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

M. Sc. David Puga

born in Cajamarca, Peru

Oral examination: September 22, 2020

Molecular characterisation of the commissural neurons

in the ventral nerve cord of the annelid *Platynereis dumerilii*

Referees: Dr. Anne Ephrussi

Juniorprofessor Dr. Steffen Lemke

“The description is not the described;
I can describe the mountain,
but the description is not the mountain,
and if you are caught up in the description,
as most people are,
then you will never see the mountain”

Jiddu Krishnamurti

Acknowledgements

I am very thankful to Detlev Arendt for giving me the opportunity to work in his group. I am particularly thankful for the balance with which he provided support, mentoring, independence, and liberty.

I am also very thankful to the other members of the Arendt lab, past and present. Many of them paved the way that I have ridden on, others have provided greatly-needed help during this ride, while others have been inspiring people to have around. I would like to single out Hernando Martinez-Vergara who supervised me when I first arrived to the lab, Emily Savage who I could always count on to offer me a helping hand, Paola Bertucci for her care and support, Jacob Musser whose curiosity inspires me, Vanessa Disela who taught me at least as much as I taught her, Ines Kübler whose determination and good heart are exemplary, Matthias Janeschik with whom I grew as an instructor, and Nikolas Papadopoulos for checking my results and providing insightful comments. Yet every member of the Arendt lab has contributed to this work in different, sometimes abstract ways, and I am thankful to all of them.

I also received the help of many more colleagues from different groups, among them Valentyna Zinchenko, Kimberly Meechan, Christian Tischer, and Constantin Pape.

I would like to thank the members of my thesis advisory committee, Anne Ephrussi, Yannick Schwab, and Steffen Lemke. Their insightful comments and suggestions throughout the years have helped me navigate this project and it is better thanks to them.

This project was carried out over several years and carries the influence of my life outside of EMBL as well. I would therefore like to acknowledge some people outside of my 'work life'.

Johannes, my husband, has loved and supported me during these years and I am very grateful to him. He also translated my Summary into German, which I am very thankful for. Finally, I will be eternally grateful to my mother for always being my number one cheerleader.

Table of Contents

List of Figures

List of Tables

List of acronyms

BNS: blastoporal nervous system

CNS: central nervous system

CoRC: core regulatory complex

DAPI: 4′,6-Diamidin-2-phenylindol

dpf: days post-fertilization

EMBL: European Molecular Biology Laboratory

FP: floor plate

Hpf: hours post-fertilization

KO: knock out

Ma: mega annum

MG: midline glia

Min: minute

PBS: phosphate-buffered saline

PNS: peripheral nervous system

ProSPr: Profiling by Single Probability Mapping

RP: roof plate

SBF-SEM: serial block face scanning electron microscopy

Sc-RNAseq: single cell RNA sequencing

SPM: signal probability map

TAE: triethanolamine

TBR: tree bisection-reconnection

TF: transcription factor

VC: virtual cell

VNC: ventral nerve cord

WMISH: whole-mount in-situ hybridisation

Summary

Commissural neurons have their cell body on one side of the body and project their axon across the midline to the other side of the body. The midline is a structure that is specific to Bilaterian animals and is formed when the blastopore closes. Whether the blastopore was closed in Urbilateria, the last common ancestor of Bilaterians, is an open question.

Since commissural neurons require the blastopore to close, finding that homologous neurons are commissural across Bilateria would suggest that the blastopore was closed in Urbilateria. On the other hand, finding that commissural neurons have acquired the ability to cross the midline convergently would suggest that the blastopore closed independently in the Protosome and Deuterostome lineages.

Transcription factors play a pivotal role in specifying neuronal cell fate. The analysis and comparison of transcription factor molecular signatures across Bilaterian animals thus represents a powerful approach to put forth and validate hypotheses about the homology of commissural neurons. One putatively-conserved commissural cell type is made up of ascending neurons that express the homeobox transcription *factor even-skipped (eve/evx*) in insects and vertebrates.

To advance the molecular characterisation and comparison of the commissural neurons in Bilateria, I have studied the commissural neurons of the ventral nerve cord of the marine annelid *Platynereis dumerilii*. Several characteristics of this animal make it particularly well suited for evolutionary studies, including its ancient mode of development and life style. It is also a model for which powerful tools have been developed to study its morphology and gene expression, such as a serial block face scanning-electron microscopy dataset of the full 6-days-post-fertilization (dpf) larva and gene expression atlases at several developmental stages.

Platynereis' commissural neurons were identified following both a biased approach searching for molecular signatures known from insects or vertebrates, and an unbiased approach taking the axonal traces of an entire segmental complement of commissural neurons as a starting point. I thus determined that in the second segment of the 6 dpf larva, fewer than 15% of the neurons are commissural. Analysis of their specific transcription factor-expression pattern allowed for the classification of seven cell types, characterised by the expression of i) *Phox2*, *Hox1*, and *Brn3*, ii) *Phox2*, *Hox1*, and *Isl*, iii) *Dbx1, Pax6, and Irx6,* iv) *Eve and Lhx1/5*, v) *Nk6*, *PitxB*, and *Maf,* vi) *Nk6*, *PitxB*, and *Tal,* and vii) *Nk6*, *PitxB*, and *Irx6.*

There are commissural neurons in the vertebrate neural tube with a similar transcription factor-expression pattern and developmental origin to the cell types i, ii and iv,

consistent with these being conserved in Bilateria. This would support the hypothesis that the midline existed in Urbilateria. More comparative research will be needed to settle this question.

Zusammenfassung

Kommissuralneurone haben ihren Zellkörper auf einer Seite des Körpers und kreuzen mit ihren Axonen die Mittellinie zur anderen Seite. Das Gewebe der Mittellinie gilt als spezifisch für Bilateria und wird im Zuge des Verschlusses des Blastoporus während der Gastrulation gebildet. Ob sich der Blastoporus bei den Urbilateria, den letzten gemeinsamen Vorfahren der Bilateria, bereits mittig verschloss ist gegenwärtig unklar; ebenso wie eine mögliche Homologie der Kommissuralneurone.

Weitere vergleichende Daten sind nötig, um diese Fragen zu klären. Hypothesen über die Homologie von Zelltypen können über die Analyse und den Vergleich der spezifischen und kombinatorischen Rolle von Transkriptionsfaktoren bei der neuronalen Spezifizierung erstellt und geprüft werden. Ein mutmaßlich konservierter Kommissuralzelltyp umfasst aufsteigende Neuronen, die den Homöobox-Transkriptionsfaktor "*even skipped (eve/evx)"* sowohl in Insekten wie auch in Wirbeltieren exprimieren.

Hier ansetzend, habe ich die Kommissuralneurone des marinen Anneliden *Platynereis dumerilii* untersucht. Mehrere Eigenschaften dieses Tieres machen es für Evolutionsstudien besonders geeignet, einschließlich seiner ursprünglichen Entwicklungs- und Lebensweise. Es ist auch ein Modelltier für das leistungsstarke Werkzeuge entwickelt wurden, um seine Morphologie und Genexpression zu untersuchen, wie beispielsweise ein Datensatz der seriellen Block-face Rasterelektronenmikroskopie für eine sechs Tage alte vollständige Larve, sowie Genexpressionatlanten mehrerer Entwicklungsstadien.

Meine Charakterisierung der kommissuralen Neuronen des Platynereis erfolgte sowohl mittels eines Kandidatengenansatzes als auch durch die unvoreingenommene Identifizierung sämtlicher Neuronen eines Segments. Meine Ergebnisse zeigen, dass im zweiten Segment der 6-dpf-Larve weniger als 15% der Neuronen kommissural sind. Die Analyse der spezifischen Expression von Transkriptionsfaktoren ermöglichte die Identifizierung von sieben Zelltypen, die durch die Expression *von i) Phox2, Hox1,* und *Brn3, ii) Phox2, Hox1* und *Isl, iii) Dbx1, Pax6* und *Irx6, iv) Eve* und *Lhx1/5, v) Nk6, PitxB* und *Maf, vi) Nk6, PitxB* und *Tal, sowie vii) Nk6, PitxB* und *Irx6* gekennzeichnet sind.

Es gibt Kommissuralneurone im Neuralrohr der Wirbeltiere mit einem ähnlichen Transkriptionsfaktor-Expressionsmuster und ähnlicher Entwicklungsherkunft wie die oben beschriebenen Zelltypen i, ii und iv. Diese Kommissuralneurone könnten daher evolutionär konserviert sein und die Hypothese stützen, dass eine geschlossene Mittellinie bereits in Urbilateria existierte. Um diese Frage abschließend zu beantworten, werden weitere Untersuchungen benötigt.

I. INTRODUCTION

1 Cell Types Are Evolutionary Units That Can Be Recognised By The Transcription Factors They Express

Cell types have traditionally been defined based on their phenotype or function. Animals and plants have also been classified according to their phenotype for millenia. But since the development of the evolutionary theory, it has been recognised that classifying living organisims according to their evolutionary history can provide insight that would otherwise remain unknown.

For example, Ctenophora and Cnidaria were classified as sister groups in a clade named Coelenterata[1] based on their radial symmetry and other morphological and ecological features. Since then we have learned that Cnidaria and Ctenophora are not as closely related as Hyman thought[1], though their position in the Metazoan phylogeny is still a matter of intense debate[2–6]. Since Ctenophora, Cnidaria, and Bilateria are the only clades that have a nervous system, the placement of these clades in the metazoan phylogeny guides our interpretation of the evolution of the nervous system (Figure I-1).

Just like the species in the biosphere can be classified according to their evolutionary history, so can the cell types in multicellular organisms. Just like species diversity can be studied from an evolutionary perspective, so can cell type diversity[7–9]. Cell types are evolutionary units, meaning that they can evolve as a unit and, at least to an extent, independently from each other[10].

The evolution of cell types is a recent field that is gaining interest among the community of evolutionary biologists, thanks to intelectual and methodological advances in the last two decades that have allowed us to recognise and reconstruct the evolutionary history of cell types[9,11–13].

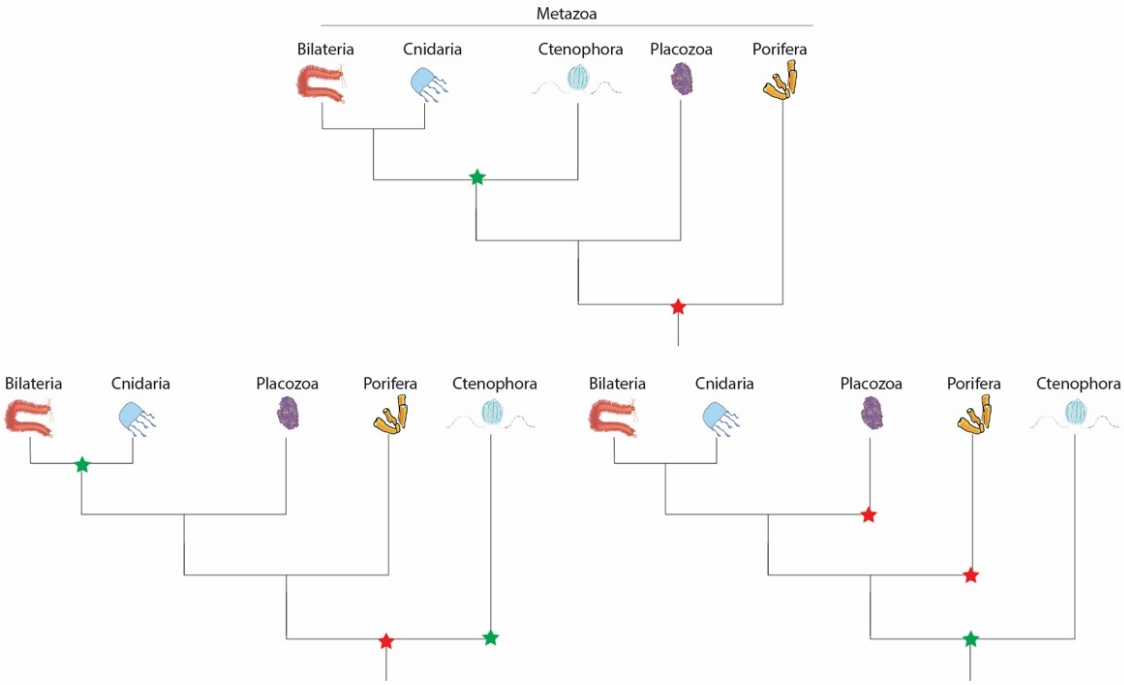


Figure I-1. Metazoan phylogeny and the origin of the nervous system.

The panel above depicts Ctenophora as sister group to Bilateria and Cnidaria. In this scenario, the most parsimonious explanation to the evolution of the nervous system is that the last common ancestor of Metazoa did not have a nervous system (red star) and that it evolved in the last common ancestor of Ctenophora, Cnidaria, and Bilateria (green star). A different phylogenetic position for Ctenophora is depicted in the panels below, as sister group to the rest of metazoan. In this scenario two equally parsimonious explanations exist. One explanation is that the last common ancestor of Metazoa had a nervous system that was independently lost in the Poriferan and Placozoan lineages. The second explanation is that the last common ancestor of Metazoa did not have a nervous system, and that it evolved independently in the Ctenophoran and Cnidarian and Bilaterian lineages.

1.1 The Core Regulatory Complex (CoRC)

These advances include the recognition that cell types can become, at least partly, evolutionary independent of other cell types through the differential use of genomic information; changes in the genome that drive the evolution of a cell type will only affect the cell types that use that genomic information. Another important advance was the insight that cell type identity is driven by Core Regulatory Complexes (CoRC) of transcription factors (TF) that control cell type-specific cellular modules, or apomeres, that are responsible for the phenotype of the cell type and make them distinct from other cell types (Figure I-2).

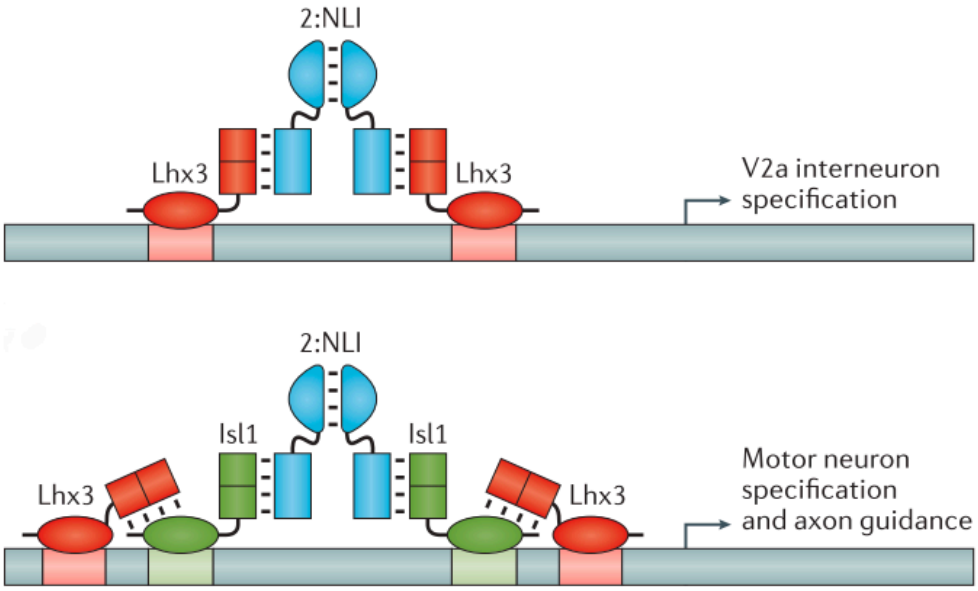


Figure I-2. V2a interneuron and motoneuron CoRCs.
A TF tetramer formed by two molecules of the LIM-homeodomain TF *Lhx3* and two of the adaptor protein *NLI* drives the specification of the V2a interneurons. The TF hexamer formed by two molecules of *Lhx3*, two of the LIM-homeodomain TF *Isl1*, and two of NLI drives the specification of motoneurons. In the second case, the hexamer has been shown to bind to an enhancer element of the homeobox TF *Hb9 (Mnx)*, whose expression triggers the differentiation of motoneurons[14]. Figure adapted from Ref [10] which was in turn adapted from Ref [15].

The emergence and accessibility of single cell-RNA sequencing (sc-RNAseq) technologies has opened the door to investigate the evolution of cell types at a greater scale than ever before, spanning whole tissues[16], whole tissues at different developmental stages[17], whole tissues in different species[18], whole organisms[19], and even whole organisms at different developmental stages[20]. But while comparing the whole transcriptomes of cells certainly helps to assess their overall similarity, this approach does not allow us to distinguish between similarity acquired through convergence or acquired by homology[7,10].

1.2 Identifying Homologous Cell Types

To address this, it is useful to think of genes as belonging to two categories: effector genes, such as enzymes and cytoskeletal proteins, which are directly responsible for the metabolism and phenotype of a cell type; and regulatory genes, or TFs and co-factors, whose effect on a cell type's phenotype is mediated through effector genes[21,22].

Effector genes are hypothesized to be less likely to carry information about homology than TFs. This is because cell types use similar effector genes for a particular function, such as contraction. All contractile cells will share a large proportion of effector genes

that are necessary for contraction. This does not, however, reflect their evolutionary relatedness.

Regulatory genes, on the other hand, seem to reflect evolutionary relatedness more faithfully[7]. This is because the differentiation of cells needs to follow an exact suite of states in order to arrive to its final state. Tampering with the regulatory program of a cell often changes its fate, and in most cases this will be deleterious to the organism and the change will not be passed on. This makes the regulatory program well conserved. The differentiation of cell types has become incorporated into developmental programs that group cell types that share a common phase of differentiation. This makes the whole system very robust and well conserved.

By studying how cell types relate to each other using TFs, one avoids the caveat that the relationships observed might reflect functional similarity rather than shared evolutionary history. This also allows us to draw the evolutionary history of cell types in phylogenetic trees, much like phylogenetic trees of species. Importantly, while a cell-type phylogenetic tree can be drawn for every multicellular organism, it is also possible to use these trees to reconstruct the cell type trees of extinct ancestors (Figure I-3).

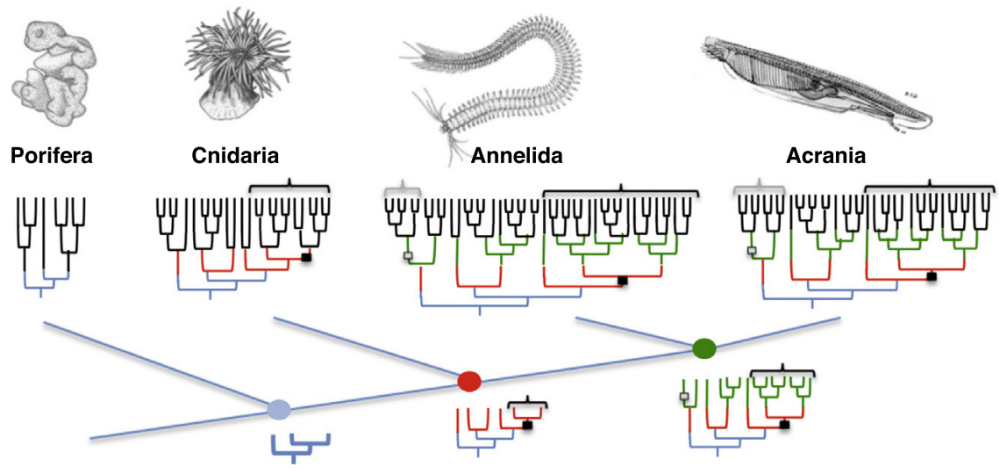


Figure I-3. Metazoan cell type trees

The last common ancestors of Metazoans, Neuralians and Bilaterians are represented by a blue, red and green dot, respectively. Under each of these dots, a tree represents the cell type complement of this ancestor, where each branch tip is a cell type. Note that the cell type complement of an ancestor is inherited by all of its descendants, and further diversified in each lineage. The black branch tips represent cell types in extant species. The relaterdness of the cell types is depicted in these trees. Taking for example the left-most branch of the blue tree, we can see that in Porifera, this ancestral cell type diversified into 3 cell types; these 3 cell types are homologous to the Cnidarian 5 cell types that occupy the same position in the tree. Figure adapted from Ref [23].

As the Ctenophore example above illustrates, adopting an evolutionary perspective to the study of cell types can provide insight that would otherwise remain hidden. It allows us, for example, to reconstruct the cell-type complement of the ancestors of different clades, as well as the emergence and disappearance of cell types, and can help us understand the mechanisms and circumstances that drive cell types to evolve or to vanish.

To summarize, cell types are evolutionary units that can be recognized by the TFs that they express. By comparing the TF-expression profile of different cell types, we can infer their phylogenetic affinities. This approach provides insight that would otherwise not be accessible.

2 Commissural Neurons Are Cell Types That Constitute A Truly Bilaterian Innovation And Could Inform Whether The Blastopore Was Closed In Urbilateria

Cnidarians, the sister-group to Bilaterians (Figure I-1), have a single gastric opening for ingesting food and disposing of waste products of digestion. Bilaterians, on the other hand, have a through gut. The through gut is formed when a developmental structure called the blastopore closes, resulting in a digestive tube with an anterior opening -the mouth- and a posterior opening -the anus.

While the closure of the blastopore in extant Bilaterians is a developmental process, it is also an evolutionary process that happened around the time that Bilaterian animals emerged, around 600 Mega annum (Ma) ago[24,25]. Ancestors of Cnidarians and Bilaterians had a single gastric opening which is the ancestral structure of the contemporary gastric opening of Cnidarians and the blastopore of Bilaterians[26,27]. After the split of Cnidarians and Bilaterians, the gastric opening in the Bilaterian lineage closed, giving rise to the digestive tube.

The digestive tube is not the only new anatomical feature that appeared upon the closure of the blastopore. The tissue immediately surrounding the blastopore of extant Bilaterians differentiates into structures that mark the midline: the floor plate (FP) in vertebrates and the neuroectodermal midline in annelids[28,29]. It can therefore be inferred that when the blastopore closed ancestrally, it formed the tissue that marks the midline.

2.1 The Blastoporal Nervous System

The neural tissue surrounding the Cnidarian gastric opening is sometimes thicker than in other parts of their body[30,31]. This neural tissue was hypothesized to be homologous to the neural tissue surrounding the blastopore of Bilaterians as early as the 19th Century[32,33]. Balfour suggested that if one were to constrict a circumoral nerve ring (like the one Cnidarians have), thereby extending it at opposite sides, one would end up with two longitudinal cords joined at the extremities (Figure I-4a-b). Recent evidence supports the homology of the neural tissue around the Cnidarian gastric opening and the Bilaterian blastopore[34,35].

This homologous neural tissue has been given the name of Blastoporal Nervous System (BNS)[31,36]. The BNS is the part of the nervous system that controls locomotion and feeding, made up of sensory neurons, interneurons, and motoneurons to coordinate these functions. In Anthozoan Cnidarians it forms the circumoral nerve ring as well as parts of the nerve net that innervate the tentacles. In Bilaterians it forms the medial-most central nervous system (CNS) of the trunk: spinal cord in vertebrates and ventral nerve cord (VNC) in Protostomes. The BNS is also thought to extend to the posterior-most portions of the brain[36], but we will focus on the BNS of the trunk.

Importantly, when the blastopore closed during the evolution of Bilateria it segregated the BNS into a left and a right side, separated by midline tissue. This set the stage for the evolution of a novel neuronal type which was able to extend its axons across midline to connect both sides of the BNS: the commissural neurons (Figure I-4b-c).

2.2 Evolutionary Origin of Commissural Neurons

Commissural neurons represent a truly Bilaterian innovation[31,37]. If the blastopore was closed in Urbilateria, the last common ancestor of Bilateria, commissural neurons could have evolved in this ancestor and homologous commissural neurons could be found in Deuterostomes and Protostomes. If the blastopore was not yet closed in Urbilateria, it would have closed independently in Protostomes and Deuterostomes. In this scenario, commissural neurons would have also evolved independently in these two lineages.

