. (2012). The Ly6 neurotoxin-like molecule target of wit regulates spontaneous neurotransmitter release at the developing neuromuscular junction in

Drosophila. *Developmental Neurobiology*, *72*(12), 1541–1558. https://doi.org/10.1002/dneu.22021

Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A., & Kirschner, M. W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*, *161*(5), 1187–1201. https://doi.org/10.1016/j.cell.2015.04.044

Kohsaka, H., & Nose, A. (2009). Target recognition at the tips of postsynaptic filopodia: accumulation and function of Capricious. *Development (Cambridge, England)*, *136*(7), 1127–1135. https://doi.org/10.1242/dev.027920

Kolodziej, P. A., Timpe, L. C., Mitchell, K. J., Fried, S. R., Goodman, C. S., Jan, L. Y., & Jan, Y. N. (1996). frazzled encodes a Drosophila member of the DCC immunoglobulin subfamily and is required for CNS and motor axon guidance. *Cell*, *87*(2), 197–204. https://doi.org/10.1016/s0092-8674(00)81338-0

Kornberg, R. (1981). The location of nucleosomes in chromatin: specific or statistical? *Nature*, *292*(5824), 579–580. https://doi.org/10.1038/292579a0

Kribelbauer, J. F., Loker, R. E., Feng, S., Rastogi, C., Abe, N., Rube, H. T., Bussemaker, H. J., & Mann, R. S. (2020). Context-Dependent Gene Regulation by Homeodomain Transcription Factor Complexes Revealed by Shape-Readout Deficient Proteins. *Molecular Cell*, *78*(1), 152-167.e11. https://doi.org/10.1016/j.molcel.2020.01.027

Kulkarni, A., Ertekin, D., Lee, C.-H., & Hummel, T. (2016). Birth order dependent growth cone segregation determines synaptic layer identity in the Drosophila visual system. *ELife*, *5*, e13715. https://doi.org/10.7554/eLife.13715

Kunst, M., Laurell, E., Mokayes, N., Kramer, A., Kubo, F., Fernandes, A. M., Förster, D., Dal Maschio, M., & Baier, H. (2019). A Cellular-Resolution Atlas of the Larval Zebrafish Brain. *Neuron*, *103*(1), 21-38.e5. https://doi.org/10.1016/j.neuron.2019.04.034

Kurusu, M., Cording, A., Taniguchi, M., Menon, K., Suzuki, E., & Zinn, K. (2008). A screen of cell-surface molecules identifies leucine-rich repeat proteins as key mediators of synaptic target selection. *Neuron*, *59*(6), 972–985. https://doi.org/10.1016/j.neuron.2008.07.037

Labrador, J. P., O'Keefe, D., Yoshikawa, S., McKinnon, R. D., Thomas, J. B., & Bashaw, G. J. (2005). The homeobox transcription factor even-skipped regulates netrin-receptor expression to control dorsal motor-axon projections in Drosophila. *Current Biology*,

15(15), 1413–1419. https://doi.org/10.1016/j.cub.2005.06.058

Landgraf, M, Roy, S., Prokop, A., VijayRaghavan, K., & Bate, M. (1999). even-skipped determines the dorsal growth of motor axons in Drosophila. *Neuron*, *22*(1), 43–52. https://doi.org/10.1016/s0896-6273(00)80677-7

Landgraf, Matthias, Bossing, T., Technau, G. M., & Bate, M. (1997a). The origin, location, and projections of the embryonic abdominal motorneurons of Drosophila. *Journal of Neuroscience*, *17*(24), 9642–9655. https://doi.org/10.1523/jneurosci.17-24-09642.1997

Landgraf, Matthias, Bossing, T., Technau, G. M., & Bate, M. (1997b). The origin, location, and projections of the embryonic abdominal motorneurons of Drosophila. *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.17-24-09642.1997

Landgraf, Matthias, Jeffrey, V., Fujioka, M., Jaynes, J. B., & Bate, M. (2003). Embryonic origins of a motor system: Motor dendrites form a myotopic map in Drosophila. *PLoS Biology*, *1*(2). https://doi.org/10.1371/journal.pbio.0000041

Landgraf, Matthias, & Thor, S. (2006). Development and Structure of Motoneurons. *International Review of Neurobiology*, *75*(06), 33–53. https://doi.org/10.1016/S0074-7742(06)75002-4