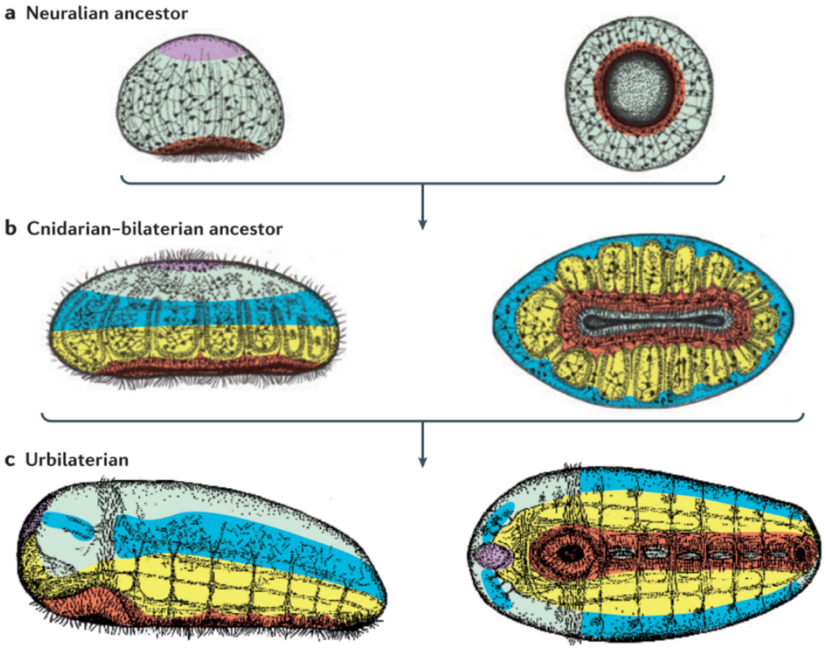


Figure I-4. Evolution of the BSN

The left panels are from a side view and the right panels, from a ventral view. The neuralian ancestor, the hypothetical last common ancestor of all organisms with a nervous system (assuming a single origin of the nervous system), had a single gastric opening surrounded by a medial neurogenic region (red). The cnidarian-bilaterian ancestor also had a single gastric opening surrounded by ectoderm that was patterned in a more complex manner: medial (red), lateral (yellow), and sensory (blue) neurogenic regions generated a greater diversity of cell types than in the neuralian ancestor. Urbilateria has, in this interpretation, a closed blastopore resulting in two gastric openings: the mouth and anus. The neurogenic regions around the blastopore are similar to the cnidarian-bilaterian ancestor. The closing of the blastopore created the midline and allowed for the evolution of commissural neurons. Figure adapted from Ref [31].

Whether the blastopore was closed in Urbilateria or closed independently in the Protostome and Deuterostome lineages is an open question. Some researchers favour the view that an ancestral gastric opening became the mouth (evo-protostomy[38–40]) or the anus (evo-deuterostomy[41,42]). These views consider that an ancestral gastric opening moved either towards the anterior or posterior of the animal to form either the mouth or the anus, respectively (Figure I-5a).

Other researchers interpret that the ancestral gastric opening, instead of moving anteriorly or posteriorly, constricted and formed a slit-like opening (evo-schistostomy) from which either the mouth (evo-protostomy), the anus (evo-deuterostomy), or both (evo-amphistomy) derived[43] (Figure I-5b).

In the first set of interpretations, where the gastric opening moved anteriorly or posteriorly, the anteroposterior axis of Bilaterians would have derived from the primary axis of a gastrula-like ancestor. In the second set of interpretations, where the gastric opening constricted into a slit, the Bilaterian anteroposterior axis would have derived from the directive axis of a gastrula-like ancestor.

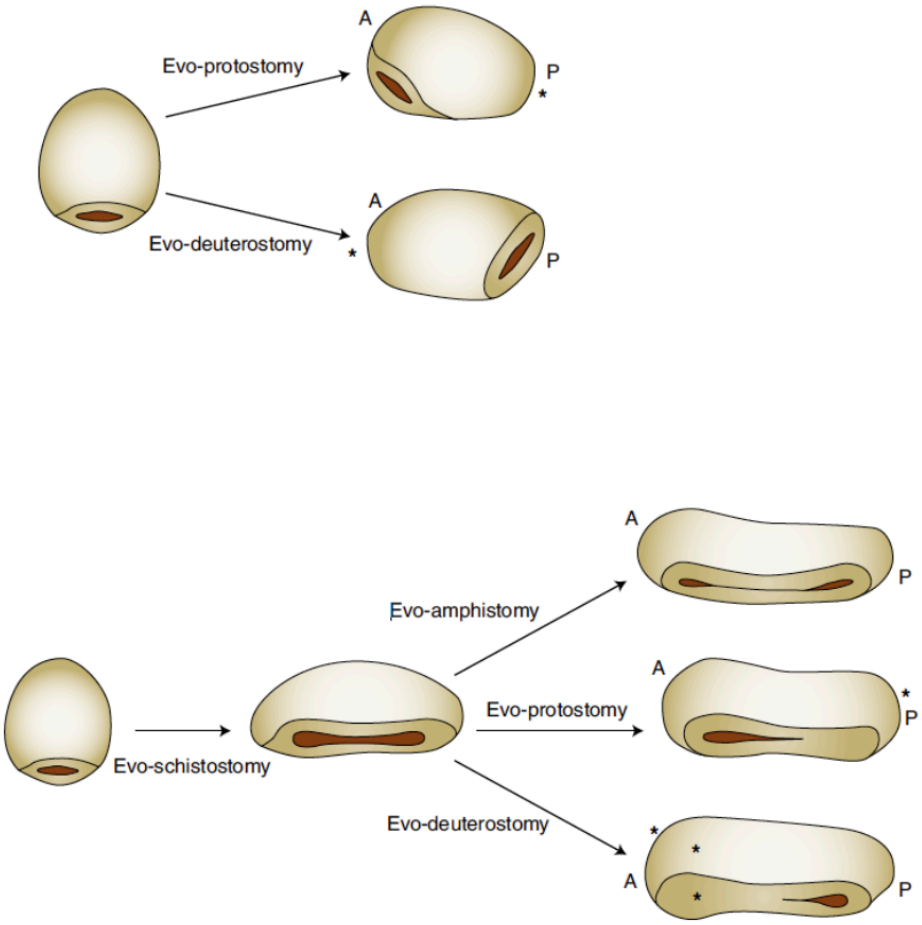


Figure I-5. The closing of the blastopore in evolution

The panels above illustrate evo-protostomy and evo-deuterostomy: the single gastric opening of an ancestor would have relocalized to the anterior (A) or posterior (P) in its descendants, giving rise secondariy (*) to the mouth or anus. The panel below illustrates evo-schistostomy: the single gastric opening of an ancestor would have first elongated in its descendants and in a second stage closed, creating both the mouth and anus, or only the mouth with the anus opening secondarily, or only the anus with the mouth opening secondarily. Figure adapted from Ref. [43].

A careful comparison of the body axes of Cnidarians and Bilaterians, supported by the study of the fate of the tissue surrounding the blastopore and other lines of evidence that escape the scope of this introduction, lead Nielsen and his collaborators [43] to favor the view that the blastopore evolved from the gastric opening by evo-amphistomy in stem Bilaterians. In this view, the blastopore had already closed in Urbilateria and homologous commissural neurons could be found in Protostomes and Deuterostomes.

To summarize, upon closure of the blastopore in evolution, the BNS became segregated into a left and a right side, with midline tissue separating both halves. Commissural neurons, a novel cell type, evolved and were able to extend their axons across the midline to connect the right and the left halves of the BNS. If the blastopore

was closed in Urbilateria, homologous commissural neurons could be found in Protostomes and Deuterostomes; otherwise, not.

3 Conservation Of The Mediolateral Patterning Of The Blastoporal Nervous System

For an animal to coordinate its feeding and locomotion, it needs to be able to sense its environment and coordinate its movements. This requires a diverse set of neurons that originate in the BNS. To create this diversity, the BNS is divided into several regions from which different populations of neurons are generated. This regionalisation is achieved by the staggered expression of TFs, whereby unique combinations of these define the different regions (Figure I-6).

Across Bilateria, the trunk CNS consists of a lateral sensory domain and a medial motor domain. The lateral domain, specified by the homeobox TF Msx, consists of sensory neurons and interneurons. The medial domain, specified by the NK homeobox TF *Nk6*, consists of the motoneurons and interneurons that make up the locomotory circuit. The conservation of these domains and of the general function of the cells that arise from these domains has led to the hypothesis that the mediolateral patterning of the BNS is ancestral[44]. In the following paragraphs I will summarize the evidence that supports this.

In vertebrates, the spinal cord has a luminal region called the ventricular zone, where the neural progenitors reside[45]. In the marine polychaete worm *Platynereis dumerilii* (see Introduction, section 6.1), the trunk neuroectoderm of the developing larva is localized ventrally. It is important to stress that the conserved patterning of the BNS specifies neural progenitors and not differentiated neurons. The comparisons between species should therefore be carried out in the neuroectoderm at pre-differentiation stages.

The vertebrate ventricular zone and the trunk neuroectoderm of the developing *Platynereis* larva are patterned by much the same TFs in the same relative position. Briefly, the medial-most region expresses the NK homeobox transcription factor *Nk2.2*, followed by the paired-domain homeobox transcription factor *Pax6*. *Nk6* spans the *Nk2.2* domain and the medial-most domain of *Pax6*. The vertebrate paired-domain homeobox transcription factors *Pax3* and *Pax7*, and their *Platynereis* ortholog *Pax3/7*, largely overlap with *Pax6* and extend further laterally. In the lateral-most region, *Msx* and the homeobox TF *Dlx* are expressed (Figure I-6).

There are also some differences. *Pax2* is expressed laterally in vertebrates and its more medial expression domain that overlaps with *Pax6* is only found in post-differentiation

stages; in the marine annelid *Platynereis dumerilii*, the *Pax2* ortholog, *Pax2/5/8*, is expressed laterally but also medially at pre-differentiation stages. Also, the vertebrate expression pattern of the bHLH TF *Sim*, the HOXL homeobox TF *Gsh* (or *Gsx*), and the NKL homeobox TF *Dbx* is not conserved in *Platynereis*, where they are only expressed after differentiation[46].

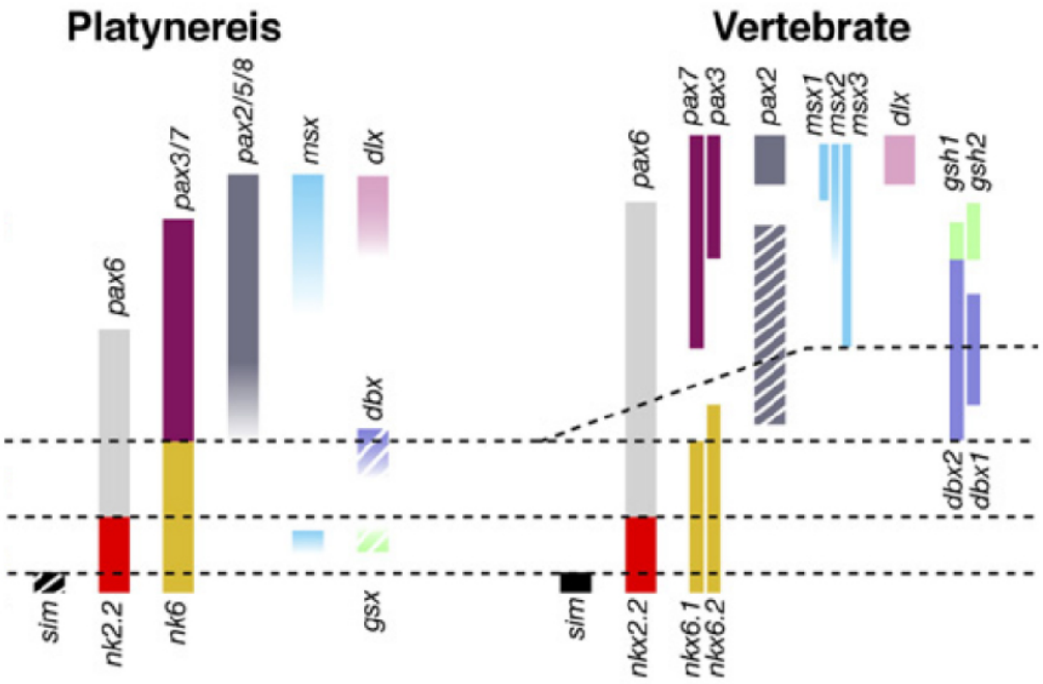


Figure I-6. Mediolateral patterning of the BNS in *Platynereis* and Vertebrates
In both panels medial is at the bottom and lateral is at the top. The dashed lines separate different progenitor domains that are characterised by the overlap of unique combinations of TFs that are, to an extent, conserved in *Platynereis* and Vertebrates (see text). After Ref.[46].

Overall, the patterning of the pre-differentiation neuroectoderm of the trunk of vertebrates and Platynereis is very similar. Subsequent studies in other species have shown a remarkable conservation of this pattern that not only extends to other Bilaterian lineages but even to Cnidaria (Figure I-7). This is why the mediolateral patterning of the BNS is hypothesized to have been present in the last common ancestor of Bilaterians and Cnidarians[47,48].

A recent paper by the lab of Andreas Hejnol put this homology hypothesis into question by providing the expression patterns of many of these genes in a remarkable amount of species and developmental stages[49]. I am critical of their interpretations of the data because they ignore whether these expression patterns occur at pre- or post-differentiation stages of the neural progenitors.

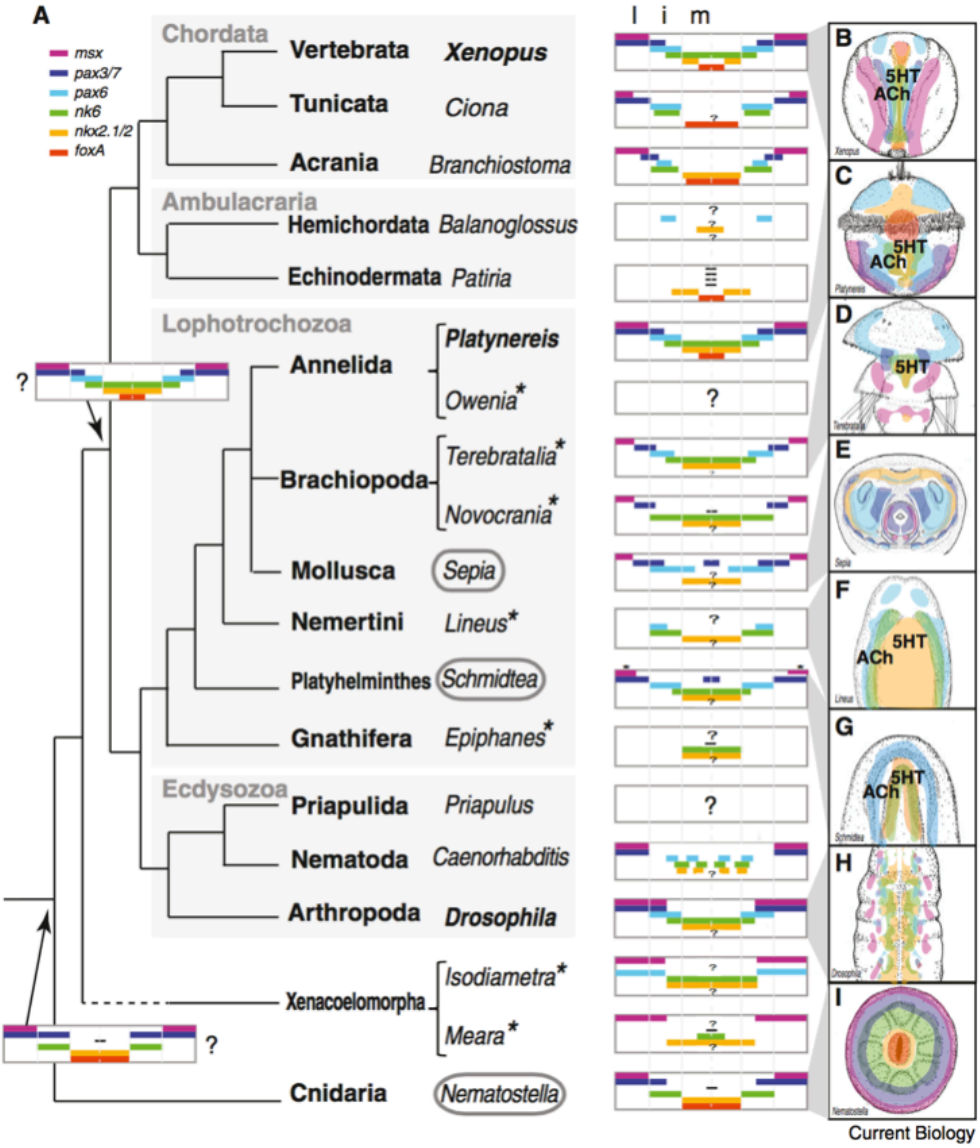


Figure I-7: Mediolateral patterning of Bilaterian and Cnidarian nervous systems

The staggered mediolateral expression of TFs that pattern the nervous system of several Bilaterians and of Nematostella are schematised in A. The hypothetical expression pattern in the last common ancestors of Bilateria and of Bilateria-Cnidaria are shown with arrows pointing at their phylogenetic position. B-I: the expression domains of these TFs in the larvae of some animals is shown as well as the position where serotonergic (5HT) and cholinergic (ACh) neurons arise. Abbreviations: l, lateral; i, intermediate; m, medial neurogenic region. Reproduced from Ref. [48].

While the regionalization of the BNS is conserved at pre-differentiation stages, we have limited knowledge about the conservation of the cell types that differentiate from these progenitors. Comparative data indicates that the serotonergic *Nk2.2*$^{+}$ and cholinergic *Hb9*+, *Pax6*$^{+}$, *Islet*$^{+}$ neurons are conserved in vertebrates and annelids[46]. Here, I will present data that the *Hb9*$^{+}$, *Pax6*$^{+}$, Islet^{+} neurons in the *Platynereis* ventral nerve cord represent motoneurons, just as in mice[50], zebrafish[51], amphioxus[52], *Ciona intestinalis*[53],

Drosophila[54], and *C. elegans*[55]. Furthermore, neurons that express *Gata123* and *Tal* are found in *Platynereis* and may relate to the vertebrate sensory Kolmer-Agduhr neurons and V2b interneurons[56].

To summarize, the neural progenitors of the trunk neuroectoderm have distinct identities that are endowed by specific combination patterns of TFs. The patterning of these neural progenitors is conserved in Cnidarians and Bilaterians, and we infer that this patterning was already present in the last common ancestor of these two clades. While the patterning of these progenitors is conserved, less is known about the cell types that these progenitors differentiate into.

4 The Commissural Neurons Of Bilaterians

Commissural neurons have been extensively studied in Bilaterians, including annelids such as *Hyalinoecia tubicola*[57] and *Nereis virnes*[58]*;* crustaceans such as *Orchestia cavimana*[59] and *Pacifastacus leniusculus*[60]; insects such as *Schistocerca gregaria*[61] and Rhynchophorus ferrugineus[62]; cephalochordates like *Branchiostoma floridae*[63]; and agnathes such as *Ichthyomyzon unicspis*[64]; and several vertebrate species. Many of these studies have focused on the morphology, electrophysiology, or neurotransmitter identity of these cells.

But identifying homologous commissural neurons in the different Bilaterian lineages is challenging. Most of the studies mentioned in the previous paragraph have no information on the TFs expressed by these cells. And while soma position and innervation pattern can be used to identify putative homologous neurons in closely-related species, this is not possible in more distantly-related species. To identify homologous cell types across species, as argued above, it is best to focus on their TF signature. In the paragraphs below I will summarize what is known about the TF signature of commissural neurons in some species where this has been investigated.

In the vertebrate spinal cord only interneurons cross the midline. In protostomes, both interneurons and motoneurons can be commissural. The specification of vertebrate commissural neurons has been best studied in the mouse and chick models, though many of these findings are also valid for all vertebrates since the cell types of the spinal cord are conserved[65,66].

4.1 Commissural Neurons in the Vertebrate Spinal Cord

The commissural neurons in the spinal cord arise from the p0, p3, pd1, pd2, pdlL, pd5, and pd6 progenitor domains[67]. Since the commissural neurons arising from the p0, p3, and pd1 domains have been studied in most detail, we will focus on them.

The p0 progenitor domain gives rise to the commissural V0 interneurons. The differentiation of the cells from the p0 domain into commissural interneurons is controlled by *Dbx1*[68]. In the presence of *Dbx1*, these cells will differentiate into commissural interneurons that coordinate left-right alternation[69], and that belong to any of three classes: V0d, V0v or V0c. *Dbx1* negatively regulates the homeobox TF engrailed, which drives the differentiation of neurons to an ipsilateral V1 fate. *Dbx1* also upregulates the homeobox TF Evx, which drives differentiation to a V0 fate[68]. V0 interneurons contact contralateral MNs and establish inhibitory synapses with them. In *Dbx1* knock-outs (KOs) the left-right alternation of limbs is affected[69].

The p3 domain gives rise to interneurons, 85% of which are commissural and coordinate left-right alternation[70]. This progenitor domain is the ventral-most, adjacent to the FP, and is defined by the expression of *Sim* [71].

The pd1 domain gives rise to the commissural dl1 neurons, a subset of which form the dorsal commissural tract of the spinal cord. They are specified by the TF Math1[72]. These neurons consist of a subpopulation of ipsilateral (dl1i) and commissural (dl1c) neurons that express the LIM-homeodomain TFs Lhx2, and Lhx2 plus Lhx9, respectively.

4.2 Commissural Neurons in Drosophila and Comparison to the Vertebrates

In *Drosophila* most of the neurons of the VNC are commissural[73,74]. In fact, all but four NBs of *Drosophila* give rise to commissural neurons, and a NB that generates commissural neurons can also give rise to ipsilateral neurons and even glia[73,75,76].

One interpretation of the *Drosophila* data is that the commissural phenotype is very easily acquired and could reflect convergent evolution rather than descent from commissural neurons in a distant ancestor. Alternatively, it could indicate that in a distant ancestor most neurons were commissural, followed by secondary loss in species with fewer commissural neuron cell types.

An interesting case relates to the commissural neurons that express *even-skipped* (shortened as *Eve* or *Evx* depending on the species). *Even-skipped* is expressed in commissural neurons of the locomotory circuit of the mouse *Mus musculus*[77], chick *Gallus gallus*[68], frog *Xenopus laevis*[78], zebrafish *Danio rerio*[79], pufferfish *Takifugu*

rubripes[80], *D. melanogaster*[81], grasshopper *Schistocerca americana*, the sow bug *Porcellio laevis*, the crayfish *Procambarus clarki*, the opossum shrimp *Mysidium columbiae*, and possibly in the branchiopod tadpole shrimp *Triops langicauditus* and the collembolan springtail *Folsomia candida*[82].

These commissural neurons correspond to the EL neurons in the arthropods and to the V0 and CoSA neurons in the vertebrates (Figure I-8). They are hypothesized to be homologous in arthropods[82,83] and even Bilaterians[81]. While other *Eve*+ neurons have been described in arthropods, only the EL are hypothesized to be homologous to the V0 of the mouse in virtue of their commissural phenotype, their projection towards the anterior of the animal, and their role in the locomotory circuit.

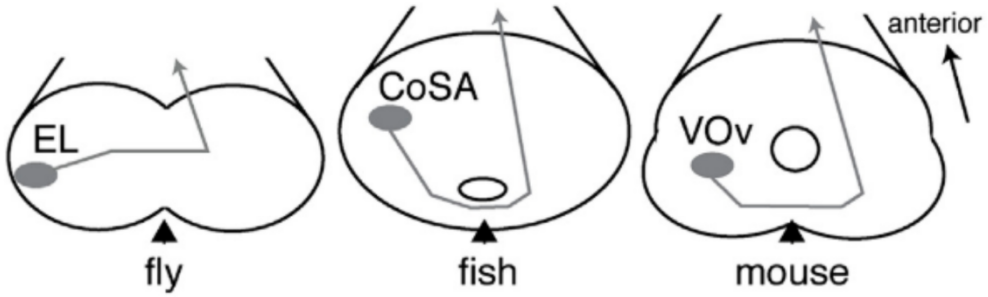


Figure I-8. Putative conserved *even-skipped*⁺ commissural neurons in Bilateria
The fly VNC and fish and mouse spinal cord are drawn in the transverse plane. Even-skipped commissural neurons are depicted with their axon crossing the midline (indicated with a black arrowhead) and then projecting anteriorly. Adapted from Ref.[81].

Further supporting this hypothesis of homology, and an observation that to the best of my knowledge has remained unnoticed, the NB3-3 from which the EL neurons are generated derives from the medial neuroectoderm that expresses the paired homeobox TFs *ey* and *toy*[84]. These are the *Drosophila* homologs of the vertebrate TF *Pax6*, which also pattern the region from which the V0 neurons derive[68].