Lee, J. H., Daugharthy, E. R., Scheiman, J., Kalhor, R., Yang, J. L., Ferrante, T. C., Terry, R., Jeanty, S. S. F., Li, C., Amamoto, R., Peters, D. T., Turczyk, B. M., Marblestone, A. H., Inverso, S. A., Bernard, A., Mali, P., Rios, X., Aach, J., & Church, G. M. (2014). Highly Multiplexed Subcellular RNA Sequencing in Situ. *Science*, *343*(6177), 1360 LP – 1363. https://doi.org/10.1126/science.1250212

Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. *Nature*, *276*(5688), 565–570. https://doi.org/10.1038/276565a0

Li, H., Horns, F., Xie, Q., Xie, Q., Li, T., Luginbuhl, D. J., Luo, L., & Quake, S. R. (2017). Classifying Drosophila Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing. *Cell*, *171*(5), 1206.e22-1220.e22. https://doi.org/10.1016/j.cell.2017.10.019

Li, H., Li, T., Horns, F., Li, J., Xie, Q., Xu, C., Wu, B., Kebschull, J. M., McLaughlin, C. N., Kolluru, S. S., Jones, R. C., Vacek, D., Xie, A., Luginbuhl, D. J., Quake, S. R., & Luo, L. (2020). Single-Cell Transcriptomes Reveal Diverse Regulatory Strategies for Olfactory Receptor Expression and Axon Targeting. *Current Biology*, *30*(7), 1189-1198.e5. https://doi.org/10.1016/j.cub.2020.01.049

Li, J., Han, S., Li, H., Udeshi, N. D., Svinkina, T., Mani, D. R., Xu, C., Guajardo, R., Xie, Q.,

Li, T., Luginbuhl, D. J., Wu, B., McLaughlin, C. N., Xie, A., Kaewsapsak, P., Quake, S. R., Carr, S. A., Ting, A. Y., & Luo, L. (2020). Cell-Surface Proteomic Profiling in the Fly Brain Uncovers Wiring Regulators. *Cell*, *180*(2), 373-386.e15. https://doi.org/10.1016/j.cell.2019.12.029

Lloyd, T. E., & Taylor, J. P. (2010). Flightless flies: Drosophila models of neuromuscular disease. *Annals of the New York Academy of Sciences*, *1184*, e1-20. https://doi.org/10.1111/j.1749-6632.2010.05432.x

Luecken, M. D., & Theis, F. J. (2019). Current best practices in single-cell RNA-seq analysis: a tutorial. *Molecular Systems Biology*, *15*(6). https://doi.org/10.15252/msb.20188746

Luo, L., Liao, Y. J., Jan, L. Y., & Jan, Y. N. (1994). Distinct morphogenetic functions of similar small GTPases: Drosophila Drac1 is involved in axonal outgrowth and myoblast fusion. *Genes & Development*, *8*(15), 1787–1802. https://doi.org/10.1101/gad.8.15.1787

Luo, Y, Raible, D., & Raper, J. A. (1993). Collapsin: a protein in brain that induces the collapse and paralysis of neuronal growth cones. *Cell*, *75*(2), 217–227. https://doi.org/10.1016/0092-8674(93)80064-l

Luo, Yuling, Shepherd, I., Li, J., Renzi, M. J., Chang, S., & Raper, J. A. (1995). A family of molecules related to collapsin in the embryonic chick nervous system. *Neuron*, *14*(6), 1131–1140. https://doi.org/https://doi.org/10.1016/0896-6273(95)90261-9

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A., & McCarroll, S. A. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, *161*(5), 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002

Mahr, A., & Aberle, H. (2006). The expression pattern of the Drosophila vesicular glutamate transporter: a marker protein for motoneurons and glutamatergic centers in the brain. *Gene Expression Patterns : GEP*, *6*(3), 299–309. https://doi.org/10.1016/j.modgep.2005.07.006

Mallo, M. S. (2014). Developmental Cell Preview Evolving Locomotion with Hoxc9. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2014.04.014

Mann, R. S., & Affolter, M. (1998). Hox proteins meet more partners. *Current Opinion in Genetics & Development*, *8*(4), 423–429. https://doi.org/10.1016/s0959-437x(98)80113-5

Martinez-Arias, A., & Lawrence, P. A. (1985). Parasegments and compartments in the Drosophila embryo. *Nature*, *313*(6004), 639–642. https://doi.org/10.1038/313639a0

Matthes, D. J., Sink, H., Kolodkin, A. L., & Goodman, C. S. (1995). Semaphorin II can function as a selective inhibitor of specific synaptic arborizations. *Cell*, *81*(4), 631–639. https://doi.org/10.1016/0092-8674(95)90084-5

Matthews, B. J., Kim, M. E., Flanagan, J. J., Hattori, D., Clemens, J. C., Zipursky, S. L., & Grueber, W. B. (2007). Dendrite self-avoidance is controlled by Dscam. *Cell*, *129*(3), 593–604. https://doi.org/10.1016/j.cell.2007.04.013

McGinnis, W., & Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell*, *68*(2), 283–302. https://doi.org/10.1016/0092-8674(92)90471-n

McInnes, L., Healy, J., Saul, N., & Großberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, *3*(29), 861. https://doi.org/10.21105/joss.00861

Menon, K. P., Kulkarni, V., Shin-Ya, T., Anaya, M., & Zinn, K. (2019). Interactions between dpr11 and dip-y control election of amacrine neurons in drosophila color ision circuits. *ELife*, *8*, 1–32. https://doi.org/10.7554/eLife.48935

Merrill, V. K. L., Diederich, R. J., Turner, F. R., & Kaufman, T. C. (1989). A genetic and developmental analysis of mutations in labial, a gene necessary for proper head formation in Drosophila melanogaster. *Developmental Biology*, *135*(2), 376–391. https://doi.org/https://doi.org/10.1016/0012-1606(89)90187-5

Miller, G. (2009). Origins. On the origin of the nervous system. In *Science (New York, N.Y.)* (Vol. 325, Issue 5936, pp. 24–26). https://doi.org/10.1126/science.325_24

Miroschnikow, A., Schlegel, P., Schoofs, A., Hueckesfeld, S., Li, F., Schneider-Mizell, C. M., Fetter, R. D., Truman, J. W., Cardona, A., & Pankratz, M. J. (2018). Convergence of monosynaptic and polysynaptic sensory paths onto common motor outputs in a drosophila feeding connectome. *ELife*, *7*, 1–23. https://doi.org/10.7554/eLife.40247

Mitchell, K. J., Doyle, J. L., Serafini, T., Kennedy, T. E., Tessier-Lavigne, M., Goodman, C. S., & Dickson, B. J. (1996). Genetic Analysis of Netrin Genes in Drosophila: Netrins Guide CNS Commissural Axons and Peripheral Motor Axons. *Neuron*, *17*(2), 203–215. https://doi.org/https://doi.org/10.1016/S0896-6273(00)80153-1

Moncada, R., Barkley, D., Wagner, F., Chiodin, M., Devlin, J. C., Baron, M., Hajdu, C. H., Simeone, D. M., & Yanai, I. (2020). Integrating microarray-based spatial transcriptomics

and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nature Biotechnology*, *38*(3), 333–342. https://doi.org/10.1038/s41587-019-0392-8

Montavon, T., & Soshnikova, N. (2014). Hox gene regulation and timing in embryogenesis. *Seminars in Cell & Developmental Biology*, *34*, 76–84. https://doi.org/10.1016/j.semcdb.2014.06.005

Moon, K. R., van Dijk, D., Wang, Z., Gigante, S., Burkhardt, D. B., Chen, W. S., Yim, K., Elzen, A. van den, Hirn, M. J., Coifman, R. R., Ivanova, N. B., Wolf, G., & Krishnaswamy, S. (2019). Visualizing structure and transitions in high-dimensional biological data. *Nature Biotechnology*, *37*(12), 1482–1492. https://doi.org/10.1038/s41587-019-0336-3

Morata, G., & Lawrence, P. A. (1975). Control of compartment development by the engrailed gene in Drosophila. *Nature*, *255*(5510), 614–617. https://doi.org/10.1038/255614a0

Morata, G., & Lawrence, P. A. (1979). Development of the eye-antenna imaginal disc of Drosophila. *Developmental Biology*, *70*(2), 355–371. https://doi.org/10.1016/0012-1606(79)90033-2