Finally, the axonal guidance of the *Drosophila* EL neurons is controlled by *Eve*[85]. The commissural phenotype of the chick V0 neurons has been shown to be controlled by *Dbx1*[68]. Interestingly, in commissural neurons of the mouse midbrain, *Dbx1* triggers expression of the axon guidance receptor *Robo3*, and this effect is thought to be mediated by *Evx*[86]. This raises the possibility that *Dbx1*'s control of the commissural phenotype in the V0 neurons is also mediated through *Evx*[87].

One must nevertheless not forget that most of *Drosophila*'s neurons are commissural. The possibility remains that the EL and V0 neurons be homologous but their commissural phenotype be convergent. Studying the TF fingerprint of the commissural

neurons of other Protostome species will allow us to draw a clearer picture about the homology of these neurons and of their commissural phenotype.

To summarize, the TF signature of commissural neurons is necessary to assess their homology across species. This is only available for few groups, like vertebrates and *Drosophila*. While in vertebrates the commissural neurons of the spinal cord can be grouped into a manageable number of cell types, in *Drosophila* almost every neuron of the VNC is commissural. This could be interpreted as indicating that in a Bilaterian ancestor most neurons were commissural or that the commissural phenotype evolved independently in many neuronal cell types. For the current literature, the best candidate for a conserved bilaterian commissural neuron are the *even-skipped*$^{+}$ neurons of the locomotory circuit.

5 Midline Axon Crossing Is Controlled By Conserved Guidance Signals And Receptors In Bilateria

While the conservation of commissural neuron cell types in Bilaterians is debatable, there is ample conservation of the mechanisms that axons use to cross the midline. In this section I will summarize some of the literature of axonal midline crossing in Bilateria.

Developing axons respond to attractive and repulsive signals to navigate the body and innervate their final target. This happens in a step-wise manner, through intermediate targets that present the guidance cues necessary to steer the axonal growth cone through the appropriate tracts[88]. Commissural axons need to cross the midline to innervate their partners in the contralateral side. For neurons in proximity of the midline-crossing point, there might be few intermediate targets before the midline; for neurons far from the midline-crossing point, the axonal path might involve several prior steps.

5.1 Midline Axon Crossing In Vertebrates

The vertebrate dorsal commissural dI1c neurons start their axon pathfinding by responding to repulsive signals from the roof plate (RP). This leads them to migrate ventrally along the pial surface. They then turn medially thanks to attractive cues from the FP and the ventricular zone, such as *Netrin* [89,90].

Netrins were among the first guidance molecules identified as playing a role in attraction of commissural neurons towards the midline[91]. *DCC* has been identified as a receptor for *Netrin-1*, and their interactions plays an important role for the commissural

projections of the *DCC*$^{+}$ neurons[92]. *DCC* mutants have fewer commissural axons that belong to different cell types, including dI1, dI5, V0, and V3[93].

DSCAM is another netrin receptor which might act in parallel to *DCC* during midline attraction[94,95]. *Netrin-1* mutant mice have fewer commissural axons in the spinal cord[96], though it mainly affects the dorsal inhibitory commissural neurons which seem to be present but to have adopted a different axonal projection[97]. The excitatory commissural neurons of the ventral spinal cord remain largely unaffected[98].

While the FP produces *Netrin-1*, it was recently shown that it is *Netrin-1* produced in the ventricular zone and not in the FP which is necessary for commissural neurons to cross the midline[99,100]. Furthermore, these studies suggest that *Netrin-1* is important for the ventral migration of dorsal commissural neurons and not for their attraction towards the midline. This could explain why only the dorsal commissural neurons are affected in the Netrin-1 KO.

Other guidance cues that direct commissural axons towards the midline are the Hedgehog morphogen *Shh* and the growth factor VEGF. Shh, produced by the FP cells, is a chemoattractant for commissural neurons via its receptor *BOC*[101,102]. *VEGF* secreted by the FP also attracts commissural neurons through the *Flk1* receptor [103].

After crossing the midline, commissural axons continue their growth towards their final targets. This is not triggered by a change in expression of guidance signals in the FP, but rather in a change in sensitivity of the axonal growth cone to these molecules. During this phase of neurogenesis, the FP constantly produces both attractant and repellent cues. Pre-crossing commissural axons must therefore be insensitive to the repellent cues from the FP and sensitive the attractive ones; after crossing the midline, the axons must loose sensitivity to the attractant cues and gain sensitivity to the repellent ones (for a recent review refer to Ref.[104]).

The switch from attraction to repulsion to *Shh* appears to be mediated by the *14-3-3* adaptor proteins, which are expressed in post-crossing commissural neurons only[105]. It is presently not clear how *14-3-3* expression is regulated.

Another guidance molecule secreted by the FP that repels post-crossing commissural axons is *Slit*[106]. This effect is mediated by the *Robo* family of receptors[107,108]. *Robo1* and *Robo2* are *Slit* receptors that mediate repulsion[109].

In vertebrates, *Robo3* has two isoforms, *Robo3.1* and *Robo3.2*. In pre-crossing commissural axons the *Robo3.1* isoform is present, whereas in post-crossing axons the *Robo3.2* isoform is present. This has led to a model that suggests that in pre-crossing axons, *Robo3.1* antagonizes the repulsive effect of *Slit* on *Robo1* and *Robo2*, and that

in post-crossing axons *Robo3.2* collaborates with *Robo1* and *Robo2* to mediate *Slit* repulsion[107].

Sema3B, secreted by the FP and ventricular zone, is another guidance molecule that repels post-crossing commissural neurons[106]. The repellent effect is mediated by the receptor *plexin-A1*. In pre-crossing axons, *plexin-A1* is cleaved by *calpain1* and the repulsive effect of *Sema3B* is thus silenced[110]. Upon reaching the midline, *gdnf* produced by the FP suppresses the cleavage of *plexin-A1* and the commissural axons become sensitive to the repulsive effect of *Sema3B*[111].

Many other molecules play roles in the axonal guidance of the vertebrate commissural axons (Figure I-9) and I would refer the reader to the following reviews:Ref. 104,112–115.

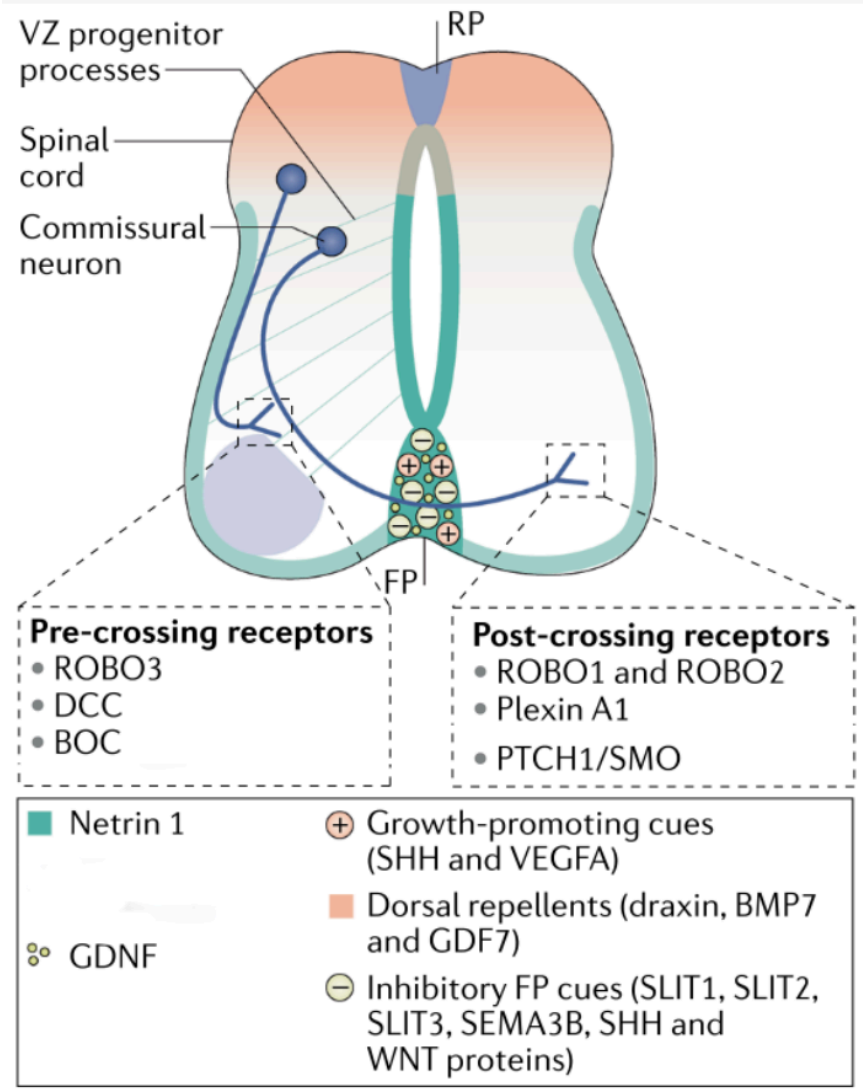


Figure I-9. Overview of guidance cues from commissural neurons in the vertebrate spinal cord

This transversal view of the spinal cord summarizes the guidance cues that commissural neurons use to cross the midline. Adapted from Ref[115].

5.2 Midline Axon Crossing In Protostomes And Comparison To Vertebrates

The mechanisms for axon guidance across the midline in protostomes have been best elucidated in *Drosophila*. Like in vertebrates, commissural axons are guided in a stepwise manner through intermediate posts until they innervate their final target.

There are fewer signalling pathways that control axon guidance across the midline of *Drosophila* than in vertebrates, and these seem to be conserved across Bilateria [113]. While in vertebrates many commissural neurons must start extending their axon in the dorsoventral axis, this is not the case in *Drosophila* since they are typically born close to the midline and aligned with their crossing point [116,117]. They start their growth by extending medially, attracted by signals produced by the midline glia (MG) [118]. Midline cells, from which many of the guidance molecules necessary for midline crossing are produced, have been postulated to be homologous in insects and crustaceans [59].

A role for netrin in guiding commissural neurons appears to be largely conserved. *C. elegans* has a single netrin ortholog, Unc-6, which is expressed in the ventral midline. The Unc-40 netrin receptor is orthologous to both DCC and neogenin, and it directs axon pathfinding towards the midline and mediates growth cone migration along the dorsoventral axis [119–121].

Netrin is also expressed in the midline of the insects *Aedes aeygypti* and *Tribolium castaneum*, the branchiopods *Artemia franciscana* and *Triops longicauditus*, the malacostracan *Porcellio laevis* and *Parhyale hawaiensis* [122], and the annellid *Platynereis dumerilii* [46]. In the planarian *Schmidtea mediterranea* netrin is not expressed in the midline but rather in the VNC [123]. Given our current understanding of Protostome phylogeny (Figure I-10 and Ref. [124]), and the fact that netrin is expressed in the midline of vertebrates, it is more parsimonious to interpret the planarian's case as derived and that netrin was indeed ancestrally expressed in the midline.

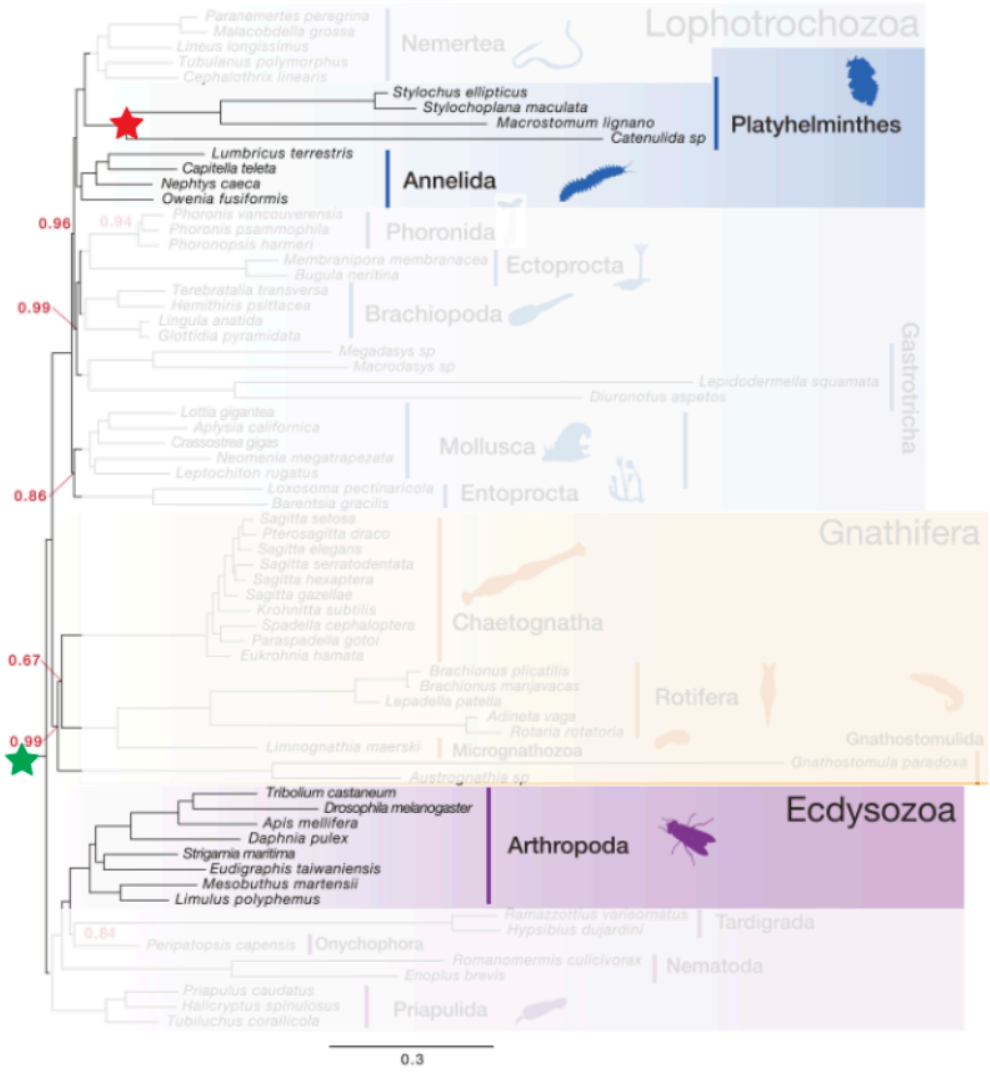


Figure I-10. Phylogenetic tree of Prostostomia and netrin expression in the midline
This tree illustrates our current understanding of the phylogeny of Protostomes. Our hypotheses about the evolution of characters have to be congruent with the species tree. Given that netrin is expressed in the midline of most Protostomes and Deuterostomes, this must have also been the case in stem protostomes (green star). The finding that netrin is not expressed in the midline of a platyhelminth is most parsimoniously explained by a secondary loss (red star). Adapted from Ref. [124].

Drosophila has two netrin paralogs expressed in the ventral midline [125]. Both appear to play a role in commissural axon guidance through short-range attraction, meaning that instead of acting as a long-range attractant towards the midline, it allows commissural neurons to cross the midline once they have reached it [126]. Their functions might be complementary since only upon deletion of both paralogs is there a significant commissural phenotype [127].

Frazzled is the *Drosophila* gene orthologous to both DCC and neogenin, and it is too a netrin receptor. It is involved in many processes taking place in the nervous system,

including motoneuron axon guidance, axon midline choice point recognition, and interneuron axon guidance [128–130].

DSCAM in *Drosophila* is remarkable since it can be alternatively spliced into tens of thousands of isophorms that are used for homophilic repulsion of axons and dendrites (reviewed in Ref. [131]). DSCAM seems to play a role in midline crossing acting as a receptor of netrin but also of other ligands [132].

Slit is the main midline repllent in *Drosophila* as it is in vertebrates. Slit is also expressed in the midline of other insects like *Anopheles gambiae* and *Apis mellifera* [133], the branchiopod *Daphnia magna* [134], the planarian *Schmidtea mediterranea* [135], and the annelid *Platynereis dumerilii* [46] just to mention some.

In *Drosophila* its repulsion in the midline is mediated through Robo receptors too [109,136,137]. It has three Robo paralogs, called Robo1, Robo2, and Robo3, but these are not 1-1 orthologs of the vertebrate Robo1, Robo2, and Robo3 since the ancestor of vertebrates and *Drosophila* had a single Robo gene that duplicated into 3 genes independently in both lineages [138].

Robo1 is of particular importance and was identified early-on in a genetic screen that showed that in the mutant, all the axons grow towards the midline [139]. Almost all neurons of the developing VNC express Robo1 [136], but the repulsion to midline Slit is silenced in commissural neurons via expression of the transmembrane protein *commissureless* (comm) [140,141]. Even though pre-crossing commissural neurons express Robo1, the protein is absent from their axonal membrane. Comm was shown to recruit the Robo1 protein into the late endosomal-lysosomal pathway, keeping it from reaching the membrane and acting as a Slit receptor (Figure I-11) [141,142]. Interestingly, the expression of comm is regulated by Frazzled independently of Netrin [143].

After crossing the midline comm expression diminishes and Robo1 protein is becomes localized on the axonal membrane; the commissural axons are thereby sensitive to slit repulsion from the midline [136,141]. Slit seems to account for all the repulsive signal from the midline in *Drosophila*, contrarily to vertebrates where other repulsive signals are in place (see above).

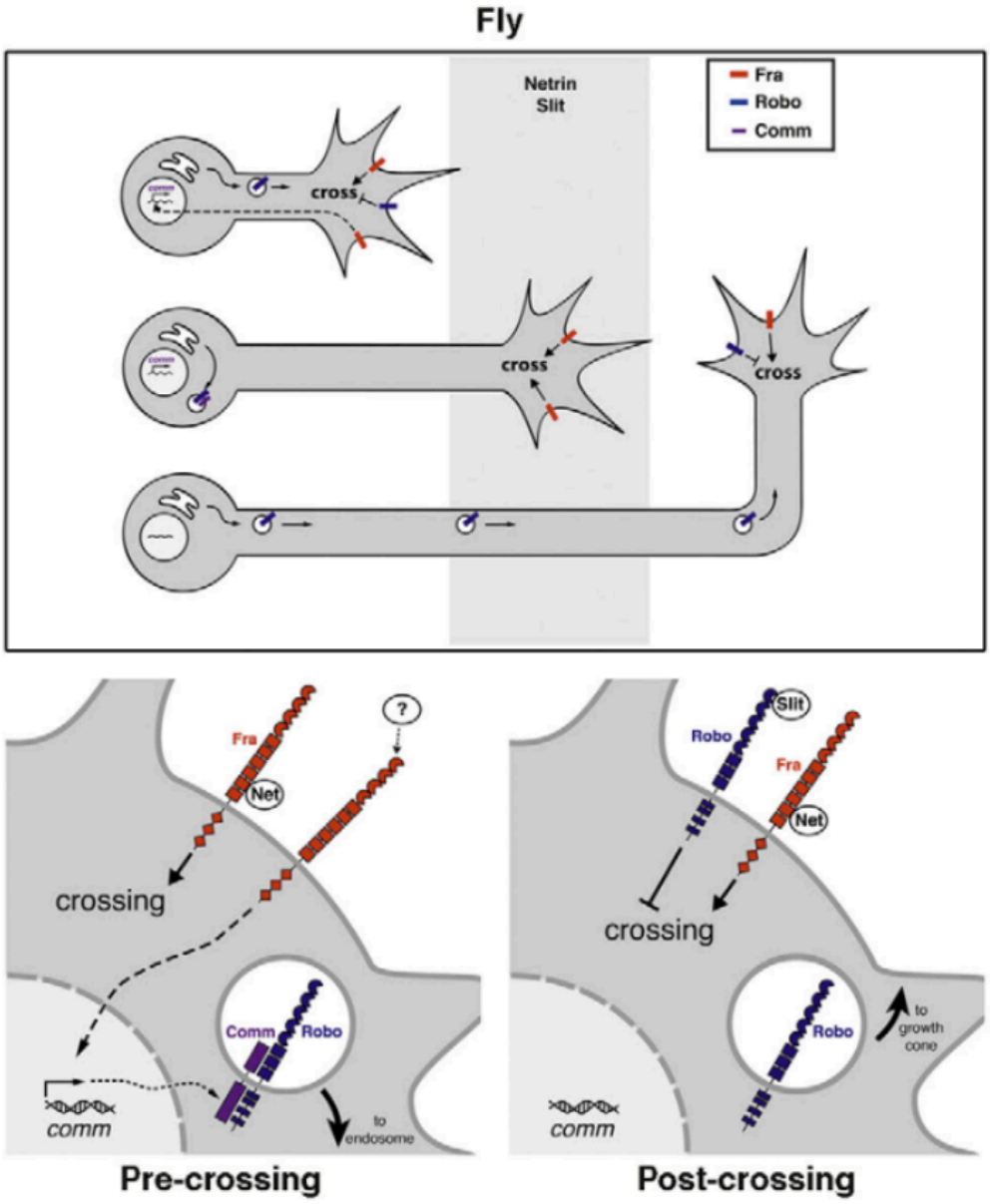


Figure I-11. Midline crossing in *Drosophila* through the interaction of Fra, Robo, and Comm

The top depicts a commissural axon at three different stages of crossing the midline. In the pre-crossing commissural axon, the intracellular domain of Fra is cleaved when it binds to Netrin. The intracellular domain induces the expression of Comm. Comm binds Robo and directs it to the endosome for proteolysis. In post-crossing axons, comm is no longer expressed and Robo becomes localized at the growth cone's membrane. After Ref. [113].

The finding that Cnidaria and many other invertebrates have a single Robo gene indicates that Urbilateria had a single Robo [138]. While several paralogs have been found in vertebrates and insects with diverging functions, phylogenetic analysis of the sequences of these paralogs indicate that the vertebrate Robo1 and Robo2 are closer to the ancestral form and function. This points to an ancestral role of Robo in mediating Slit repulsion [138].

To summarize, the most parsimonious interpretation of these data suggest Urbilateria had a slit and netrin-expressing midline and a single ancestral Robo receptor that mediated slit repulsion. I would also interpret the current data as indicating that an ancient Neogenin/DCC receptor could have played a role in netrin-mediated midline crossing, with a possible role for DSCAM too. Nevertheless, these interpretations assume that Urbilateria already had a midline formed by the closure of the blastopore. It is also possible that the midline did not yet exist and that tissue that would form the blastopore margin expressed slit and netrin. In this scenario, axons could have been guided by these signals through the expression of the receptors Robo, Neo/DCC, and DSCAM, without this guidance involving crossing the midline (which in this scenario did not yet exist).

6 Studying The Commissural Neurons Of *Platynereis dumerilii* Could Provide Insight Into The Evolution Of This Cell Type

In the absence of direct evidence of evolution, in the form of fossils, for example, one has to use a comparative approach to make inferences about evolution. In this approach, extant species are compared and their similarities and differences are assessed to draw inferences about their evolution. But for this approach to be meaningful, many species in key locations of the phylogenetic tree have to be studied.

Much of the research in biology of the past several decades has focus on a small number of model species, such as *Drosophila*, *C. elegans*, or the mouse. These species became so widely used for varied reasons, but their suitability for evolutionary studies is not among them. In fact, research has shown that these species are quite derived, meaning that they have evolved at a fast rate and have diverged for their ancestors more than other species. This is often the case when species colonize new environments, such as the (independent) colonization of land by the ancestors of mice, and of *Drosophila* and *C. elegans*. *C. elegans* in particular is very derived, as its ancestors went through major simplifications of their morphology leading to the stereotypic cell lineage and reduced number of cells that we observe in the extant nematode species. Furthermore, *Drosophila*, *C. elegans* and many other model species belong to the Ecdysozoan clade of Protostomia (Figure I-10), whereas few model species belong to the Lophotrochozoan and Gnathifera clades of Protostomia. Only studying Ecdysozoans can help us learn much about this clade, but if we want to study larger-encompassing clades, such as Protostomia or Bilateria, it is important to also include members of these understudied clades in our comparisons, such as Lophotrochozoa.

6.1 *PLATYNEREIS* IS A SLOW-EVOLVING LOPHOTROCHOZOAN MODEL SPECIES

Platynereis dumerilii is a model species that is well suited for evolutionary studies. The protein sequences of *Platynereis* have diverged less from those of vertebrates than other Protostomes[144], and the orthologues of many vertebrate genes are found in its genome[144–146], suggesting that it is a slow-evolving species. Moreover, *P. dumerilii* is thought to have conserved many ancestral features of Urbilateria, such as cell-types, development, and ecology[147–149].

P. dumerilii is well suited for evolutionary studies because it belongs to the Lophotrochozoa clade. These features of *Platynereis* allow for relevant comparisons to be made with Deuterostomes and thus help to draw inferences about Urbilateria.