Morgan, T. H. (1910). SEX LIMITED INHERITANCE IN DROSOPHILA. *Science (New York, N.Y.)*, *32*(812), 120–122. https://doi.org/10.1126/science.32.812.120

Nakamura, M., Baldwin, D., Hannaford, S., Palka, J., & Montell, C. (2002). Defective Proboscis Extension Response (DPR), a Member of the Ig Superfamily Required for the Gustatory Response to Salt. *Journal of Neuroscience*, *22*(9), 3463–3472. https://doi.org/10.1523/jneurosci.22-09-03463.2002

Nassif, C., Noveen, A., & Hartenstein, V. (1998). *Embryonic De ve lopme nt of the Dros op hila Brain . I . Patte rn of Pione e r Tracts*. *31*, 10–31.

Ni, J.-Q., Zhou, R., Czech, B., Liu, L.-P., Holderbaum, L., Yang-Zhou, D., Shim, H.-S., Tao, R., Handler, D., Karpowicz, P., Binari, R., Booker, M., Brennecke, J., Perkins, L. A., Hannon, G. J., & Perrimon, N. (2011). A genome-scale shRNA resource for transgenic RNAi in Drosophila. *Nature Methods*, *8*(5), 405–407. https://doi.org/10.1038/nmeth.1592

Noro, B., Culi, J., McKay, D. J., Zhang, W., & Mann, R. S. (2006). Distinct functions of homeodomain-containing and homeodomain-less isoforms encoded by homothorax. *Genes and Development*, *20*(12), 1636–1650. https://doi.org/10.1101/gad.1412606

Nose, A., Mahajan, V. B., & Goodman, C. S. (1992). Connectin: a homophilic cell adhesion molecule expressed on a subset of muscles and the motoneurons that innervate them in Drosophila. *Cell*, *70*(4), 553–567. https://doi.org/10.1016/0092-8674(92)90426-d

Nose, A., Umeda, T., & Takeichi, M. (1997). Neuromuscular target recognition by a homophilic interaction of connectin cell adhesion molecules in Drosophila. *Development (Cambridge, England)*, *124*(8), 1433–1441.

Noyes, M. B., Christensen, R. G., Wakabayashi, A., Stormo, G. D., Brodsky, M. H., & Wolfe, S. A. (2008). Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell*, *133*(7), 1277–1289. https://doi.org/10.1016/j.cell.2008.05.023

Okumura, M., Sakuma, C., Miura, M., & Chihara, T. (2015). Linking cell surface receptors to microtubules: Tubulin folding cofactor d mediates Dscam functions during neuronal morphogenesis. *Journal of Neuroscience*, *35*(5), 1979–1990. https://doi.org/10.1523/JNEUROSCI.0973-14.2015

Özkan, E., Carrillo, R. A., Eastman, C. L., Weiszmann, R., Waghray, D., Johnson, K. G., Zinn, K., Celniker, S. E., & Garcia, K. C. (2013). An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. *Cell*, *154*(1), 228–239. https://doi.org/10.1016/j.cell.2013.06.006

Perrimon, N. (1994). The genetic basis of patterned baldness in Drosophila. *Cell*, *76*(5), 781–784. https://doi.org/10.1016/0092-8674(94)90351-4

Philippidou, P., & Dasen, J. S. (2015). Sensory-Motor Circuits: Hox Genes Get in Touch. In *Neuron*. https://doi.org/10.1016/j.neuron.2015.10.035

Picelli, S., Faridani, O. R., Björklund, Å. K., Winberg, G., Sagasser, S., & Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols*, *9*(1), 171–181. https://doi.org/10.1038/nprot.2014.006

Ranganayakulu, G., Elliott, D. A., Harvey, R. P., & Olson, E. N. (1998). Divergent roles for NK-2 class homeobox genes in cardiogenesis in flies and mice. *Development (Cambridge, England)*, *125*(16), 3037–3048.