It also has many advantages as a model species. It can be cultivated in the laboratory, providing mature worms for mating every day of the year. Fertilization is external and thousands of zygotes can be obtained from a single pair of worms. The larvae are transparent until >6 days post fertilization (dpf). Tested protocols for most of the common molecular biology techniques, such as immunostaining and whole-mount in-situ hybridization (WMISH) are also available. CRISPR-Cas9 has been successfully used to generate mutants. In addition, The genome has been sequenced and both long-read and short-read transcripts have been mapped. The gene models are currently being tested.

6.1.1 PROSPR GENE-EXPRESSION ATLASES

The ProSPr atlases[56] are gene-expression atlases of the entire *Platynereis* larva, and are currently available for the following developmental stages: 16, 20, 24, 28, 36, 48, 56, 72 hpf, and 6 dpf.

Briefly, for each developmental stage, gene expression patterns that are obtained through WMISH are overlaid and aligned onto a common spatial framework. This spatial framework is a high-resolution reference generated from DAPI-stained nuclei of many individual larvae (in the case of the 6 dpf,153 individuals were used for the reference).

For each gene, the expression pattern from dozens of larvae (usually around 30) are aligned to this reference. By overlapping the expression patterns of every larva, a 3D expression probability map -or ‘signal probability map’ (SPM)- of this gene is generated.

The SPM can be subsequently used to calculate the average location of gene expression. The SPM’s region of highest probability was found to map to the reference

at cellular resolution[56], something that is only possible because the development of Platynereis is highly stereotypical until 6 dpf. Every gene that is incorporated is aligned to the same reference (for each developmental stage), which makes it possible to study the coexpression of genes at cellular resolution.

6.1.2 *PLATYNEREIS* SBF-SEM DATASET

The SBF-SEM *Platynereis* dataset[150] is a ~2.5 TB image stack from an entire 6 dpf *Platynereis* larva (300 x 150 x 150 μm^3) acquired with a pixel size of 10 x 10 nm^2. It consists of 11731 slices, 25 nm thick; no slice was lost.

Immobilized 6 dpf *Platynereis* larvae were immobilized and fixed following [151], starting directly with the cacodylate buffer incubation of step 1.To alleviate the undesired effects of the electron beam on the sample, the samples were removed from the Epon resin before it was cured and transferred into an epoxy glue with silver particles, also known as EE embedding [152].

The images were acquired with a ZEISS Merlin SEM at 1.8 keV landing energy, 270pA beam current, and 0.8 µs dwell time. Please refer to Ref. [150] for a detailed description of the methods followed to process these images.

I used this dataset to trace neurons using the software PyKnossos. It is also part of the PlatyBrowser.

6.1.3 PLATYBROWSER

The PlatyBrowser[150] is a multimodal atlas that combines the 6 dpf ProSPr gene-expression atlas with the *Platynereis* SBF-SEM dataset. We used it to extract the gene-expression profile of the commissural neurons of the second segment. The paper describing this astonishing resource has been submitted for peer review and available on bioRxiv[150]. For a detailed explanation of the resource please refer to this publication. I will briefly explain the aspects of this resource that are most relevant to the this thesis.

The nuclei of the SBF-SEM dataset were segmented using an extension of the mutex watershed algorithm[153]. This algorithm segments through agglomerative clustering, similarly to seeded watershed, but incorporates not only attractive but also repulsive cues. It requires no seeds nor stopping criteria.

The segmented nuclei were then registered to the ProSPr DAPI reference using the software package elastix. 43 reference landmarks were found to guide the alignment. First, they computed rigid transformations of the ProSPr DAPI reference, which allows

rotation, translation, and uniform scaling. This registration was already very good and resulted in a median discrepancy of 10.09 µm between the landmarks (4.7 µm is the diameter of a cell[56]). To refine the registration, the ProSPr DAPI reference underwent local deformations using a sequence BSpine transformations with 5.5 µm grid spacing[154]. The median discrepancy between landmarks was thus reduced to 2.99 um, less than a cell diameter.

This results in a multimodal atlas that incorporates the gene-expression data from ProSPr with the SBF-SEM dataset. For each segmented nucleus in the SBF-SEM we can extract how much of its volume overlaps with the genes in the 6 dpf ProSPr atlas.

The gene expression data is given as a proportion of the volume of the segmented nucleus that overlaps with the expression domain of a gene. (overlap value).

> Overlap value: fraction of a segmented nucleus that is covered by the expression domain of a gene in the PlatyBrowser[150]. Every overlap value relates to a single segmented nucleus and a single gene. There are currently 201 genes in the PlatyBrowser, meaning that every segmented nucleus has 201 overlap values. 1 = total overlap, 0 = no overlap.
>
> Definition 1

Finally, the segmented nuclei were also used to introduce top-down constraints to cell boundaries and to automatically segment larger tissues, among them the VNC. This tissue information is also incorporated into the PlatyBrowser, and one can subset the segmented cells according to their tissue.

6.1.4 Commissural Neurons in *Platynereis dumerilii*

Some research into the commissural neurons of *Platynereis dumerilii* has already been published. A description of the development of the pioneering commissural axons of each commissure of the VNC can be found in [155]. Briefly, the VNC consists of three large commissures where many axons cross the midline. The first commissural axon crosses the midline at approximately 38 hpf and pioneers the first commissure. The pioneering axon of the second commissure crosses the midline at approximately 44 hpf. And the pioneering axon of the third commissure, at approximately 48 hpf. There are, nevertheless, many commissural axons that cross the midline outside of these three commissures.

At 50 hpf, the longitudinal nerves of the VNC are furthest apart, and the commissural axons that are perpendicular to them are easily distinguishable. After 50hpf, the

longitudinal nerves will grow closer together, and the individual commissural axons will become increasingly difficult to distinguish. For this reason, many of the experiments where commissural axons are quantified are carried out at 50 hpf [156,157].

Hernando Martinez-Vergara[156] found that *Dbx1* and Neo/DCC were coexpressed in cells of the VNC of the developing *P. dumerilii* larva. He also found, in collaboration with Ryan Prestil, that knocking out *Dbx1* resulted in a 20% reduction of the commissural density in the VNC. Ines Kübler[157] found that knocking out Neo/DCC also decreased the commissural density in the VNC, though not significantly when compared to control larvae. But as I will show in this thesis, a reanalysis of the data shows that the reduction of commissural axons in the Neo/DCC mutant is significantly higher than in control larvae.

7 Aim of this Thesis

My aim in this thesis is to describe the commissural neuron cell types of *Platynereis dumerilii* in such a way as to allow for the search and comparison of putative homologous cell types in other species. Ultimately, I wish to provide some insight into the evolution of commissural neurons through these comparisons.

For this, I will build on the resources that have been developed by my colleagues. To analyse gene expression patterns, I will use the ProSPr gene-expression atlases [56] at different developmental stages to analyse gene expression patterns.

In a candidate gene approach, I set out to identify neurons that share the conserved molecular fingerprint for the vertebrate p0 and V0 commissural cell types (Supplementary figure 1) and investigate their possible commissural phenotype in *Platynereis*.

To gain first insight into the full complement of commissural neurons of *Platynereis dumerilii,* I did an unbiased search of the commissural neurons of the second segment. I did this on the 6 dpf larva using an SBF-SEM dataset [150] and the tracing software PyKnossos [152] (see Materials and Methods, section IV.13). I traced all of the commissural axons of the second segment in this dataset. My aim was the full coverage of the second segment because its gene-expression profile is most comparable to the vertebrate spinal cord at this stage [46,56].

To interrogate the relatedness of the commissural neurons that I identified, I used their transcription factor-expression profile to cluster them using maximum parsimony. Finally, to gain insight into the evolution of commissural neurons, we searched for conserved commissural cell types in vertebrates and *Platynereis*.

II. RESULTS

1 Populating The ProSPr Atlases For This Study

Intending to make guided comparisons to the vertebrate spinal cord, we complemented the ProSPr atlases with conserved genes of interest involved in the specification and differentiation of commissural neurons in the model organisms.

Most of these genes had already been added to at least one of the atlases. For these genes we already had an RNA probe for the WMISH or a plasmid from which this probe could be transcribed. In the cases where the gene had not been added to any of the ProSPr atlases, we had to design the RNA probe for the WMISH. Also, following a suggestion by Dr. Rob Meijers, I designed the probes for *Ephrin* and *Glypican1/2/4/6* (Materials and Methods, section IV.6).

The genes that I helped add to the ProSPr atlases are (Figure II-1):

- *Beta3* (*BHLHE23, bHLH5*): *bHLH* transcription factor expressed in the postmitotic dI6, V0d, V1 and V2 neurons of the spinal cord[158].
- *Chx10* (*Vsx2*): this transcription factor of the visual system homeobox family is important for the differentiation of the V2 neurons of the spinal cord[159]
- *Dbx1*: this NKL homeobox transcription factor is important for the differentiation of the V0 commissural neurons of the spinal cord[68].
- *DSCAM:* this gene is a member of the immunoglobulin superfamily of cell adhesion molecules. It plays a role in axon guidance and I have discussed its relevance to this study in the section I.5 of the Introduction.
- *En:* this homeobox gene plays an important role in the differentiation of the V1 neurons of the spinal cord[160].
- *Eph:* Ephrins are guidance queues that bind to Ephrin receptors. It is involved in the guidance of commissural neurons in the vertebrate spinal cord[161]. See Figure IV-1 for a gene tree.
- *Eve*: *even-skipped* is a homeobox transcription factor that plays an important role in the differentiation of the V0 commissural neurons of the spinal cord[77]. It relevance for this study has also been discussed in the section I.4 of the Introduction.
- *FoxD:* this forkhead box transcription factor is expressed in the V1 and dI2 neurons of the spinal cord[158].
- *FoxP*: this forkhead box transcription factor is expressed in the V1 interneurons and in the motoneurons of the spinal cord[158].
- *Gbx1:* this homeobox transcription factor is expressed in the dILa neurons of the spinal cord[158].

- *Gly1/2/4/6:* Glypicans are membrane-associated proteins that make up the core of heparan sulfate proteoglycans. Glypicans have been found to be implicated in the axonal guidance of commissural neurons[162,163]. See Figure IV-2 for a gene tree.
- *Hb9 (Mnx):* this homeobox transcription factor plays a crucial role in the development of motoneurons in vertebrates[164] and *Drosophila*[54]
- *Netrin:* this guidance molecule plays a conserved role in the axonal projection of commissural neurons in vertebrates and *Drosophila*. Please refer to section I.5 of the Introduction for a wider discussion about the relevance of Netrin to this study.
- *Tal (Scl):* this *bHLH* transcription factor is expressed in the V2b neurons of the spinal cord[158]. A putative homologous cell type has been found in the ventral nerve cord of *Platynereis*[56]*.*

Adding these genes to the ProSPr atlases was a collaborative effort with Emily Savage helping with experimental work and Paola Bertucci overseeing the bioinformatic processing. I was involved in different capacities for each of these genes/stages (Table II-1).

The end result was that all the TFs that pattern the mouse spinal cord (Supplementary figure 1) were part of the ProSPr atlases of 48 hpf, 72 hpf, and 6 dpf. This allowed us to make comparisons to every cell type of the vertebrate spinal cord.

Table II-1. Genes that I contributed to add to the ProSPr atlases

The probes I designed are detailed in the Material and Methods, section IV.6. Pre-WMISH can be anything from fixing larvae, transcribing the RNA probe, to transforming bacteria to produce more of the plasmid for the transcription of the probe, for example. The WMISH and Imaging protocols are referred to in the Material and Methods, sections IV.7 and IV.9. The Processing refers to the computational pipeline to add the acquired images into the ProSPr atlases. It involves many steps, and I was usually involved in rotating the 48 hpf larvae to a ventral position. This is something that has to currently be done manually.

Gene	**ProSPr atlas**	**My involvement**	
Dbx1	48 hpf	2, 3, 4, 5	1 = Probe design
DSCAM	48 hpf	2, 3, 4, 5	2 = Pre-WMISH work
Eph	48 hpf	1, 2, 3, 4, 5	3 = WMISH
Eve	48 hpf	2, 3, 4, 5	4 = Imaging
Gbx1	48 hpf	2	5 = Processing
Hb9	48 hpf	2, 3, 4, 5	
Gly1246	48 hpf	1, 2, 3, 4, 5	
Netrin	48 hpf	2, 3, 4, 5	
Beta3	72 hpf	2	
Chx10	72 hpf	2, 3, 4	
Eph	72 hpf	1, 2, 3, 4	
FoxD	72 hpf	2	
Gbx1	72 hpf	2	
Hb9	72 hpf	2, 3, 4	
Netrin	72 hpf	2, 3, 4	
Tal	72 hpf	2, 3, 4	
Netrin	6 dpf	2, 3, 4	
FoxP	6 dpf	2	
En	6 dpf	2,3, 4	
DSCAM	6 dpf	2, 3, 4	

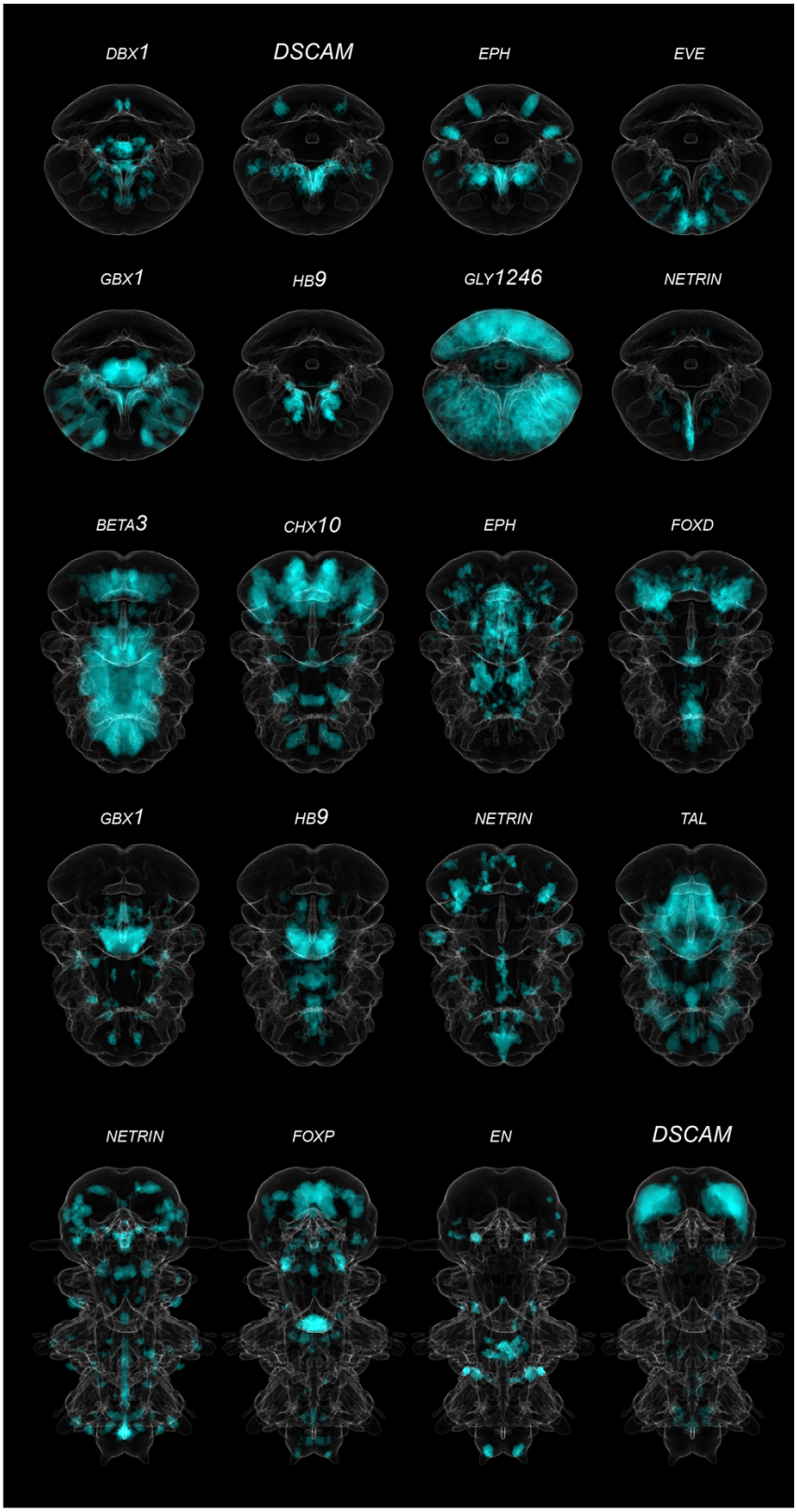


Figure II-1. Expression panels of genes for which I initiated and contributed to the addition to ProSPr atlases for several developmental stages

Ventral views of the expression patters of the genes that I helped add to the ProSPr atlases. The first two rows show the panels of the 48 hpf atlas. The third and fourth row show the 72 hpf atlas. The last row shows the 6 dpf atlas. See text for information on each gene.

2 Assessing Gene-Expression From The PlatyBrowser

To characterise the commissural neurons of *Platynereis*, we can extract the gene expression values for each of the segmented cells in the animal using the *PlatyBrowser*. For this we can use the overlap in volume between the gene-expression domain of every gene on the 6 dpf ProSPr atlas and every segmented nucleus of the SBF-SEM (hereafter referred to simply as overlap value, see Definition 1). I use the segmented nuclei and not the segmented cells because their segmentation was better and because they occupy the vast majority of a neuron's volume[150]. We therefore gain in accuracy of the segmentation without losing much of the coverage of the cell's volume.

These overlap values do not reflect how strongly a gene is expressed, but rather indicate whether a gene is expressed or not in a given cell. To determine how the overlap values inform us about gene expression, I analysed how they are distributed across all of *Platynereis'* segmented nuclei at 6 dpf.

Since the registration of the PlatyBrowser is done using the segmented nuclei of the SBF-SEM dataset and the 6 dpf ProSPr DAPI reference, it also performs best in regions with a high nuclear density. One such region is the VNC (Christian Tischer, personal communication). The overlap values will be little affected by registration artefacts in the VNC. Taking advantage that the VNC is segmented in the PlatyBrowser (see Introduction, section 6.1.3), I studied the overlap values in the segmented nuclei of the VNC.

Plotting these values in a frequency polygon reveals that they have a binomial distribution, with a large peak close to 0 and a smaller peak close to 1 (Figure II-2). This means that genes are either not expressed or expressed, respectively.

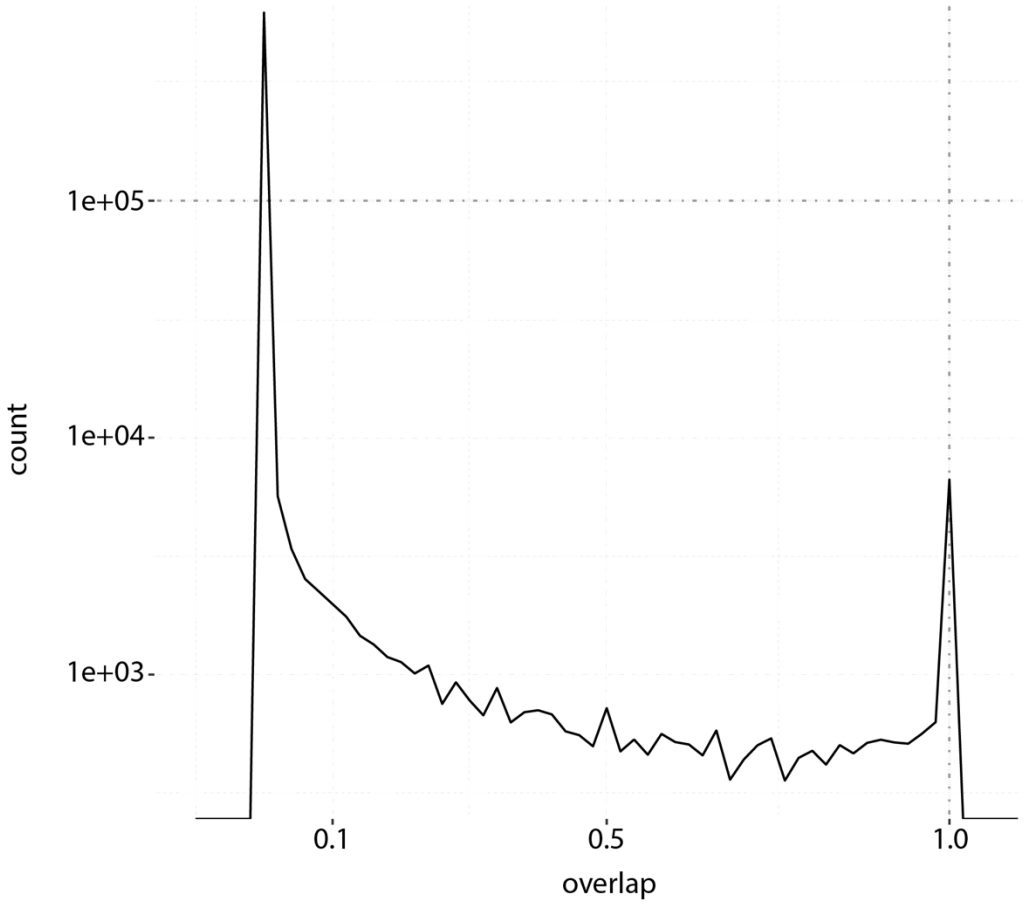


Figure II-2. Distribution of the overlap values in VNC nuclei

This is a frequency polygon that shows the distribution of the overlap values divided into bins of size 0.02. Each overlap value is a count, and relates to one gene and one cell. The distribution of the overlap values is binomial, with the largest peak close to 0 and a smaller peak close to 1. Only the segmented nuclei of the VNC were used for this plot, but the distribution was very similar when all cells were added (data not shown).

Nevertheless, such binomial distribution is not representative of all genes. Broadly-expressed genes, like *Elav*, fully cover (overlap = 1) thousands of cells. *Elav* is a conserved RNA-binding protein with a role in neuronal differentiation. Other genes, like the NKL homeobox transcription factor *Dbx1*, are sparsely expressed, with small regions of expression that do not fully cover many cells (Figure II-3a).

I classified as broadly-expressed (broad) any gene that had an overlap value of 1 in more than 10 cells. Sparsely-expressed genes (sparse) were all the rest, with overlap values of 1 in 10 cells or less. Of the 201 genes in the 6dpf ProSPr atlas, 147 are expressed in the VNC. 16 were classified as broad and 131 as sparse (Supplementary table 1).

In contrast to the binomial distribution of the overlap values for the broad genes, the distribution for the sparse genes has a single peak, close to 0 (Figure II-3b).

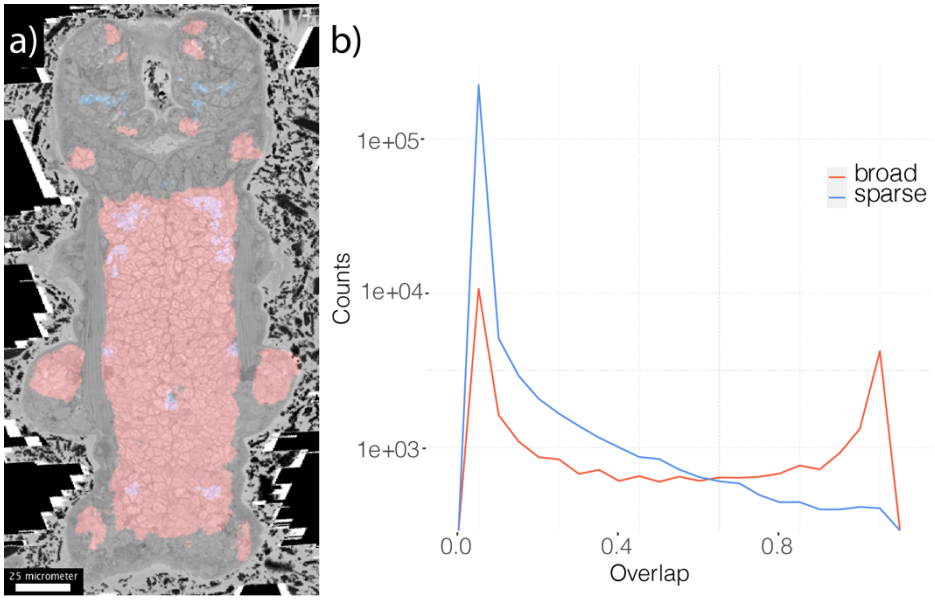


Figure II-3. Broad and sparse genes can be distinguished by their expression patterns and by the distribution of their overlap values

a) The expression patterns of *Elav* (red) and *Dbx1* (blue) are shown in a ventral view of the PlatyBrowser. Dbx is a sparsely-expressed gene and completely covers very few cells. b) Frequency polygon showing the distribution of overlap values divided into bins of size 0.02 for broad and sparse genes in the segmented nuclei of the VNC. Every overlap value is a count. Broad genes follow a binomial distribution, similar to when all genes are confounded together. Sparse genes do not follow a binomial distribution and have a single peak close to 0.

Since the overlap values are a function of the nuclear volume, I reasoned that the interpretation of these values could be different in nuclei of very different volumes. I decided to continue our analysis in nuclei with a homogeneous volume.

While most VNC nuclei have a volume of around 100 um^3, some were much larger, reaching up to 739 um^3 (Figure II-4). I removed from my subsequent analyses all nuclei with a volume higher than 150 um^3.

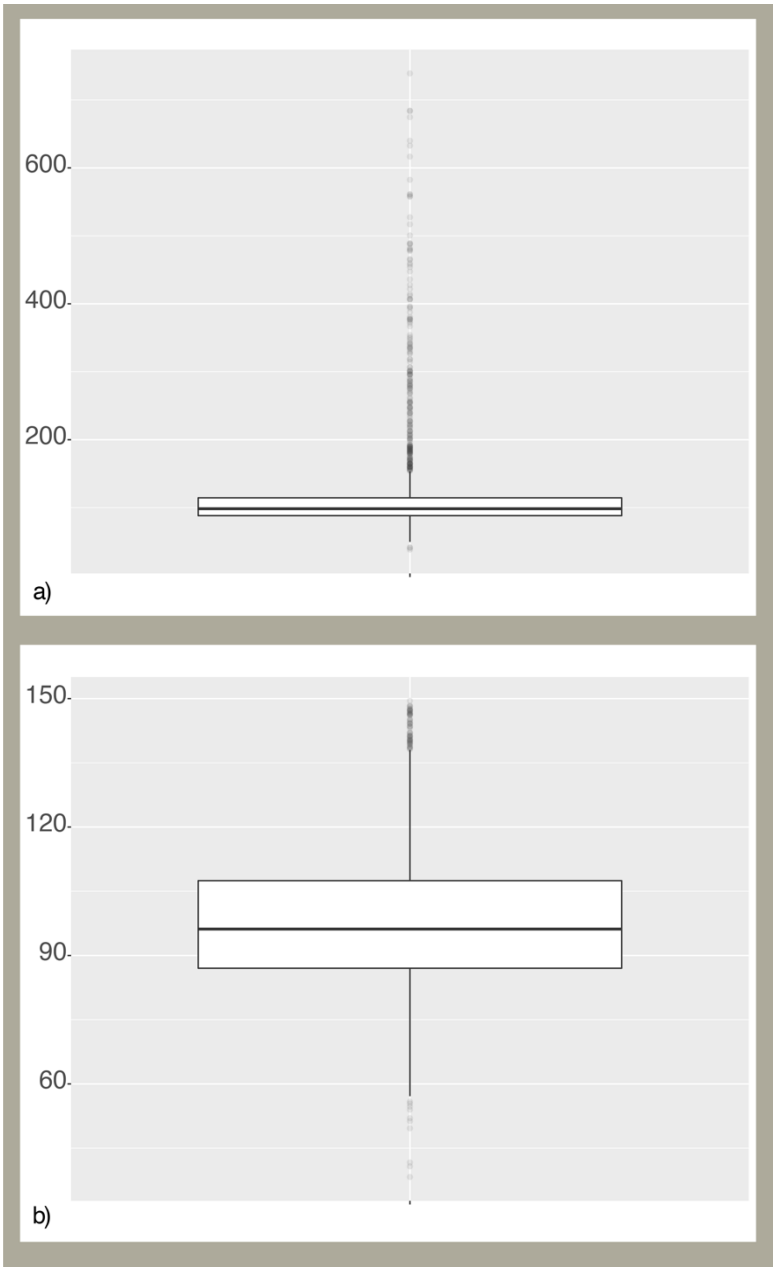


Figure II-4. Volume of the segmented nuclei of the VNC before (a) and after filtering out nuclei above 150 um^3 (b)

The y-axis shows the volume in um^3. The white box is delimited by the first quartile (Q1) at the bottom and the third quartile (Q3) at the top. The black line between these two quartiles is the median. The whisker at the top of the boxplot extends from Q3 until Q3 + a, where a = 1.5 x (Q3 – Q1). The whisker at the bottom of the boxplot extends from Q1 until Q1 – a. Any observation beyond the whiskers is represented with a translucent dot.

An overlap value of a nucleus is a fraction its volume. 1 overlap means that the entire volume of the nucleus overlaps with a gene's expression. In the filtered nuclei of the VNC, this will be close to 95 um^3, their median volume. An overlap value of 0.5 in two neighbouring cells means that the expression region of this gene has a volume of around 100 um^3, the equivalent of one cell volume. One could argue that this expression domain corresponds to a single nucleus and not two. To estimate how many nuclei

express a gene, one could divide the volume that the expression domain of this gene occupies by the median volume of a nucleus. This can be found by simply adding all the overlap values of a gene and rounding it up.

Another method to calculate how many nuclei express a gene is by deciding that all overlap values above a certain threshold will be considered as indicating that the gene to which the overlap value relates is expressed in the corresponding nucleus.

Both of these methods calculate the number of cells that express a gene in a slightly different manner. To visualize this I plotted the number of cells that express a gene as calculated through the volume of the gene's expression domain (x-axis) and by the thresholding method (y-axis) (Figure II-5).

I found that the threshold value at which both methods gave the most similar results was 0.4 for sparse genes, and 0.5 for broad genes (see legend of Figure II-5 for additional explanation).

I used these results to guide my interpretation of the overlap values. In some cases I found it useful to binarize the data. I used thresholds for this. High thresholds provide conservative results, less likely to contain false positives but more likely to contain false negatives. This is desirable when we want to determine with some confidence the gene-expression profile of a cell, for example.

But there are cases in which using a lower threshold is desirable. Lower thresholds are less likely to contain false negatives but more likely to contain false positives. This is desirable, for example, when we want to search for cells based on their gene-expression profile. In this case it is better to start the search with a low threshold value and then validate the cells through other methods, such as the presence of a bilateral homolog on the other side of the body (which proved to be the case in many instances).

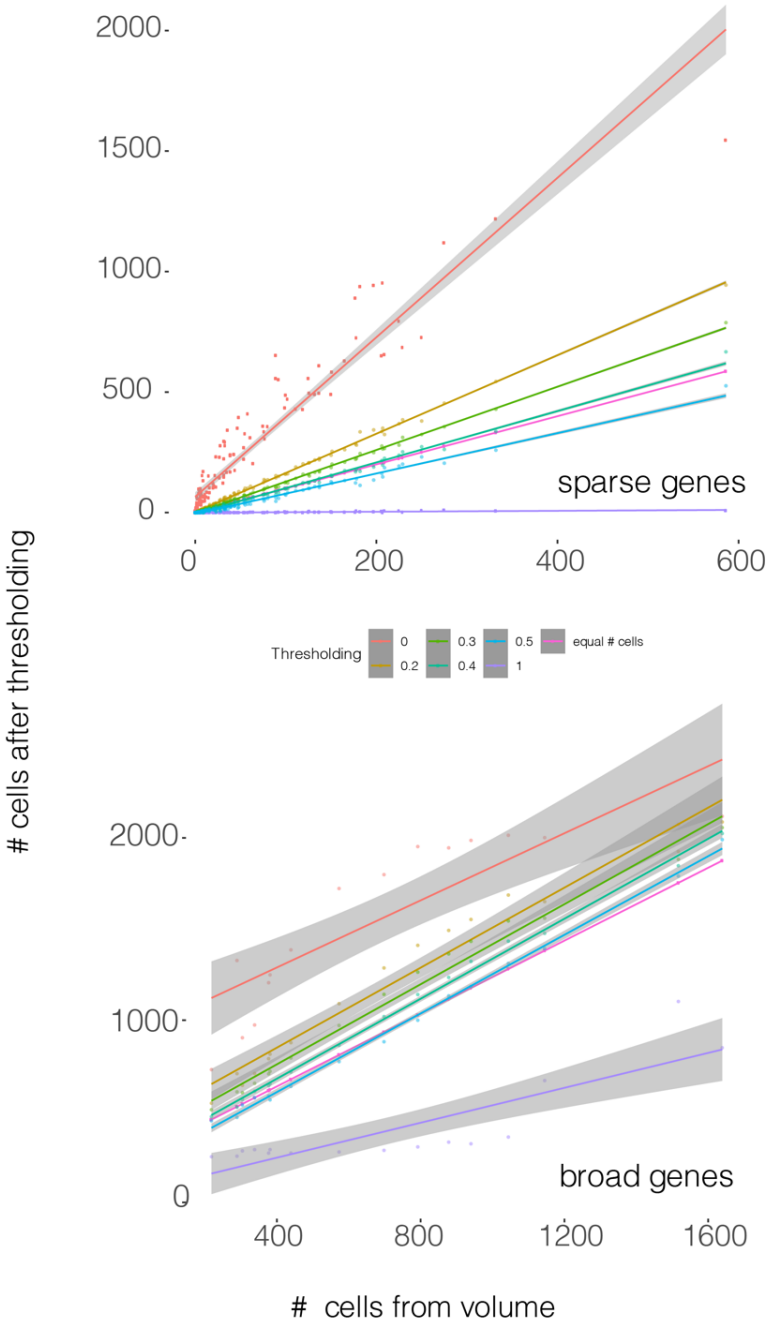


Figure II-5. Number of cells that express sparse or broad genes, as calculated from the threshold and the volume of the expression pattern

Six different thresholds (0, 0.2, 0.3, 0.4, 0.5, and 1) were used on every sparse and broad gene. For each threshold, a dot represents a different gene. The value that it takes in the x-axis corresponds to the number of cells that express this gene, estimated as a function of the volume that this gene's expression domain. Every gene will have the same value on the x-axis, regardless of the threshold. The y-axis corresponds to the number of cells that express this gene, estimated by deciding that every overlap value equal or above the threshold means that it is expressed in the corresponding nucleus. A regression line (coloured line) and its standard error (grey area) are shown for every threshold. This was estimated using the R package ggplot2. The magenta line is a reference straight line where x and y are equal. I reasoned that a good threshold value should result in a similar number of cells, regardless of which of the two methods described earlier is used to estimate it. For the sparse genes, the threshold that is closest to our reference magenta line is 0.4; for the broad genes it is 0.5.

3 Biased Search Of Commissural Neurons Guided By The Vertebrate Spinal Cord.

As discussed in the Introduction, commissural neurons that express the homeobox transcription factor even-skipped have been found in many Ecdysozoans and Vertebrates, and in the case of *Drosophila*, mouse, and zebrafish, it has been shown that they project anteriorly[81]. In the mouse and chick, they correspond to the V0 commissural interneurons.

The vertebrate V0 commissural interneurons arise from the progenitor domains p0 (Supplementary figure 1). They are best-characterised in the mouse, though their gene-expression program, throughout development, is well conserved across vertebrates[66].

Table II-2. TFs expressed in the p0 and V0 cell types of the mouse spinal cord
The p0 progenitor domain gives rise to three V0 cell types. This table lists the TFs that each cell type expresses (expression data taken from Ref.[158].

P0	*Olig3*	*Dbx1*	*Dbx2*	*Irx3*	*Neurog1*	*Neurog2*	*Pax6*
V0d	*Bhlhb5*	*Lhx1*	*Lhx5*	*Pax2*			
V0v	*Evx1*	*Evx2*	*Lhx1*	*Lhx5*			
V0c	*Evx1*	*Evx2*	*Lhx1*	*Lhx5*	*Pax2*	*Pitxb2*	

The *Platynereis* orthologs of the *bHLH* TFs *Olig3, Bhlhb5,* and *Neurog1* were first identified by Simionato *et al.*, 2008[165], and of the homeobox TF *Evx2* by de Rosa *et al.,* 2005[166]. The *Platynereis* orthologs of the NKL homeobox TF *Dbx1* and the LIM-homeodomain TFs *Lhx1* and *Lhx5* were first identified by Denes *et al., 2007*[46]. The ortholog of the Iroquois TF *Irx3* was first identified by Marlow *et al.,* 2014[167], and of the paired-homeobox TFs *Pax2, Pitxb2,* and *Pax6* by Backfisch *et al.,* 2013[168], Tessmar-Raible, 2004[169], and by Raible *et al.*, 2005[144], respectively (Table II-3).

Table II-3. *Platynereis* orthologs of the mouse TFs that are expressed in the p0 and V0 cell types of the spinal cord

The first row lists the mouse TFs that are expressed in the p0 and V0 cell types, as well as their paralogs. The second row lists the *Platynereis* TFs that are orthologous to the TFs in the first row.

Mouse paralogs	olig1, olig2, olig3	*Dbx1*, dbx2	irx1, irx2, irx3, irx4, *Irx6*	neurog1, neurog2	*Pax6*	evx1, evx2	bhlhb5	lhx1, lhx5	pax2	pitx1, pitx2, pitx3
***Platynereis* homologs**	olig	*Dbx1*	*Irx6*	*Ngn*	*Pax6*	eve	beta3	lhx1/5	pax2/5/8	*PitxB*

I used the ProSPr and *PlatyBrowser* atlases to explore the cellular expression patterns in the VNC of the *Platynereis* orthologs of the vertebrate TFs that are expressed in the p0, V0d, V0v, and V0c cell types of the mouse spinal cord. I investigated the expression of each gene as well as the coexpression of several combinations at 48 hpf, 72 hpf, and 6 dpf (Supplementary figure 2-4).

I took into account the divergence time between vertebrates and Platynereis and did not expect for the full TF signature to be necessarily conserved. This could also be due to evolutionary divergence time or due to the fact that we are dealing with non-corresponding differentiation stages. Importantly, the different segments of the larvae harbour neurons that are at differentially advanced developmentally: the neurons of the second segment will be at a later developmental stage than those in the third segment (see Introduction, section I.3). It is therefore important to look at as many larval stages as possible.

3.1 Biased Search For Commissural Neurons Corresponding To The p0 Progenitor Domain Of The Vertebrate Spinal Cord

3.1.1 *Dbx1*$^{+}$ Neurons In The VNC Of *Platynereis*

The vertebrate p0 progenitor domain gives rise to the commissural V0 interneurons, a process that is controlled by the NKL homeobox transcription factor *Dbx1*[68]. In the presence of *Dbx1*, these cells differentiate into commissural interneurons that coordinate left-right alternation[69]. *Dbx1* negatively regulates the homeobox transcription factor *engrailed* (*En*), which drives the differentiation of neurons to an ipsilateral V1 fate. *Dbx1* also upregulates the homeobox transcription factor *Evx*, which drives differentiation to a V0 fate [68].

This crucial role of *Dbx1* in the differentiation of the p0 progenitor cells into postmitotic V0 commissural neurons drove us to study this *TF* in Platynereis. We studied the cells that express *Dbx1* as well as the role that *Dbx1* plays in their axonal projection.

I started by investigating the axonal projection of the neurons in the VNC that express *Dbx1* in the 48 hpf larva. To do this, I combined WMISH against *Dbx1* with immunostaining against acetylated tubulin (see Materials and Methods section IV.8). The WMISH provided us with signal from the nuclei that express *Dbx1*. The immunostaining provided us signal from the acetylated tubulin of the axons. I used this signal to trace the axons of *Dbx1*$^{+}$ neurons of the VNC with the Fiji plug in Simple Neurite Tracer (see Materials and Methods section IV.10). I traced some of these axons across the midline (Figure II-6), meaning that they are commissural.

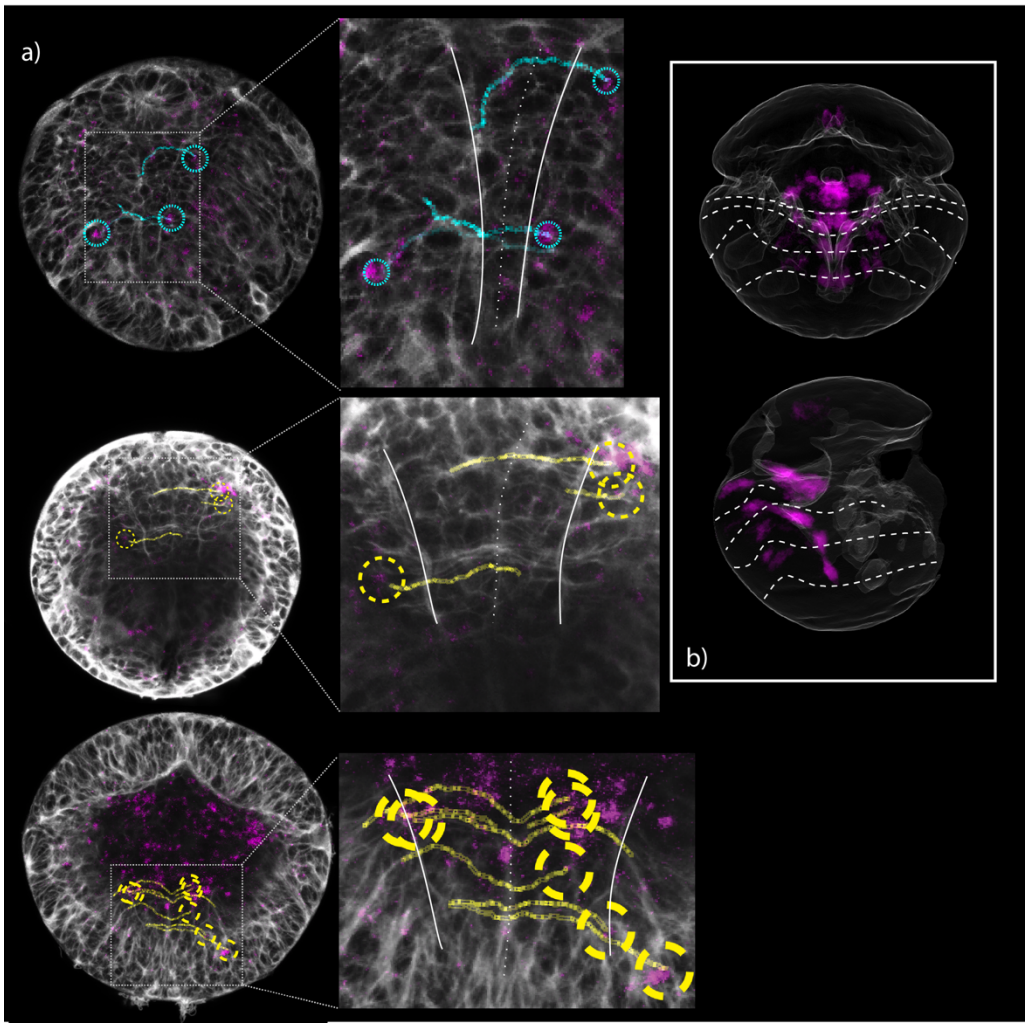


Figure II-6. Commissural axons of *Dbx1*$^{+}$ neurons at 48 hpf

a) Z-projections of confocal stacks of three larvae at 48 hpf that underwent a WMISH against *Dbx1* (magenta) and an immunostaining against acetylated tubulin (grey). In each case, a panel zooming on the VNC is shown where two white lines delineate the nerve cords and the dashed line delineates the midline. I traced the axons of the neurons that express *Dbx1* (see Materials and Methods, section IV.10) and found that some of them were commissural. The traces of these commissural axons are shown in cyan or yellow. The dashed circles indicate the location of the somas of the traced axons. **b)** ProSPr expression panel of *Dbx1* at 48 hpf. The dashed lines indicate the anterior boundaries of segments 1, 2, 3 and 4 (see Materials and Methods, section IV.14).

3.1.1.1 Functional Analysis Of *Dbx1* Mutant Larvae

To assess whether *Dbx1* is necessary in *Platynereis* for the axonal projection of the commissural neurons of the VNC, a functional analysis of *Dbx1* was performed by Hernando Martinez-Vergara and Ryan Prestil[156]. I was not involved in the experimental part of this study nor in the first analysis, but I did perform a second analysis of the results.

Dbx1 mutants were obtained by microinjecting zygotes with Cas9 mRNA and a sgRNA against *Dbx1*. At 50 hpf, when the ventral nerve cords are furthest apart and the commissural axons can be seen best, they fixed the larvae. Next, they stained them with an antibody against acetylated tubulin (AcTu) and imaged the larvae on a confocal microscope (Figure II-7). The protocols they used for this are the same as I describe in sections IV.4, IV.5, and IV.9 of the Materials and Methods.

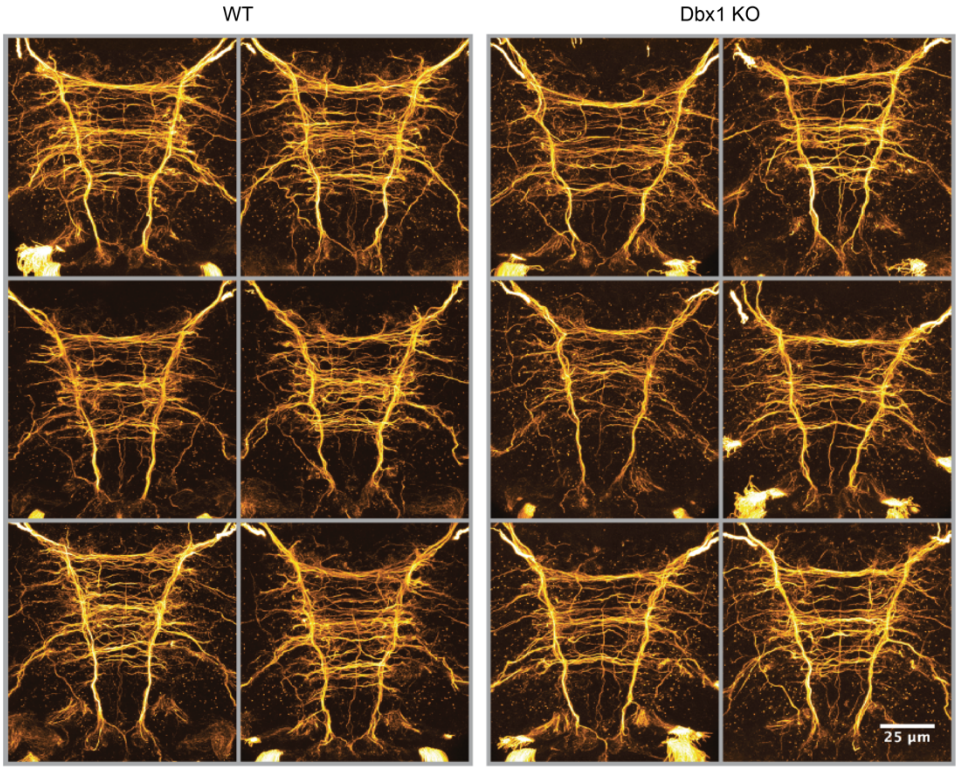


Figure II-7. VNC acetylated tubulin immunostaining - Reproduced from Ref. [156].
These panels show maximum projections of confocal stacks of wildtype (WT) and *Dbx1* mutant (Dbx KO) larvae that were stained with an antibody against acetylated tubulin.

They assessed the mutant phenotype by analysing the confocal stacks with the AcTu staining. They estimated the commissural density using a Fiji macro that analysed the AcTu fluorescent signal. First, a rectangular region of interest (ROI) between the two nerve cords was selected where there was only signal from the commissural axons and that covered the whole antero-posterior extent of the commissural axons. This ROI was then projected to the YZ view. Next, the peaks in fluorescence along the antero-posterior axis at each X plane were quantified. These peaks are measured from to signal from the commissural axons, and the number of peaks was used as an estimate of the quantity of commissural axonal bundles, a measure they refer to as commissural density. Please refer to Ref. [156] for more details. They found that the *Dbx1* mutant has a significantly lower commissural density than control larvae that were injected with the Cas9 mRNA but no sgRNA[156].

I reanalysed the confocal stacks from this study. Instead of measuring the commissural density through image analysis, I manually counted the number of commissural axons in the VNC (see Materials & Methods section IV.12 and Supplementary table 3). This method, while more time-consuming and labour-intensive, directly counts the number of commissural axons. I think that it is a better representation of the quantity of commissural axons than the image-analysis pipeline.

Similarly to the analysis in Ref. [156], I found that the *Dbx1* mutant had significantly fewer commissural axons in the VNC than the control or wildtype larvae (Figure II-8). Also in agreement with the analysis in Ref. [156], I found no significant difference in the number of commissural axons between the control and uninjected control larvae.