Reilly, M. B., Cros, C., Varol, E., Yemini, E., & Hobert, O. (2020). Unique homeobox codes delineate all C . elegans neuron classes. *In Press*, *February*, 1–28. https://doi.org/10.1038/s41586-020-2618-9

Reiter, L. T., Potocki, L., Chien, S., Gribskov, M., & Bier, E. (2001). A systematic analysis of

human disease-associated gene sequences in Drosophila melanogaster. *Genome Research*, *11*(6), 1114–1125. https://doi.org/10.1101/gr.169101

Replogle, J. M., Norman, T. M., Xu, A., Hussmann, J. A., Chen, J., Cogan, J. Z., Meer, E. J., Terry, J. M., Riordan, D. P., Srinivas, N., Fiddes, I. T., Arthur, J. G., Alvarado, L. J., Pfeiffer, K. A., Mikkelsen, T. S., Weissman, J. S., & Adamson, B. (2020). Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nature Biotechnology*, *38*(8), 954–961. https://doi.org/10.1038/s41587-020-0470-y

Riechmann, V., & Ephrussi, A. (2001). Axis formation during Drosophila oogenesis. *Current Opinion in Genetics & Development*, *11*(4), 374–383. https://doi.org/10.1016/s0959-437x(00)00207-0

Risso, D., Perraudeau, F., Gribkova, S., Dudoit, S., & Vert, J.-P. (2019). Publisher Correction: A general and flexible method for signal extraction from single-cell RNA-seq data. *Nature Communications*, *10*(1), 646. https://doi.org/10.1038/s41467-019-08614-2

Ritzenthaler, S, Suzuki, E., & Chiba, A. (2000). Postsynaptic filopodia in muscle cells interact with innervating motoneuron axons. *Nature Neuroscience*, *3*(10), 1012–1017. https://doi.org/10.1038/79833

Ritzenthaler, Sarah, & Chiba, A. (2003). Myopodia (postsynaptic filopodia) participate in synaptic target recognition. *Journal of Neurobiology*, *55*(1), 31–40. https://doi.org/10.1002/neu.10180

Rust, K., Byrnes, L. E., Yu, K. S., Park, J. S., Sneddon, J. B., Tward, A. D., & Nystul, T. G. (2020). A single-cell atlas and lineage analysis of the adult Drosophila ovary. *Nature Communications*, *11*(1), 5628. https://doi.org/10.1038/s41467-020-19361-0

Ryoo, H. D., Marty, T., Casares, F., Affolter, M., & Mann, R. S. (1999). Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex. *Development (Cambridge, England)*, *126*(22), 5137–5148.

Sanes, J. R., & Zipursky, S. L. (2020). Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. *Cell*, *181*(3), 536–556. https://doi.org/10.1016/j.cell.2020.04.008

Sanyal, S. (2009). Genomic mapping and expression patterns of C380, OK6 and D42 enhancer trap lines in the larval nervous system of Drosophila. *Gene Expression Patterns : GEP*, *9*(5), 371–380. https://doi.org/10.1016/j.gep.2009.01.002

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction

of single-cell gene expression data. *Nature Biotechnology*, *33*(5), 495–502. https://doi.org/10.1038/nbt.3192

Schmid, A., Chiba, A., & Doe, C. Q. (1999). Clonal analysis of Drosophila embryonic neuroblasts: neural cell types, axon projections and muscle targets. *Development (Cambridge, England)*, *126*(21), 4653–4689.

Schmidt, H., Rickert, C., Bossing, T., Vef, O., Urban, J., & Technau, G. M. (1997). The embryonic central nervous system lineages of Drosophila melanogaster. II. Neuroblast lineages derived from the dorsal part of the neuroectoderm. *Developmental Biology*, *189*(2), 186–204. https://doi.org/10.1006/dbio.1997.8660

Schoofs, A., Hanslik, U., Niederegger, S., Heinzel, H. G., & Spieß, R. (2010). The thoracic muscular system and its innervation in third instar calliphora vicina larvae. II. Projection patterns of the nerves associated with the pro- and mesothorax and the pharyngeal complex. *Journal of Morphology*, *271*(8), 969–979. https://doi.org/10.1002/jmor.10853

Shapiro, L. (2007). Self-Recognition at the Atomic Level: Understanding the Astonishing Molecular Diversity of Homophilic Dscams. *Neuron*, *56*(1), 10–13. https://doi.org/https://doi.org/10.1016/j.neuron.2007.09.024

Shishido, E., Takeichi, M., & Nose, A. (1998). Drosophila synapse formation: regulation by transmembrane protein with Leu-rich repeats, CAPRICIOUS. *Science (New York, N.Y.)*, *280*(5372), 2118–2121. https://doi.org/10.1126/science.280.5372.2118

Siegler, M. V, & Jia, X. X. (1999). Engrailed negatively regulates the expression of cell adhesion molecules connectin and neuroglian in embryonic Drosophila nervous system. *Neuron*, *22*(2), 265–276. https://doi.org/10.1016/s0896-6273(00)81088-0

Sink, H, & Whitington, P. M. (1991). Pathfinding in the central nervous system and periphery by identified embryonic Drosophila motor axons. *Development (Cambridge, England)*, *112*(1), 307–316.