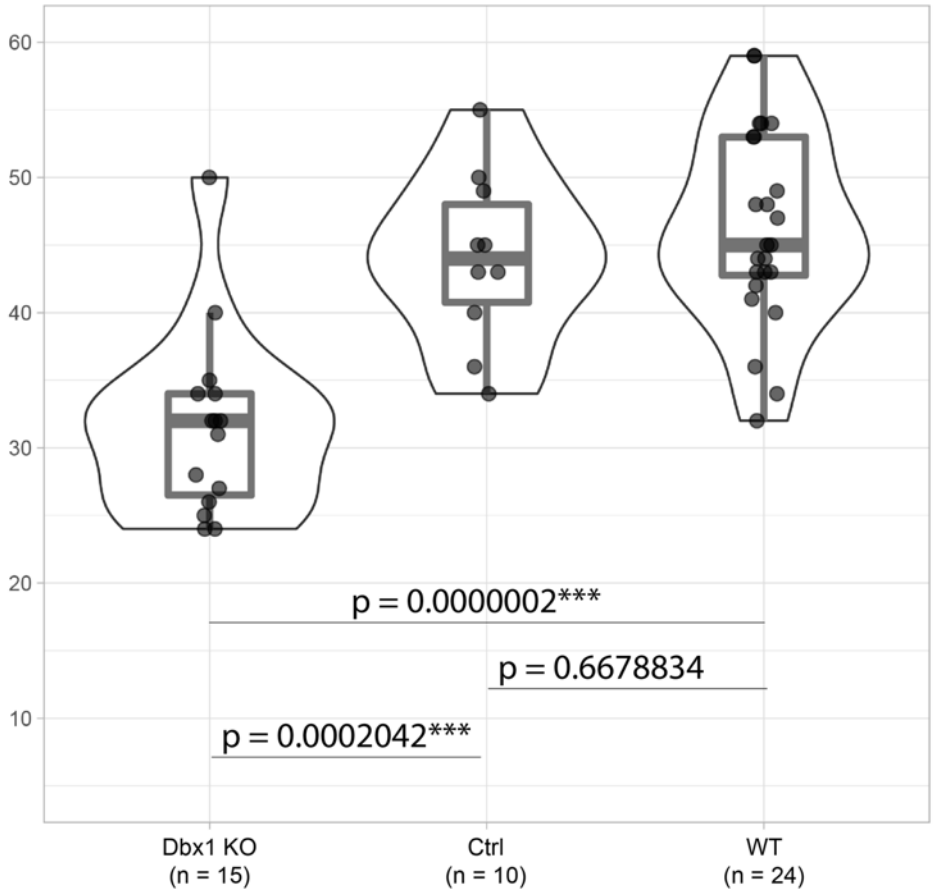


Figure II-8. Comparison of the number of commissural axons in *Dbx1* KO, injected control (Ctrl), and uninjected control larvae (WT)

Each point in these violin plots is a single larva. The value that each point takes in the y-axis is the number of commissural axons I counted in the VNC in that larva. For each violin plot, the white box is delimited by the first quartile (Q1) at the bottom and the third quartile (Q3) at the top. The black line between these two quartiles is the median. The whisker at the top of the boxplot extends from Q3 until Q3 + a, where a = 1.5 x (Q3 – Q1). The whisker at the bottom of the boxplot extends from Q1 until Q1 – a. The curves surrounding the boxplot (violins) represent the kernel probability density at different y-values. N is the number of larvae. The p values are from Tukey's Test after ANOVA. Sig. * < 0.5, ** < 0.1, *** < 0.01.

3.1.1.2 Functional Analysis Of Neo/DCC Mutant Larvae

Using the 48 hpf ProSPr atlas, we found a striking overlap in the expression patterns of *Dbx1* and Neo/DCC (Figure II-9). Neo/DCC is the *Platynereis* homolog of the *Drosophila frazzled* gene and the vertebrate *Neogenin* and *DCC* genes. These are Netrin receptors with a conserved role in axonal guidance across the midline (see Introduction, section I.5).

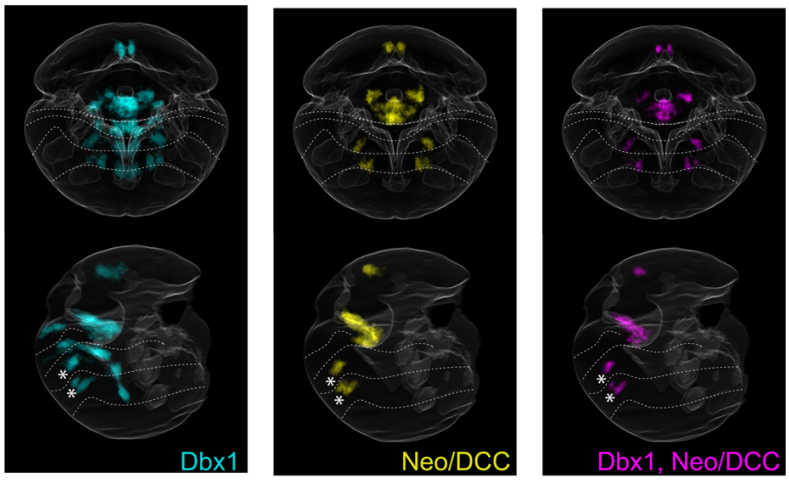


Figure II-9. Expression panels for the *Dbx1*, Neo/DCC, and for the coexpression of *Dbx1* and Neo/DCC

These panels were generated using the ProSPr atlas of 48 hpf. The expression domains of *Dbx1* and Neo/DCC are shown in the left and middle panels. The right panel shown the overlap in the expression domains of *Dbx1* and Neo/DCC. The dashed delineate the anterior boundaries of segments 1, 2, 3, and 4, as assessed by expression of Engrailed (see Materials and Methods, section IV.14). The asterisks show two expression domains in the second and third segments where the expression patterns of *Dbx1* and Neo/DCC overlap.

We hypothesised that Neo/DCC could play a similar role in *Platynereis* and tested this through a functional study of Neo/DCC similar to *Dbx1* (Materials and Methods, section IV.11). This experiment was done in collaboration with Ines Kübler, a master student under my supervision. Ines did most of the experimental work, while I planned the study, designed sgRNAs supervised injections and analysed the data as presented in this thesis. Ines published this work as her Master's thesis with an analysis that she did using the commissural density measure. I counted the commissural axons manually (Supplementary table 3) and analysed the results.

The mutant larvae were generated by injecting a sgRNA directed to the Netrin receptor sequence in the first exon of the Neo/DCC gene and Cas9 mRNA (see Ref. [157], Fig. 8). Injected control larvae were generated by injecting a sgRNA whose sequence has no blast hits in the *Platynereis* genome. We also kept uninjected control larvae as a reference.

I found that the mutant larva have significantly fewer commissural axons than either the injected (Ctrl) and uninjected (WT) control (Figure II-11). In this case the injected control larvae also have significantly fewer commissural axons than uninjected control larvae, which may be due to a developmental delay due to the injection.

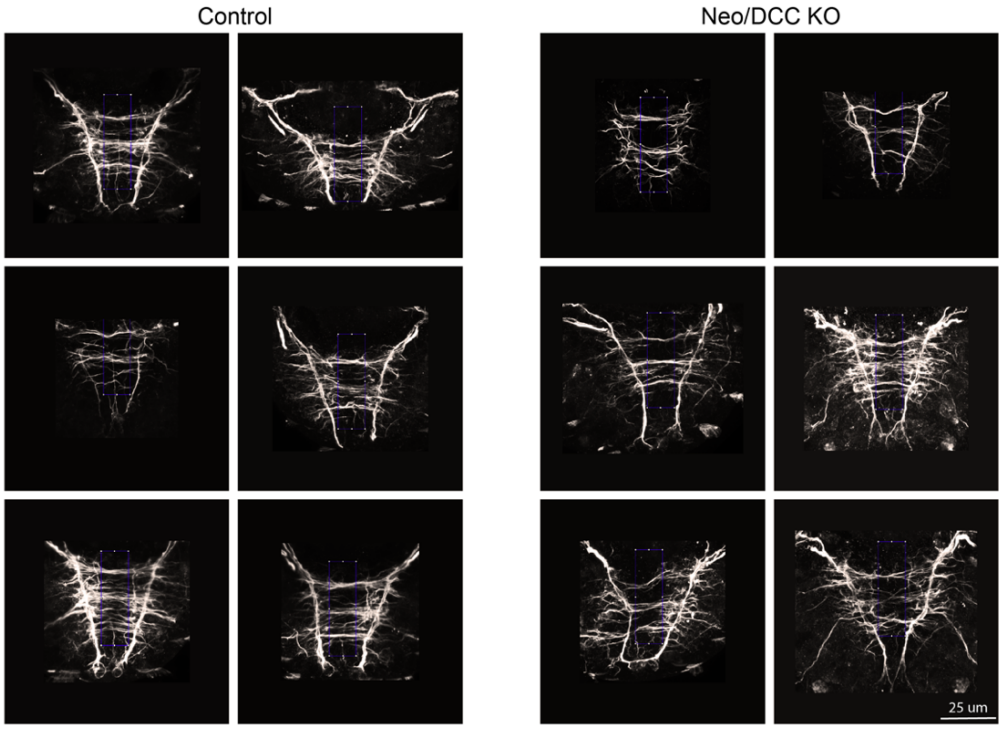


Figure II-10. VNC with acetylated tubulin immunostaining

These panels show maximum projections of confocal stacks of injected control and Neo/DCC mutant (Neo/DCC KO) larvae that were stained with an antibody against acetylated tubulin.

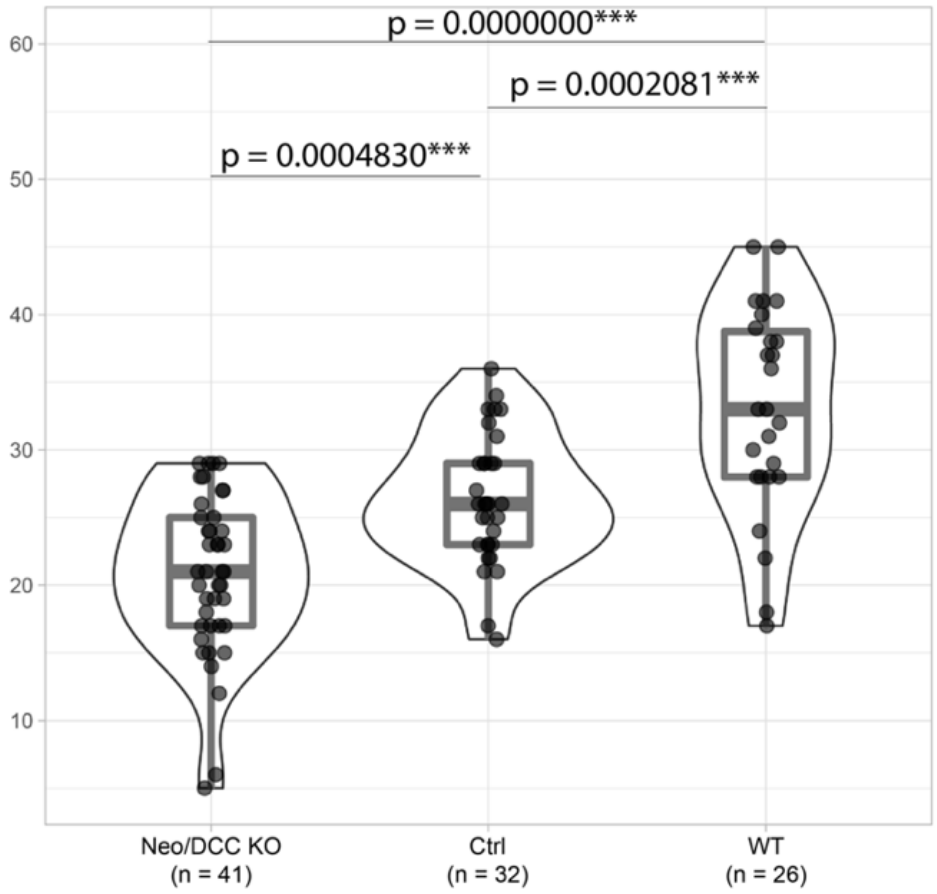


Figure II-11. Comparison of the number of commissural axons in Neo/DCC KO, injected control (Ctrl), and uninjected control (WT) larvae

Each point in these violin plots is a single larva. The value that each point takes in the y-axis is the number of commissural axons I counted in the VNC in that larva. For each violin plot, the white box is delimited by the first quartile (Q1) at the bottom and the third quartile (Q3) at the top. The black line between these two quartiles is the median. The whisker at the top of the boxplot extends from Q3 until Q3 + a, where a = 1.5 x (Q3 – Q1). The whisker at the bottom of the boxplot extends from Q1 until Q1 – a. The curves surrounding the boxplot (violins) represent the kernel probability density at different y-values. N is the number of larvae. The p values are from Tukey's Test after ANOVA. Sig. * < 0.5, ** < 0.1, *** < 0.01.

3.1.1.3 Development Of The *Dbx1*$^{+}$, Neo^{+} Commissural Neurons

I analysed the 48 hpf ProSPr atlas to determine which other TFs are expressed in the *Dbx1*+, Neo/DCC+ commissural neurons that we identified in the previous section. I found that the expression domains of *Dbx1* and Neo/DCC overlap with the zinc-finger transcription factor *Gata123* and the homeobox transcription factor *Hb9* (Figure II-12).

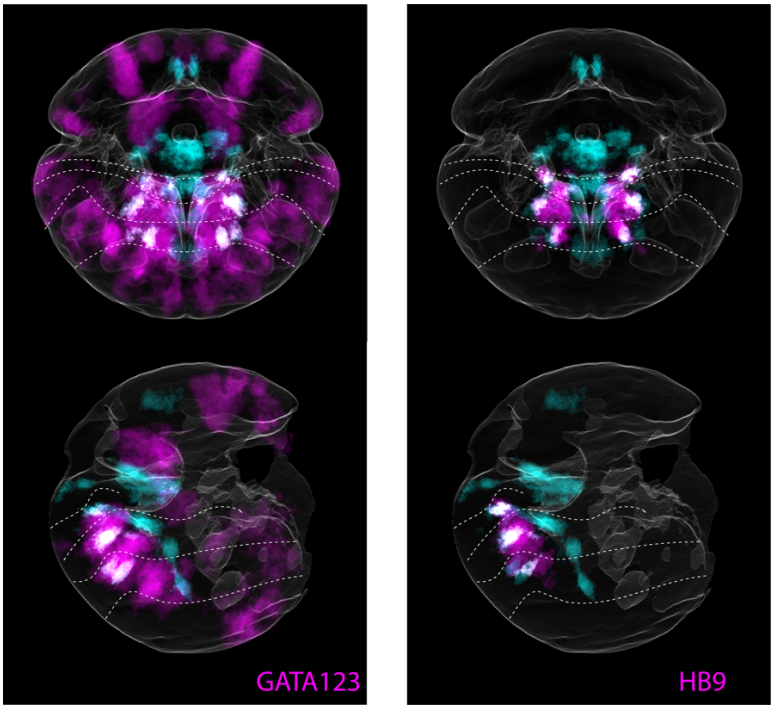


Figure II-12. TFs whose expression domains overlap with Dbx in segments 2 and 3
These panels were generated using the ProSPr atlas of 48 hpf. The expression domains of *Dbx1* are shown in cyan. The expression domains of TFs that overlap with *Dbx1* in segments 2 and 3 are shown in magenta. The dashed delineate the anterior boundaries of segments 1, 2, 3, and 4, as assessed by expression of Engrailed (see Materials and Methods, section IV.14). The top panels are a ventral view of the 48 hpf larva. The bottom panels are a lateral view with the ventral side on the left.

At 72 hpf, *Dbx1*, Gata123, and *Hb9* are coexpressed in smaller bilateral regions in segments 1, 2, and 3 (Figure II-13). Their location is very similar to the 48 hpf larva. At 6 dpf the coexpression region has disappeared from the segments 2 and 3, but remains in the first segment.

There is a bilateral region in the first segment where the expression domains of *Dbx1*, *Hb9*, and *Gata123* overlap at 48 hpf, 72 hpf, and 6 dpf. It is plausible that these are developmentally related.

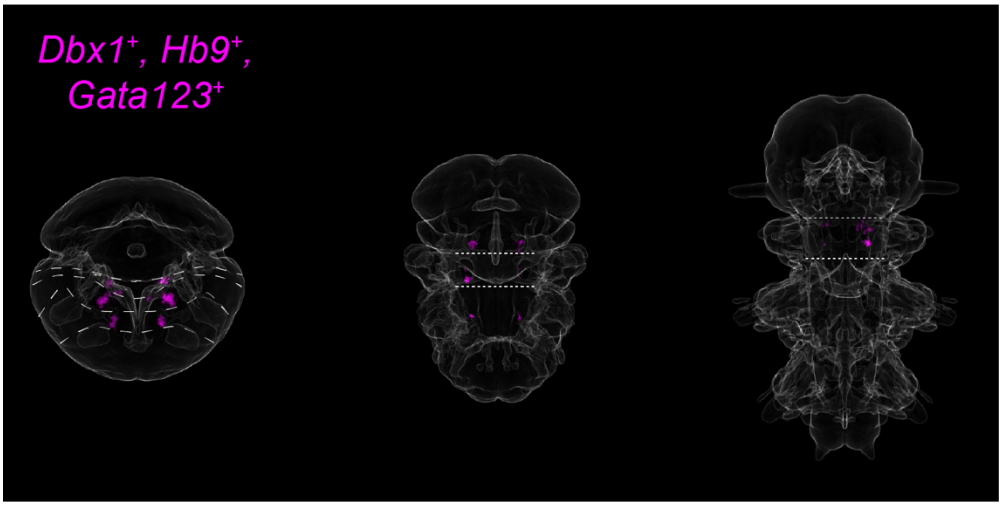


Figure II-13. Putative trajectory of *Dbx1*$^{+}$, *Hb9*$^{+}$, *Gata123*$^{+}$ cells

These panels were generated using the ProSPr atlases of gene expression at 48 hpf (left), 72 hpf (middle), and 6 dpf (right). The overlap in the expression domains of *Dbx1*, *Hb9*, and *Gata123* is shown in magenta at each stage. The dashed lines at 48hpf delineate the anterior segmental boundaries of the trunk, as assessed by expression of *Engrailed* (see Materials and Methods, section IV.14). The dashed lines at 72 hpf and 6 dpf delineate the anterior and posterior boundaries of segment 1.

Nevertheless, the *Dbx1*$^{+}$ and *Neo/DCC*$^{+}$ commissural neurons of the 48 hpf are located in segments 2 and 3.The fate of these neurons at 6 dpf remains unclear because we do not find regions in segments 2 and 3 that coexpress *Dbx1*, *Hb9*, and *Gata123*. I hypothesise that they would be found where *Hb9* and *Gata123* are coexpressed (Figure II-14), and that Dbx is downregulated at this stage of differentiation (as it happens with differentiating commissural neurons in the vertebrate neural tube; see Table II-2).

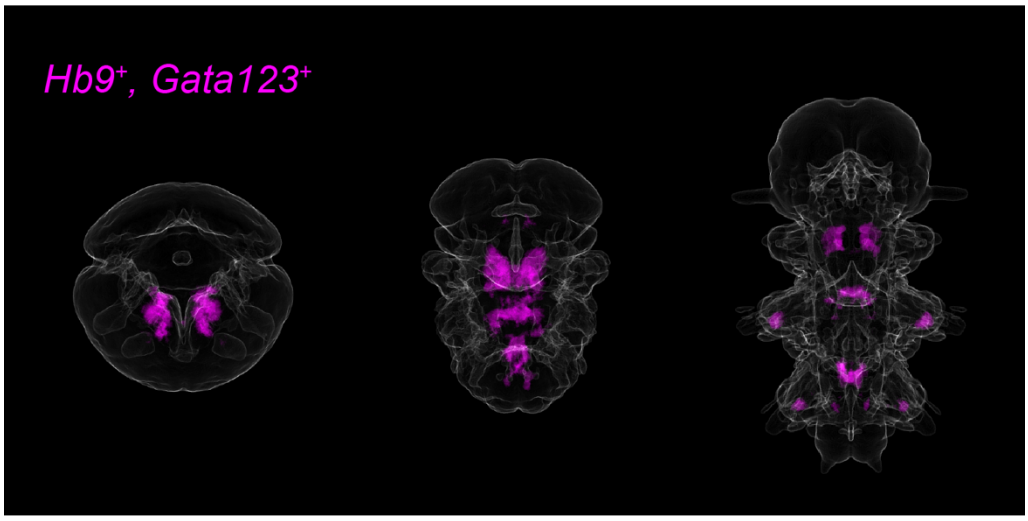


Figure II-14. Overlap of expression domains of *Hb9* and gata123

These panels were generated using the ProSPr atlases of gene expression at 48 hpf (left), 72 hpf (middle), and 6 dpf (right). The overlap in the expression domains of *Hb9* and *Gata123* is shown in magenta at each stage.

Since the first segment appears to harbour *Dbx1*$^{+}$ neurons that can be followed through development, I analysed its expression pattern more carefully. I searched if the *Dbx1*$^{+}$ cells of the first segment could be defined better by a different set of TFs.

I found that the paired-domain homeobox transcription factor *Pax6*, *Dbx1*, and the Iroquois homeobox transcription factor *Irx6* defined the cells of the first segment (Figure II-15). These are also all markers of the p0 progenitor domain of the vertebrate neural tube (Supplementary figure 1). Their expression does not overlap in segments 2 and 3.

3.1.1.4 Developmental Fate Of The *Pax6*+, *Dbx1*+, *Irx6*+ Domain

I found a bilateral region that expresses the p0 markers *Pax6*, *Dbx1*, and *Irx6* as early as 48 hpf (Figure II-15, Supplementary figure 2-2). It is located ventrally in the first segment, within the *Pax6* domain (Figure I-6) of the mediolateral patterning of the VNC. The bHLH transcription factor *Neurogenin* (*Ngn*) is also coexpressed, suggesting it might be a progenitor domain.

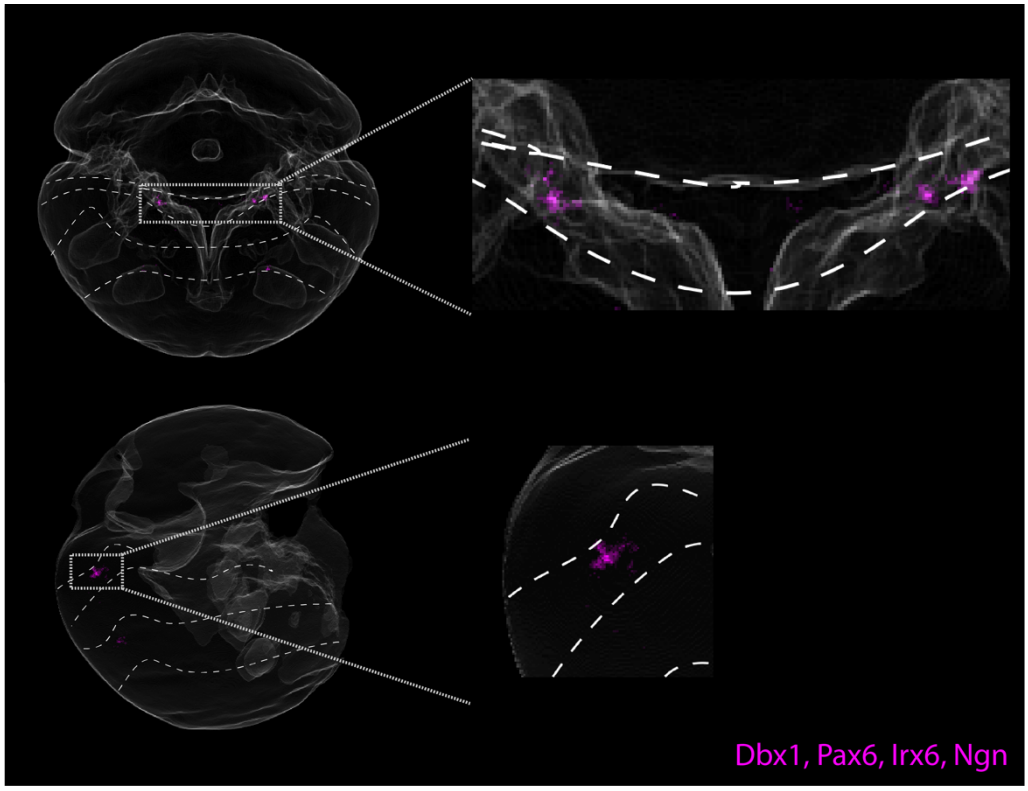


Figure II-15. Coexpression of *Dbx1*, *Pax6*, *Irx6*, and *Ngn* in the 48 hpf ProSPr atlas

These panels were generated using the ProSPr atlas of 48 hpf. They show the overlap in the expression of *Dbx1*, *Pax6*, *Irx6*, and *Ngn* in magenta. The dashed lines delineate the anterior segmental boundaries of the trunk, as assessed by expression of Engrailed (see Materials and Methods, section IV.14). The top panels are a ventral view of the 48 hpf larva. The bottom panels are a lateral view with the ventral side on the left. The inlays zoom in on the same coexpression region located in the ventral first segment.