Sink, Helen, & Whitington, P. M. (1991). Location and connectivity of abdominal motoneurons in the embryo and larva of Drosophila melanogaster. *Journal of Neurobiology*. https://doi.org/10.1002/neu.480220309

Skeath, J B. (1999). At the nexus between pattern formation and cell-type specification: the generation of individual neuroblast fates in the Drosophila embryonic central nervous system. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, *21*(11), 922–931. https://doi.org/10.1002/(SICI)1521-

1878(199911)21:11<922::AID-BIES4>3.0.CO;2-T

Skeath, J B, & Carroll, S. B. (1992). Regulation of proneural gene expression and cell fate during neuroblast segregation in the Drosophila embryo. *Development (Cambridge, England)*, *114*(4), 939–946.

Skeath, James B, & Carroll, S. B. (1994). The achaete-scute complex: generation of cellular pattern and fate within the Drosophila nervous system. *The FASEB Journal*, *8*(10), 714–721. https://doi.org/https://doi.org/10.1096/fasebj.8.10.8050670

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, H. J., & Mann, R. S. (2011). Cofactor binding evokes latent differences in DNA binding specificity between hox proteins. *Cell*, *147*(6), 1270–1282. https://doi.org/10.1016/j.cell.2011.10.053

Song, M. R., & Pfaff, S. L. (2005). Hox genes: The instructors working at motor pools. In *Cell*. https://doi.org/10.1016/j.cell.2005.10.014

Sorge, S., Ha, N., Polychronidou, M., Friedrich, J., Bezdan, D., Kaspar, P., Schaefer, M. H., Ossowski, S., Henz, S. R., Mundorf, J., Rätzer, J., Papagiannouli, F., & Lohmann, I. (2012). The cis-regulatory code of Hox function in Drosophila. *EMBO Journal*, *31*(15), 3323–3333. https://doi.org/10.1038/emboj.2012.179

Soroka, V., Kolkova, K., Kastrup, J. S., Diederichs, K., Breed, J., Kiselyov, V. V, Poulsen, F. M., Larsen, I. K., Welte, W., Berezin, V., Bock, E., & Kasper, C. (2003). Structure and interactions of NCAM Ig1-2-3 suggest a novel zipper mechanism for homophilic adhesion. *Structure (London, England : 1993)*, *11*(10), 1291–1301. https://doi.org/10.1016/j.str.2003.09.006

Sperry, R. W. (1963). CHEMOAFFINITY IN THE ORDERLY GROWTH OF NERVE FIBER PATTERNS AND CONNECTIONS. *Proceedings of the National Academy of Sciences of the United States of America*, *50*(4), 703–710. https://doi.org/10.1073/pnas.50.4.703

Stagg, S. B., Guardiola, A. R., & Crews, S. T. (2011). Dual role for drosophila lethal of scute in CNS midline precursor formation and dopaminergic neuron and motoneuron cell fate. *Development*, *138*(11), 2171–2183. https://doi.org/10.1242/dev.056507

Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J. O., Huss, M., Mollbrink, A., Linnarsson, S., Codeluppi, S., Borg, Å., Pontén, F., Costea, P. I., Sahlén, P., Mulder, J., Bergmann, O.,

… Frisén, J. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, *353*(6294), 78 LP – 82. https://doi.org/10.1126/science.aaf2403

Steinmetz, P. R. H., Kostyuchenko, R. P., Fischer, A., & Arendt, D. (2011). The segmental pattern of otx, gbx, and Hox genes in the annelid Platynereis dumerilii. *Evolution & Development*, *13*(1), 72–79. https://doi.org/10.1111/j.1525-142X.2010.00457.x