Regions with the same expression pattern can be found at 72 hpf and 6 dpf, but they no longer express *Ngn* (Figure II-16). The similarities in TF-expression profile and location in the first segment lead us to believe that these cells are developmentally related.

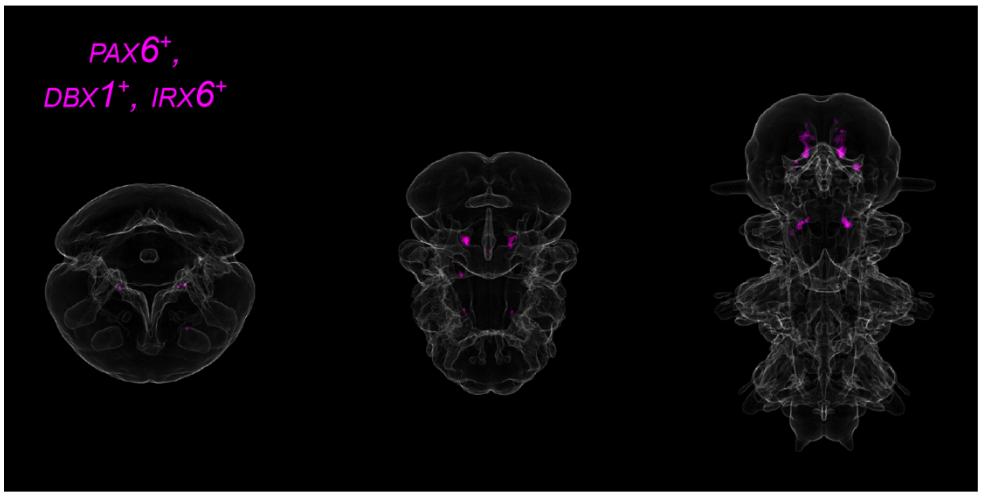


Figure II-16. Putative development of *Pax6*$^{+}$, *Dbx1*$^{+}$, *Irx6*$^{+}$ cells

These panels were generated using the ProSPr atlases of gene expression at 48 hpf (left), 72 hpf (middle), and 6 dpf (right). The overlap in the expression domains of *Pax6*, *Dbx1*, and *Irx6* is shown in magenta at each stage. The dashed lines at 48hpf delineate the anterior segmental boundaries of the trunk, as assessed by expression of Engrailed (see Materials and Methods, section IV.14). The dashed lines at 72 hpf and 6 dpf delineate the anterior and posterior boundaries of segment 1.

3.1.1.5 Molecular Characterisation And Axonal Tracing Of The *Pax6*$^{+}$, *Dbx1*$^{+}$, *Irx6*$^{+}$ Neurons In The 6 Dpf Larva

We investigated the neurons that express *Pax6*, *Dbx1*, and *Irx6* at 6 dpf. We used the PlatyBrowser[150] to identify them further and characterize them molecularly. I also used the *Platynereis* SBF-SEM dataset to trace them using the software PyKnossos (see Materials and Methods, sectionIV.13). Since our interest is in the commissural neurons of the trunk, we limited our search to the segmented nuclei of the VNC (see Introduction, section 6.1.3).

To locate these cells in the PlatyBrowser, we extracted the overlap values (Definition 1) of every segmented neuron of the VNC and used a threshold of 0.1 to decide whether a gene was expressed. I considered all values below 0.1 as absence of expression, and all values equal or above 0.1 as expression. I chose this permissive threshold to include many cells that might coexpress these TFs, yet have low overlap values. I later validated these cells using their full gene-expression profile, morphology, and bilaterally symmetrical body position.

I found 30 segmented nuclei that coexpress *Pax6*, *Dbx1*, and *Irx6*. 18 are in the first segment of the VNC, 2 in the second segment, and 1 in the third segment (Figure II-17). Note, however, that the ProSPr panel on (Figure II-16) shows there is no region in segments 2 or 3 where the expression domains of *Pax6*, *Dbx1*, and *Irx6* overlap.

This apparent contradiction can be reconciled by remembering that I consider that a gene is expressed in a cell when its expression domains cover at least 0.1 of the volume of the corresponding segmented nucleus. The volume of the segmented nucleus that these genes occupy does not have to be overlapping, which is why they are identified in the PlatyBrowser but not in the ProSPr panels. We can validate the segmented nuclei by taking into account their full gene-expression pattern, position in the body, and bilateral symmetry.

The segmented nuclei that I found in segments 2 and 3 are identified as 25068, 16459, and 16803 in the heatmap of Figure II-17. Their overlap values for *Pax6*, *Dbx1*, and *Irx6* are low. They are also not found as bilateral pairs. This suggests that they are ‘false positives’ and I discarded them from our subsequent analysis. I then proceeded to trace the 18 validated neurons with PyKnossos.

Both PyKnossos and the PlatyBrowser use the same SBF-SEM dataset. It is therefore possible to identify the exact same neuron on both PyKnossos and the PlatyBrowser. In PyKnossos the coordinates are given in voxels (1 voxel = 10 x 10 x 25 nm), and in the PlatyBrowser they are given in µm.

I recovered the coordinates of the validated segmented nuclei of the PlatyBrowser and transformed them to voxels. I moved to the transformed coordinates in PyKnossos and ensured that they pointed to the same nucleus. At a random point within this nucleus I placed a seeding point (Figure II-18b).

To trace in PyKnossos one places nodes along the axon that become connected with a line (Figure II-18a). The PyKnossos software has four navigation panes: XY, YZ, XZ, and a fourth pane that is orthogonal to the line that connects the nodes of a trace.

Tracing the whole neuron was not always possible because the membranes of more than one axon were sometimes fused and/or ruptured, an artefact that was likely caused during the fixation step (Figure II-19a). The volume dataset also has some alignment problems, but navigating with the 4 view panes usually made it possible to still trace most axons (Figure II-19b).

I determined that an axon was commissural if I could trace it crossing the midline. I used the axochord[170] as a reference point for the midline (Figure II-20). The axochord is a muscular structure at the base of the VNC that has been suggested to be homologous to the vertebrate notochord[170]. It expresses Slit and Netrin and could therefore be a major guidepost for commissural axonal guidance.

The 18 validated *Pax6+, Dbx1+,* and *Irx6+* are all in the first segment. I was able to determine that 4 of them were commissural (Figure II-17b). I could not trace the remaining 14 neurons far enough to decide whether they are commissural or not.

Of the four commissural neurons I found, two were on the left side of the body and two on the right side. Those on the left have a similar axonal projection: they project posteriorly before decussating medially and crossing the midline (Figure II-17b). Those on the right project posteriorly into the second segment before crossing the midline. This means they are describing commissural neurons.

The four commissural neurons express glutamate decarboxylase (gad) and choline acetyltransferase (chat). Unlike the p0 cell type of the vertebrate neural tube, they express *Hb9*. Also unlike the p0 progenitors, the cells in *Platynereis* are not progenitors but postmitotic, mature neurons, based on their expression of the potassium channels kv33b and kv33e (Figure II-17a).

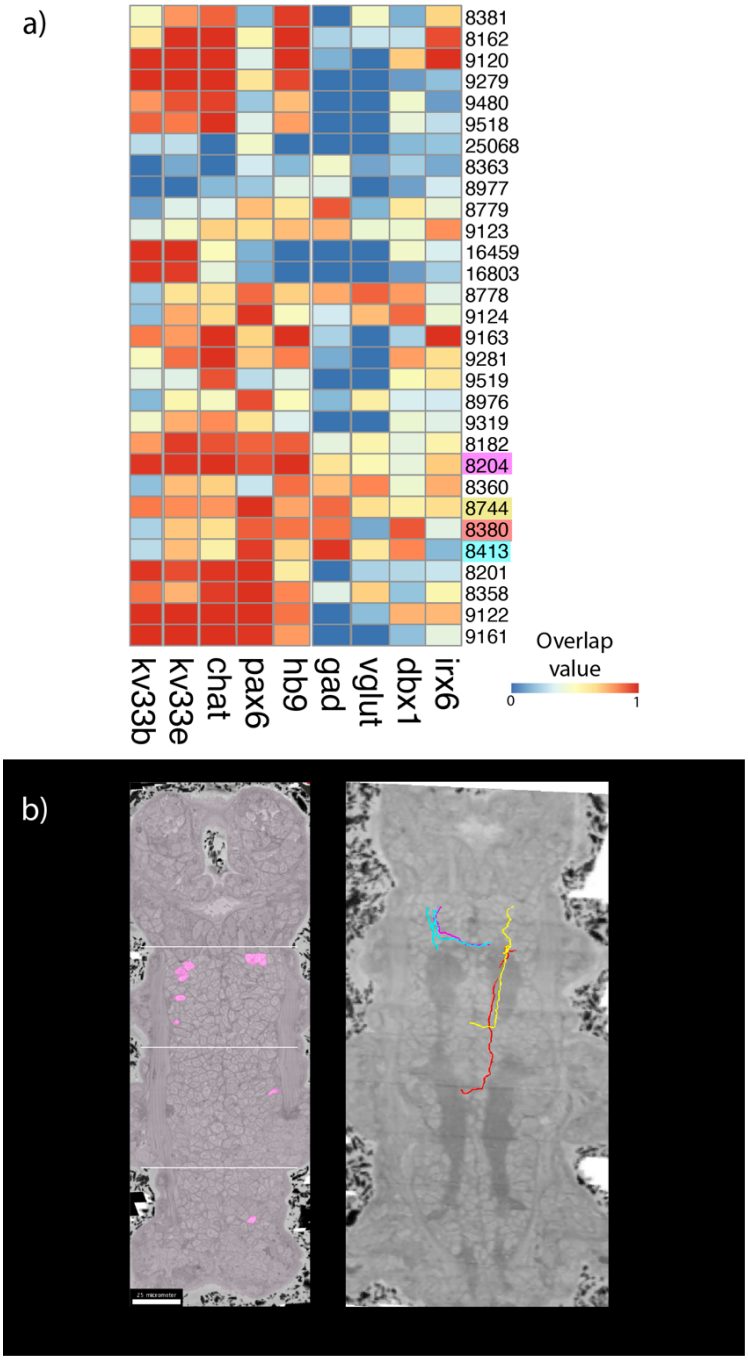


Figure II-17. Segmented nuclei that coexpress *Pax6*, *Dbx1*, and *Irx6*

a) Heatmap of all the segmented nuclei with overlap values (Definition 1) above 0.1 for *Pax6*, *Dbx1*, and *Irx6*. Selected genes are shown in the x-axis, and the y-axis is labelled with the IDs of the segmented nuclei. The labels that are coloured correspond to the commissural neurons whose traces are shown in the right panel of (b). Chat is choline acetyltransferase, an enzyme that catalyses the synthesis of acetylcholine. Kv33e and kv33b are voltage-gated potassium channels. Gad is glutamate decarboxylase, an enzyme that catalyzes the decarboxylation of glutamate to GABA. Vglut is vesicular glutamate transporter, a protein involved in loading synaptic vesicles with glutamate. *Pax6*, *Hb9*, *Dbx1*, and *Irx6* are TFs. This heatmap was created with the R package 'pheatmaps'. **b)** The left panel show a ventral plane of the PlatyBrowser where some of the segmented nuclei of (a) are highlighted in magenta. White horizon*tal* bars indicate the approximate anterior boundaries of segments 1, 2, and 3 from top to bottom. Scale bar: 25 µm The right panel shows the traces of the commissural neurons I identified among the segmented nuclei of (a). Each neuron is colored differently. A single ventral plane of the SBF-SEM dataset is visible, but belongs to a slightly more dorsal position to traces so as not to interfere with them.

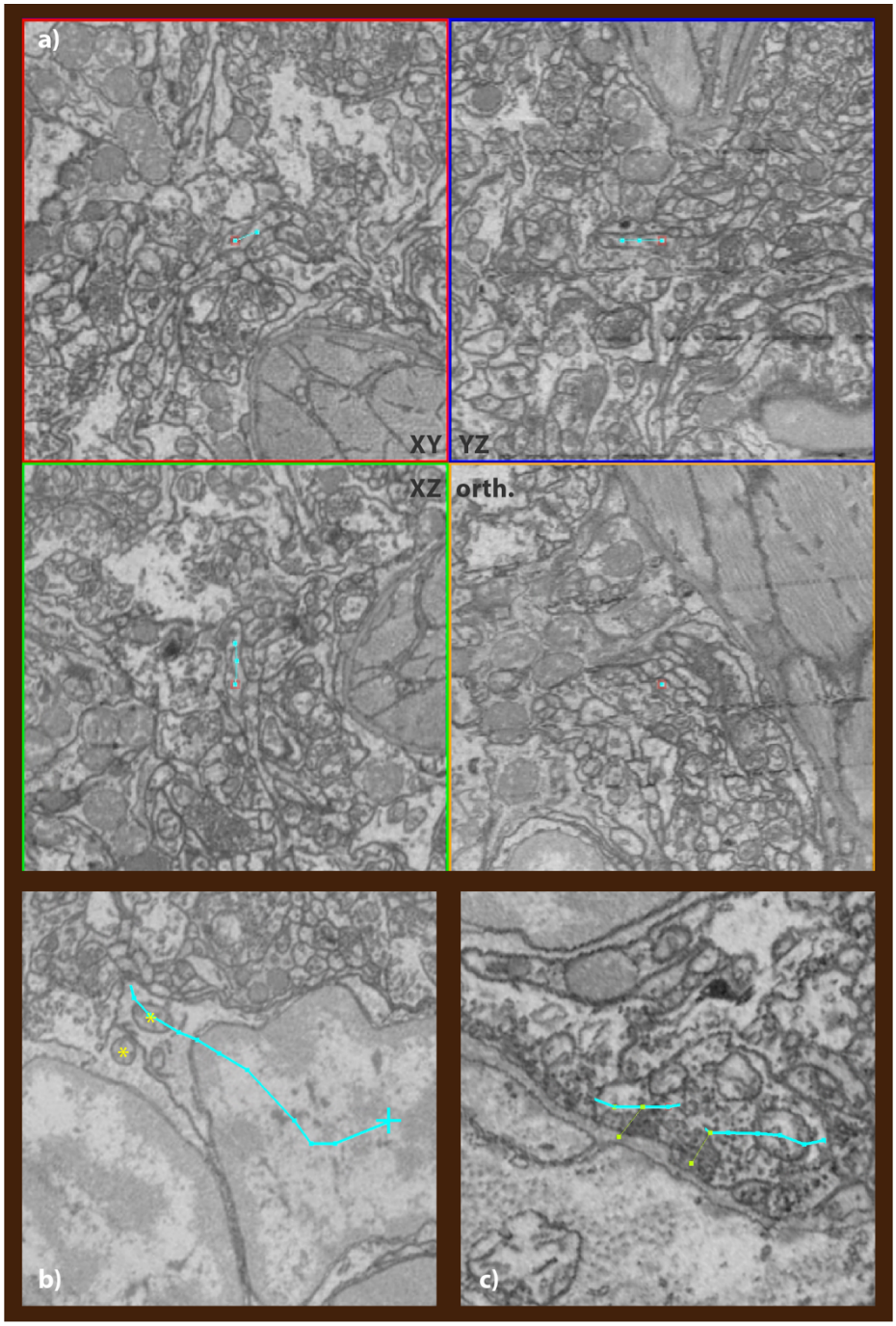


Figure II-18. Tracing neurons using PyKnossos

A) A neuron is traced by placing nodes along the axon (cyan squares). These nodes become automatically connected by a line (cyan line). The four navigation panes of PyKnossos allow one to explore the dataset in the XY, YZ, and XZ viewpoints. It also has a fourth pane that is orthogonal to the line that connects the nodes of a trace. B) The first step to trace a neuron starting from the nucleus is to set a seeding point (cyan cross). One must then search for the axon hillock. Axon hillocks have a characteristic funnel-like structure with a mitochondrion (yellow asterisks mark mitochondrion in neighbouring axon hillocks). C) Synapses could be identified as electron dense condensations adjacent to the axonal membrane. The neuron traced in this panel is a motoneuron, as it is synapsing onto muscle cells (green nodes/lines).

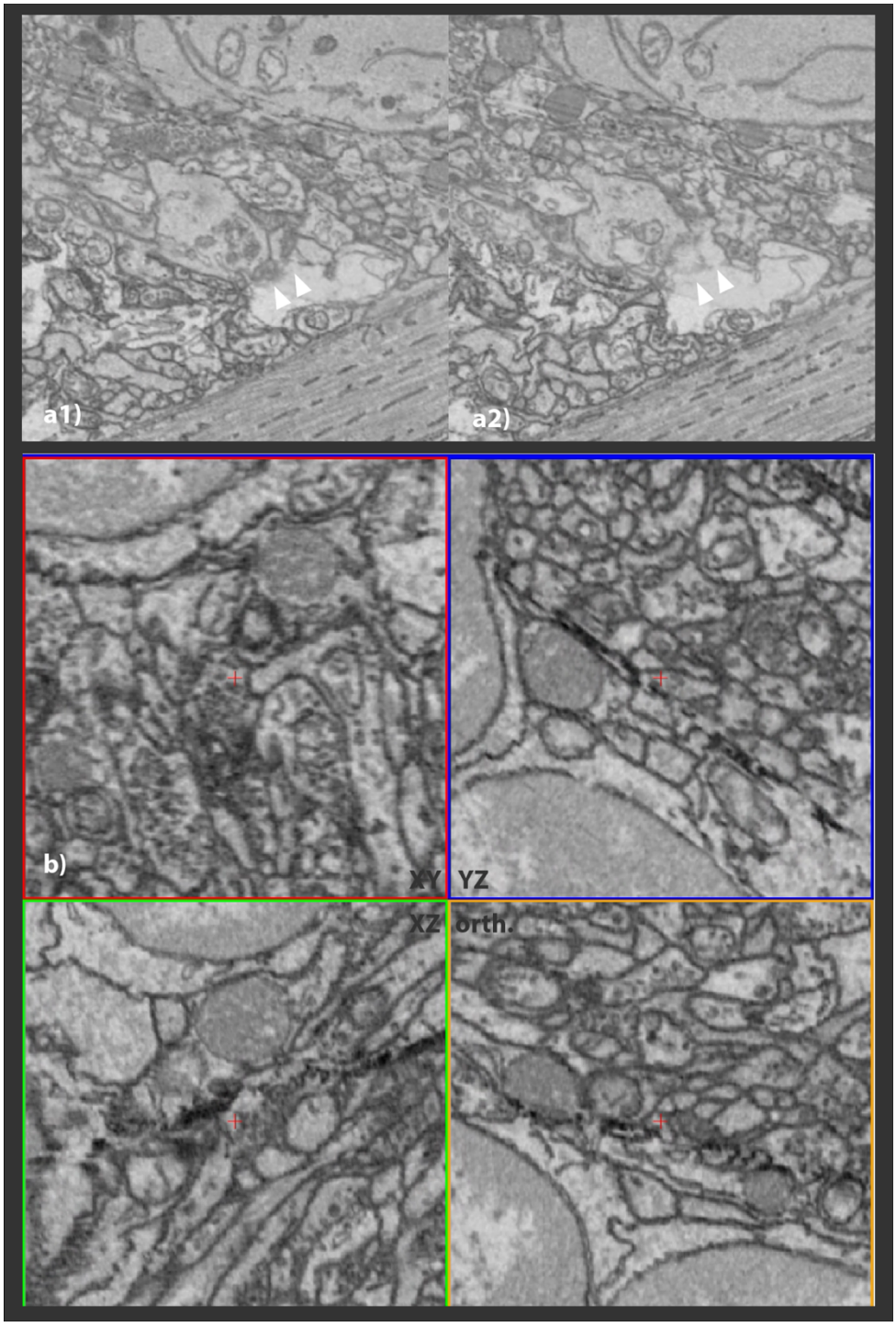


Figure II-19. Artefacts of the *Platynereis* SBF-SEM dataset

a) The same coordinates are shown in two consecutive z-slices where the membrane-fusion artefact can be seen. The white arrowheads point at segments of an axonal membrane that can be distinguished in a1) but become fused in a2). b) A misalignment can be seen in the four viewports of PyKnossos (XY, YZ, XZ, orth.).The red cross in the middle points at the same coordinates of an axon in the four viewports. While an artefact might make it impossible to continue tracing a neuron in one viewport, it might still be possible to continue on a different viewport.

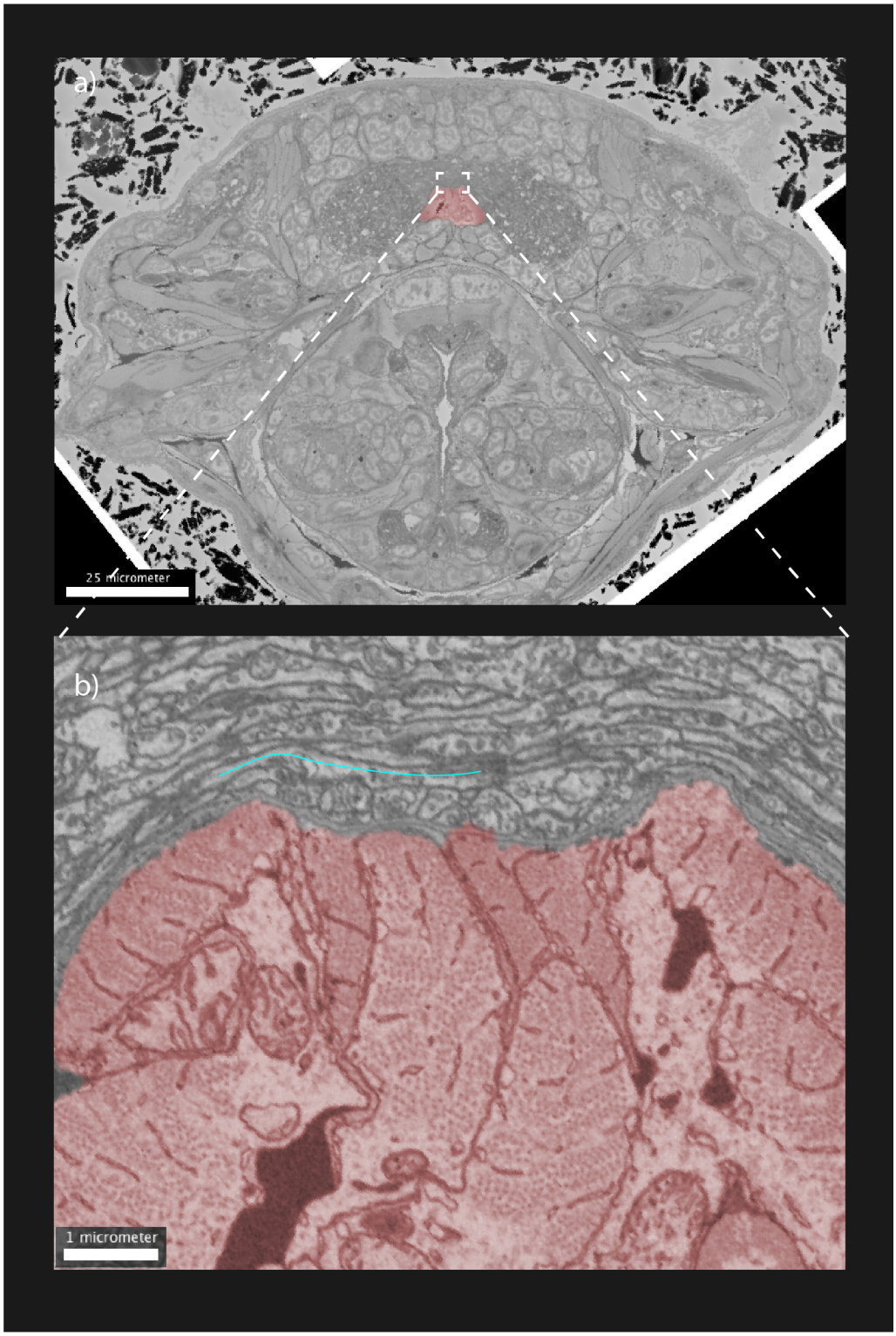


Figure II-20. Commissural neurons can be identified because they cross the midline at the level of the axochord

a) Cross section at the level of the first segment. The axochord[170] is coloured in red. The VNC neuropil is dark grey but has not been coloured, it is just more electron dense. **b)** The neuropil and axochord at higher magnification. A segment of a commissural axon's trace can be seen crossing the midline (turquoise line). Scale bar in panel a: 25 µm and b: 1 µm

3.2 Biased Search For Commissural Neurons Corresponding To The V0 Commissural Cell Type of the Vertebrate Spinal Cord

3.2.1 Developmental Fate Of The Lhx1/5^{+}, *Eve*+ Domain

The V0 cell types of the vertebrate spinal cord belong to three groups: V0d, V0c, and V0v (Table II-2). Our preliminary analysis showed that the TF signature of the V0v neurons is also found in *Platynereis* at 36 hpf, 48 hpf, 72 hpf, and 6 dpf (Figure II-21, Supplementary figure 4, and Supplementary figure 5)

We found a bilateral region in the pygidium that expresses the homeodomain transcription factor *Eve* and the LIM-homeodomain transcription factor *Lhx1/5*, both matkers of the V0v cell type of the vertebrate neural tube. We could detect this coexpression domain as early as 36 hpf (Figure II-21). It is found at 48 hpf and 6 dpf, but not at 72 hpf.