Stratoulias, V., & Heino, T. I. (2015). MANF silencing, immunity induction or autophagy trigger an unusual cell type in metamorphosing Drosophila brain. *Cellular and Molecular Life Sciences : CMLS*, *72*(10), 1989–2004. https://doi.org/10.1007/s00018-014-1789-7

Sugino, K., Clark, E., Schulmann, A., Shima, Y., Wang, L., Hunt, D. L., Hooks, B. M., Nkner, D. T., Chandrashekar, J., Picard, S., Lemire, A. L., Spruston, N., Hantman, A. W., & Nelson, S. B. (2019). Mapping the transcriptional diversity of genetically and anatomically defined cell populations in the mouse brain. *ELife*, *8*, 1–29. https://doi.org/10.7554/eLife.38619

Sweeney, L. B., Chou, Y., Wu, Z., Joo, W., Potter, C. J., Kolodkin, A. L., & Garcia, K. C. (2012). *NIH Public Access*. *72*(5), 734–747. https://doi.org/10.1016/j.neuron.2011.09.026.Secreted

Takeichi, M. (2018). Historical review of the discovery of cadherin, in memory of Tokindo Okada. *Development, Growth & Differentiation*, *60*(1), 3–13. https://doi.org/10.1111/dgd.12416

Tan, L., Zhang, K. X., Matthew, Y., Bellen, H. J., Morey, M., Tan, L., Zhang, K. X., Pecot, M. Y., Nagarkar-jaiswal, S., Lee, P., & Takemura, S. (2015). Ig Superfamily Ligand and Receptor Pairs Expressed in Synaptic Partners in Drosophila Article Ig Superfamily Ligand and Receptor Pairs Expressed in Synaptic Partners in Drosophila. *Cell*, *163*(7), 1756–1769. https://doi.org/10.1016/j.cell.2015.11.021

Tang, F., Barbacioru, C., Nordman, E., Bao, S., Lee, C., Wang, X., Tuch, B. B., Heard, E., Lao, K., & Surani, M. A. (2011). Deterministic and stochastic allele specific gene expression in single mouse blastomeres. *PLoS ONE*, *6*(6). https://doi.org/10.1371/journal.pone.0021208

Thiery, J. P., Brackenbury, R., Rutishauser, U., & Edelman, G. M. (1977). Adhesion among neural cells of the chick embryo. II. Purification and characterization of a cell adhesion

molecule from neural retina. *The Journal of Biological Chemistry*, *252*(19), 6841–6845.

Thor, S., Andersson, S. G., Tomlinson, A., & Thomas, J. B. (1999). A LIM-homeodomain combinatorial code for motor-neuron pathway selection. *Nature*, *397*(6714), 76–80. https://doi.org/10.1038/16275

Urbach, R., Jussen, D., & Technau, G. M. (2016). Gene expression profiles uncover individual identities of gnathal neuroblasts and serial homologies in the embryonic CNS of Drosophila. *Development (Cambridge)*, *143*(8), 1290–1301. https://doi.org/10.1242/dev.133546

Urbach, R., Volland, D., Seibert, J., & Technau, G. M. (2006). Segment-specific requirements for dorsoventral patterning genes during early brain development in Drosophila. *Development (Cambridge, England)*, *133*(21), 4315–4330. https://doi.org/10.1242/dev.02605

van Oostrum, M., Campbell, B., Seng, C., Müller, M., tom Dieck, S., Hammer, J., Pedrioli, P. G. A., Földy, C., Tyagarajan, S. K., & Wollscheid, B. (2020). Surfaceome dynamics reveal proteostasis-independent reorganization of neuronal surface proteins during development and synaptic plasticity. *Nature Communications*, *11*(1), 1–16. https://doi.org/10.1038/s41467-020-18494-6

Van Vactor, D., Sink, H., Fambrough, D., Tsoo, R., & Goodman, C. S. (1993). Genes that control neuromuscular specificity in Drosophila. *Cell*, *73*(6), 1137–1153. https://doi.org/10.1016/0092-8674(93)90643-5

Velten, L., Haas, S. F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B. P., Hirche, C., Lutz, C., Buss, E. C., Nowak, D., Boch, T., Hofmann, W. K., Ho, A. D., Huber, W., Trumpp, A., Essers, M. A. G., & Steinmetz, L. M. (2017). Human haematopoietic stem cell lineage commitment is a continuous process. *Nature Cell Biology*, *19*(4), 271–281. https://doi.org/10.1038/ncb3493

von Ohlen, T., & Doe, C. Q. (2000). Convergence of dorsal, dpp, and egfr signaling pathways subdivides the drosophila neuroectoderm into three dorsal-ventral columns. *Developmental Biology*, *224*(2), 362–372. https://doi.org/10.1006/dbio.2000.9789

Watanabe, H., Fujisawa, T., & Holstein, T. W. (2009). Cnidarians and the evolutionary origin of the nervous system. *Development, Growth & Differentiation*, *51*(3), 167–183. https://doi.org/10.1111/j.1440-169X.2009.01103.x

Weiss, J. B., Von Ohlen, T., Mellerick, D. M., Dressler, G., Doe, C. Q., & Scott, M. P.

(1998). Dorsoventral patterning in the Drosophila central nervous system: the intermediate neuroblasts defective homeobox gene specifies intermediate column identity. *Genes & Development*, *12*(22), 3591–3602. https://doi.org/10.1101/gad.12.22.3591

Weitkunat, M., Kaya-Çopur, A., Grill, S. W., & Schnorrer, F. (2014). Tension and Force-Resistant Attachment Are Essential for Myofibrillogenesis in Drosophila Flight Muscle. *Current Biology*, *24*(7), 705–716. https://doi.org/10.1016/j.cub.2014.02.032

Wheeler, W. M. (1893). *A contribution to insect embryology*. Ginn.

Winberg, M. L., Mitchell, K. J., & Goodman, C. S. (1998). Genetic analysis of the mechanisms controlling target selection: complementary and combinatorial functions of netrins, semaphorins, and IgCAMs. *Cell*, *93*(4), 581–591. https://doi.org/10.1016/s0092-8674(00)81187-3

Xu, S., Xiao, Q., Cosmanescu, F., Sergeeva, A. P., Yoo, J., Lin, Y., Katsamba, P. S., Ahlsen, G., Kaufman, J., Linaval, N. T., Lee, P. T., Bellen, H. J., Shapiro, L., Honig, B., Tan, L., & Zipursky, S. L. (2018). Interactions between the Ig-Superfamily Proteins DIP-α and Dpr6/10 Regulate Assembly of Neural Circuits. *Neuron*, *100*(6), 1369-1384.e6. https://doi.org/10.1016/j.neuron.2018.11.001

Yogev, S., & Shen, K. (2014). Cellular and molecular mechanisms of synaptic specificity. *Annual Review of Cell and Developmental Biology*, *30*(August), 417–437. https://doi.org/10.1146/annurev-cellbio-100913-012953

Zeisel, A., Hochgerner, H., Lönnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Häring, M., Braun, E., Borm, L. E., La Manno, G., Codeluppi, S., Furlan, A., Lee, K., Skene, N., Harris, K. D., Hjerling-Leffler, J., Arenas, E., Ernfors, P., Marklund, U., & Linnarsson, S. (2018). Molecular Architecture of the Mouse Nervous System. *Cell*, *174*(4), 999-1014.e22. https://doi.org/10.1016/j.cell.2018.06.021

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., Winblad, N., Choudhury, S. R., Abudayyeh, O. O., Gootenberg, J. S., Wu, W. Y., Scott, D. A., Severinov, K., van der Oost, J., & Zhang, F. (2017). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nature Biotechnology*, *35*(1), 31–34. https://doi.org/10.1038/nbt.3737

Zhao, X., Yip, P. M., & Siu, C. H. (1998). Identification of a homophilic binding site in

immunoglobulin-like domain 2 of the cell adhesion molecule L1. *Journal of Neurochemistry*, *71*(3), 960–971. https://doi.org/10.1046/j.1471-4159.1998.71030960.x

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*, *8*(1), 14049. https://doi.org/10.1038/ncomms14049

Zhu, L., Lei, J., Klei, L., Devlin, B., & Roeder, K. (2019). Semisoft clustering of single-cell data. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(2), 466–471. https://doi.org/10.1073/pnas.1817715116

Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., & Enard, W. (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. *Molecular Cell*, *65*(4), 631-643.e4. https://doi.org/10.1016/j.molcel.2017.01.023