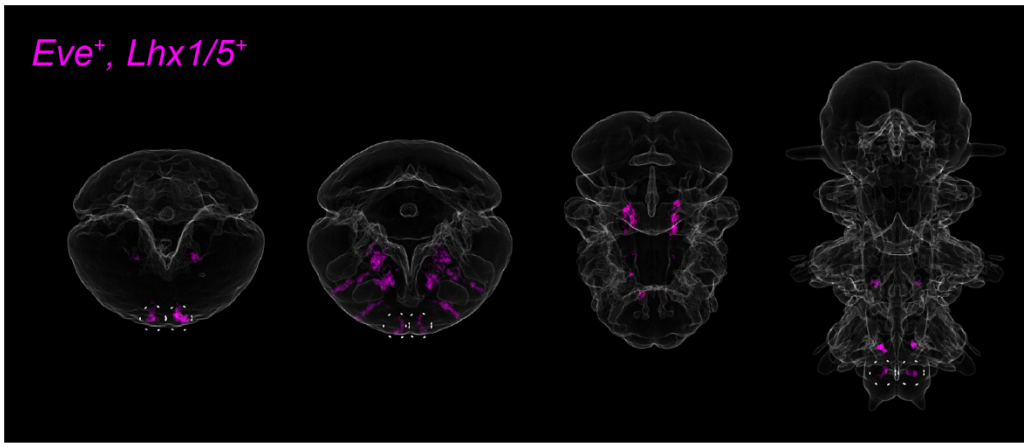


Figure II-21. Putative development of the *Eve*+, *Lhx1/5*$^{+}$ cells
These coexpression panels were generated using the ProSPr atlases of gene expression at 48 hpf (left), 72 hpf (middle), and 6 dpf (right). The magenta text above the panels lists the TFs that are coexpressed. The coexpression region in the pygidium is surrounded by a dashed circle.

3.2.2 Molecular Characterisation And Axonal Tracing Of The *Eve+*, *Lhx1/5* Neurons In The 6 dpf Larva

I followed the same strategy followed in 3.1.1.5 , with the difference that I added the segmented nuclei of the pygidium to this analysis. I found 14 segmented nuclei that coexpress *Eve* and *Lhx1/5*. They all express the homeobox transcription factor *Otp* and the paired-domain homeobox transcription factor *Pax6* as well (Figure II-22a).

Seven are located in the second segment and seven in the pygidium. We can distinguish them based on their expression of *Hox* genes: *Hox*1 in the second segment, and *Hox*4 and *Hox*7 in the pygidium. Based on the EdU staining, we can determine that the cells in the second segment were dividing between the 3rd and the 4th day after fertilisation, whereas the pygidial cells were dividing earlier, between 42 and 48 hpf. They also differ in the remaining TFs that they express.

The cells in the second segment express the forkhead box transcription factor *Fezf*, the non-basic HLH transcription factor *COE*, and some of them express the POU Class 4 homeobox transcription factor *Brn3*. The pygidial cells express the homeobox transcription factor Uncx, the forkhead box Foxn4, the zinc-finger transcription factor Sp8, and the Iroquois homeodomain transcription factor *Irx6*.

With respect to the effector genes, only the neurons in the second segment express the potassium channel kv33b, glutamate synthase (glt1), and choline acetyltransferase (chat). On the other hand, there are many other genes that are only expressed in the pygidial cells, such as the presynaptic protein *Rab3* and the guidance molecule *Slit*.

These differences lead us to hypothesise that the commissural *Lhx1/5*$^{+}$, *Eve+*, *Pax6*$^{+}$, *Otp*$^{+}$ neurons belong to two cell types: one forms earlier in the pygidium and the other forms later in the second segment (also see Results, section 5.2).

All of the *Lhx1/5*$^{+}$*, Eve*$^{+}$*, Pax6*$^{+}$*, Otp*$^{+}$ commissural neurons project anteriorly after crossing the midline (Figure II-22b). This is in contrast to the *Dbx1*$^{+}$*, Pax6*$^{+}$*, Irx6*$^{+}$ commissural neurons of the first segment, which appear to be descending commissural neurons.

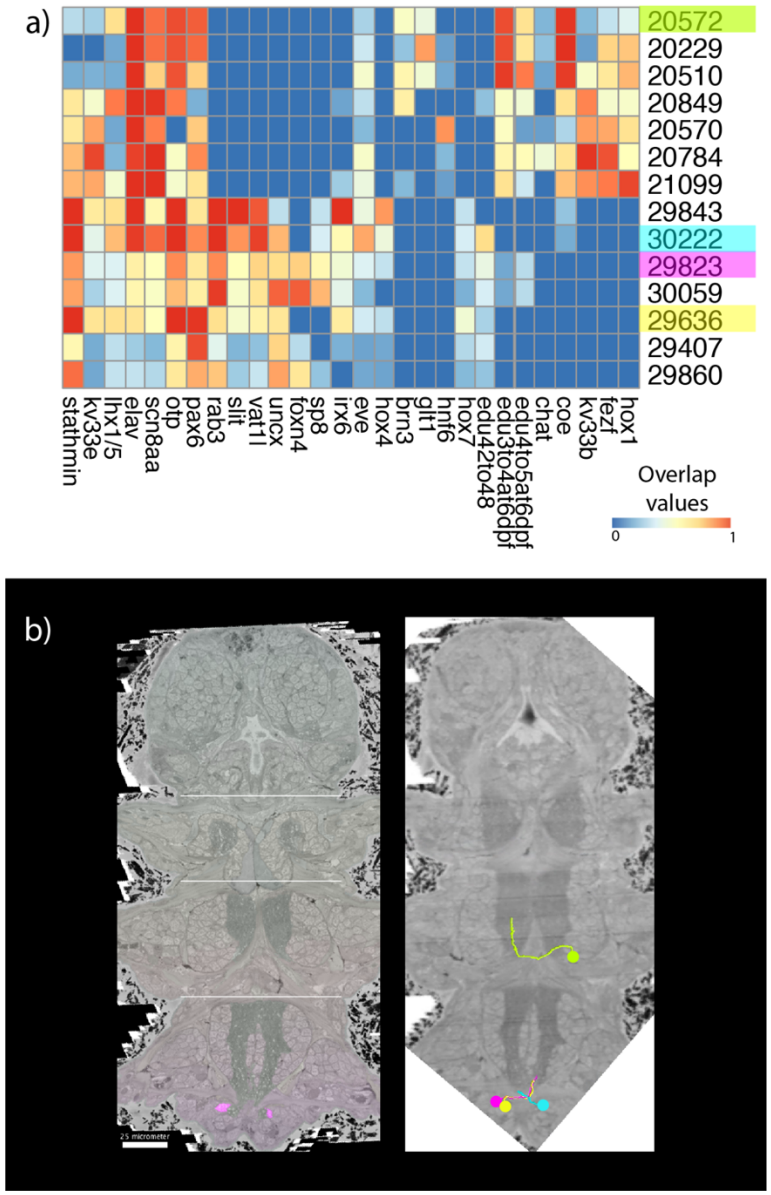


Figure II-22. Segmented nuclei that coexpress *Lhx1/5* and Eve

a) Heatmap of all the segmented nuclei with overlap values (Definition 1) above 0.1 for *Lhx1/5* and Eve. Selected genes are shown in the x-axis, and the y-axis is labelled with the IDs of the segmented nuclei. The labels that are coloured correspond to the commissural neurons whose traces are shown in the right panel of (b).This heatmap was created with the R package 'pheatmaps'. **b)** The left panel show a ventral plane of the PlatyBrowser where some of the segmented nuclei of (a) are highlighted in magenta. White horizon*tal* bars indicate the approximate anterior boundaries of segments 1, 2, and 3 from top to bottom. Scale bar: 25 µm The right panel shows the traces of the commissural neurons I identified among the segmented nuclei of (a). Each neuron is coloured differently. A single ventral plane of the SBF-SEM dataset is visible, but belongs to a slightly more dorsal position to traces so as not to interfere with them. Stathmin is a protein involved in the regulation of the microtubule filament system. Kv33e and kv33b are voltage-gated potassium channels. Elav is a conserved RNA-binding protein with a role in neuronal differentiation. Scn8aa is a sodium voltage-gated channel. Rab3 is a protein involved in synaptic vesicle exocytosis. Glt1 is a glutamate transport that clears glutamate from the synaptic cleft. Edu42to48 marks the nuclei that divided between 42 and 48 hpf, as determined by Edu pulses. Edu3to4at6dpf marks the nuclei that divided between 3 and 4 dpf, as determined by Edu pulses. Edu4to5at6dpf marks the nuclei that divided between 4 and 5 dpf, as determined by Edu pulses. Chat is choline acetyltransferase, an enzyme that catalyses the synthesis of acetylcholine. *Lhx1/5*, *Otp*, *Pax6*, *Uncx*, *Foxn4*, Sp8, *Irx6*, *Eve, Hox*4, *Brn3*, *Hnf6*, *Hox7*, *COE*, *Fezf*, and *Hox*1 are TFs (see text).

4 Unbiased Study Of The Commissural Neurons Of The Second Segment VNC Of A 6dpf Larva Of *P. Dumerilii*

4.1 Identification Of The Commissural Neurons In The Second Segment

The next approach we took to study the commissural neurons of *Platynereis* was to unbiasedly look for them in the second segment. Only through an unbiased approach can we study the totality of the commissural neurons in the second segment. We will investigate their identity through the analysis of their TF-expression profile.

An advantage of using the larva at 6 dpf is that most cells in the VNC have already differentiated by this stage[156]. This allows us to study progenitor cell types in earlier larval stages and differentiated neurons in the 6 dpf larva.

We chose to carry out this study in the second segment because it is most representative of the differentiated neurons of the VNC. On the one hand, the first segment has a very unique TF-expression pattern that suggests that many cell types are unique to this segment[56]. At this stage, the first segment, contrarily to the second and third segments, is not used for crawling; at later stages, over the course of the second metamorphosis, it is transformed into a head segment[155]. On the other hand, the third segment is the last one to differentiate[46] and is therefore the most likely to be still composed of progenitor cells and lacking fully grown commissures.

To identify and trace all the commissural neurons of the second segment I used the tracking software PyKnossos (see Materials and Methods, section IV.13). I started by setting a seeding point for each commissural axon at the location where it crosses the midline. I used the axochord[170] as a reference point for the midline (Figure II-23).

By setting a seeding point for every axon that crossed the midline in the second segment, I was able to identify 167 commissural axons in this segment. These seeding points were then used to trace the corresponding neurons.

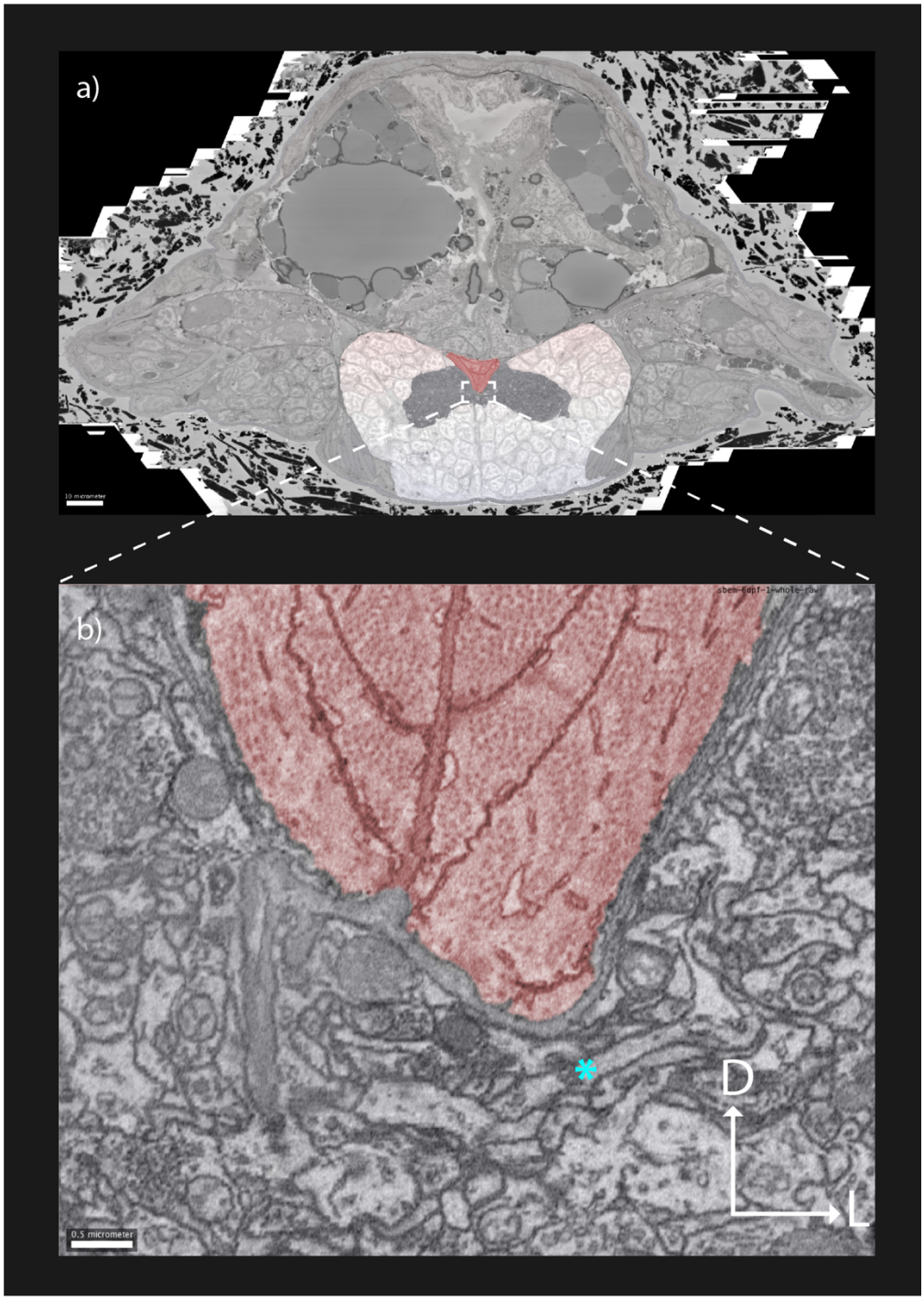


Figure II-23. Setting a seeding point to trace a commissural axon

A) Cross section at the level of the second segment. The VNC soma are coloured in white and the axochord[170] in red. The VNC neuropil is dark grey but has not been coloured, it is just more electron dense. B) The neuropil and axochord at higher magnification. Individual axons can be distinguished. I set a seeding point (cyan asterisk) in every axon of the second segment that I could see crossing from one side of the axochord to the other. (Scale bar in panel a: 10 µm and b: 0.5 um).

4.2 Tracing Of The Commissural Axons in The Second Segment

To trace neurons in PyKnossos, one places nodes along the axon that become connected with a line (Figure II-18a). The PyKnossos software has four navigation panes: XY, YZ, XZ, and a fourth pane that is orthogonal to the line that connects the nodes of a trace (Figure II-18a).

From the midline, axons were traced towards the left and the right until the soma was reached on one side (Figure II-18b), and as far as possible on the other, where most of the post-synaptic partners of the respective neurons are found (Figure II-18c).

Tracing the whole neuron was not always possible because the membranes of more than one axon were sometimes fused and/or ruptured, an artefact that was likely caused during the fixation step (Figure II-19a). The volume dataset also has some alignment problems, but navigating with the 4 view panes usually made it possible to still trace most axons (Figure II-19b). Because of these artefacts, 41/167 cell bodies could not be found.

All of the 125 soma I found were located in the second segment (Figure II-26), meaning that intersegmental commissural neurons project into other segments only after crossing the midline.

On the contralateral side, neurons were traced as far as possible. As far as I could tell, most of the commissural neurons I traced are interneurons. But I could determine through the observation of synapses onto muscle cells that at least 10 of the neurons I traced were commissural motoneurons (Figure II-18c). All the motoneurons that I found also synapsed onto other neurons (data not shown). I did not systematically annotate synapses, given that so many projections remained incomplete towards their dis*tal* ends.

I annotated the axonal projection of these 125 neurons. 66 project laterally, meaning that they remain in the second segment (Figure II-24). 33 project anteriorly, defined here as projecting into the first segment. 20 project posteriorly and project into the third segment. I also found 7 bidirectional axons, which bifurcate and project into the first and third segments.

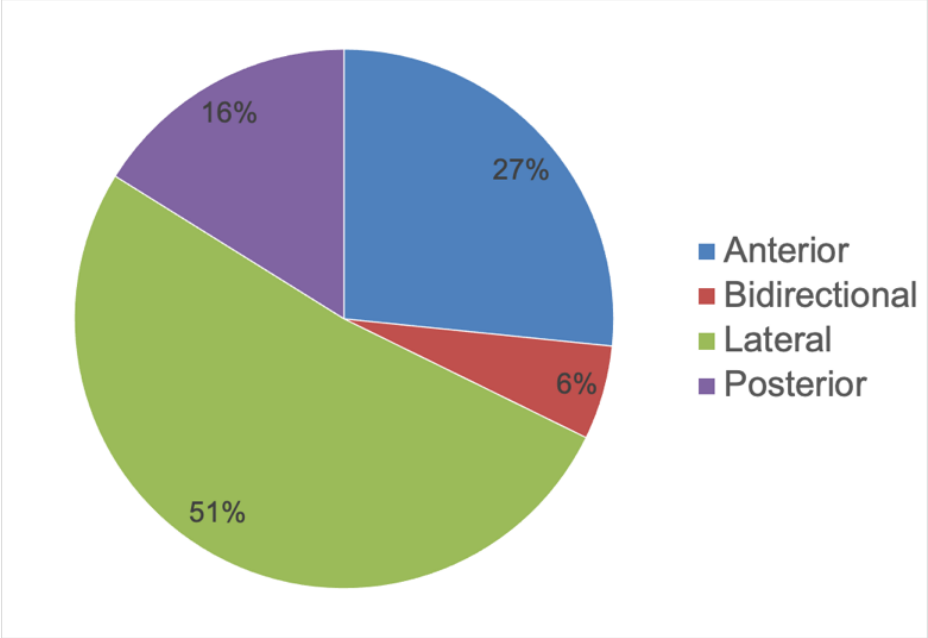


Figure II-24. Axonal projection of the commissural neurons of the second segment
The axonal projection of the 125 commissural neurons of the second segment. Lateral: the whole trace is within one segment. Anterior: crosses the anterior segmental boundary. Posterior: crosses the posterior segmental boundary. Bidirectional: axons bifurcate into an anteriorly- and a posteriorly-projecting segment.

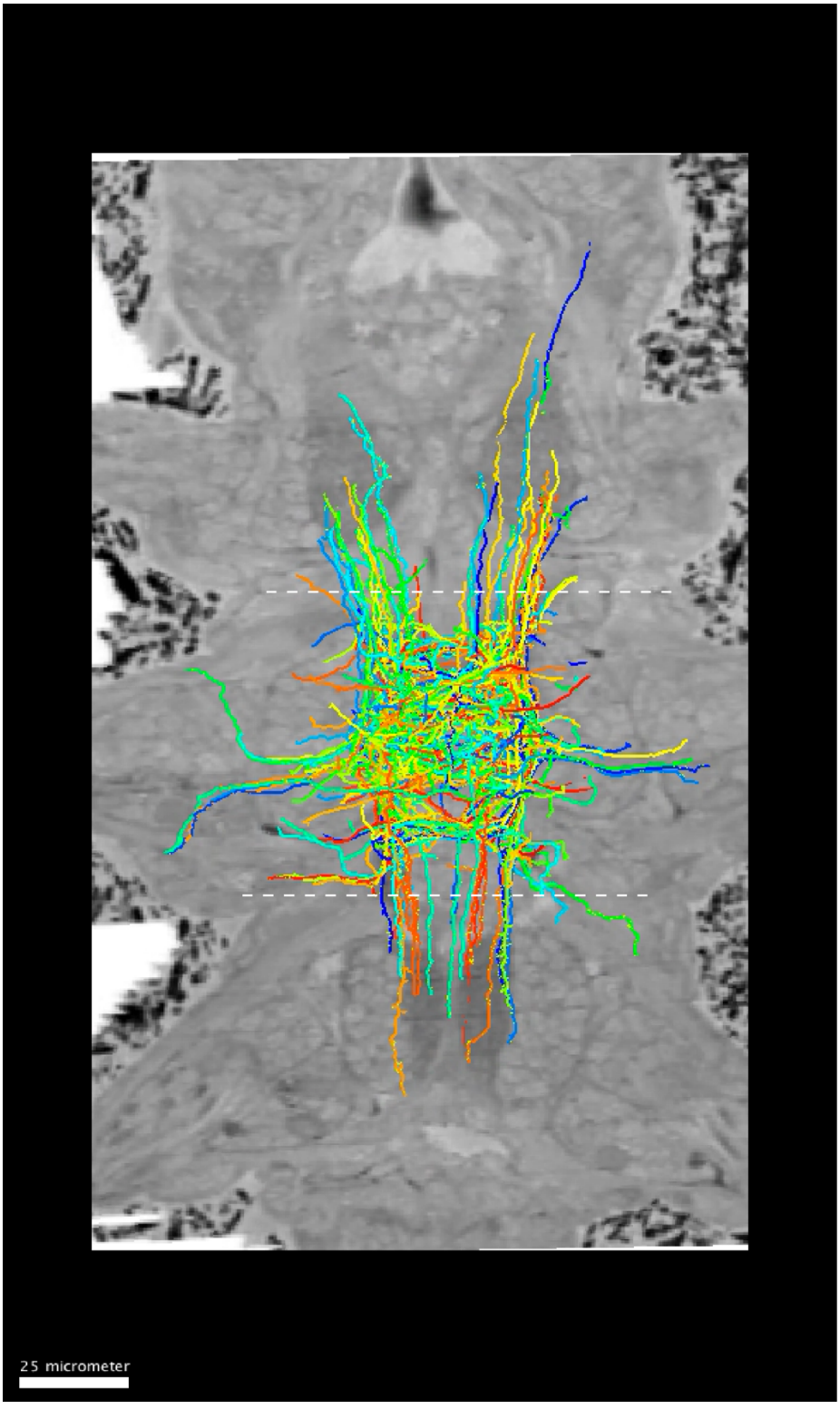


Figure II-25. Commissural neuron traces overlaid on the 6 dpf SBF-SEM dataset

This is a ventral view of the traces of the commissural neurons, each in a different colour. A single plane of the SBF-SEM dataset is visible, but belongs to a slightly more dorsal position so as not to interfere with the commissural traces. The segmental boundaries are indicated with dashed lines.

4.3 Molecular Characterisation Of The Commissural Neurons In The Second Segment

I recovered the coordinates of every soma that belonged to a commissural neuron that I traced on PyKnossos. After transforming them to um, Valentyna Zinchenko retrieved the segmented cells of the PlatyBrowser located at these coordinates. I manually verified that they were the same cells (Figure II-26).

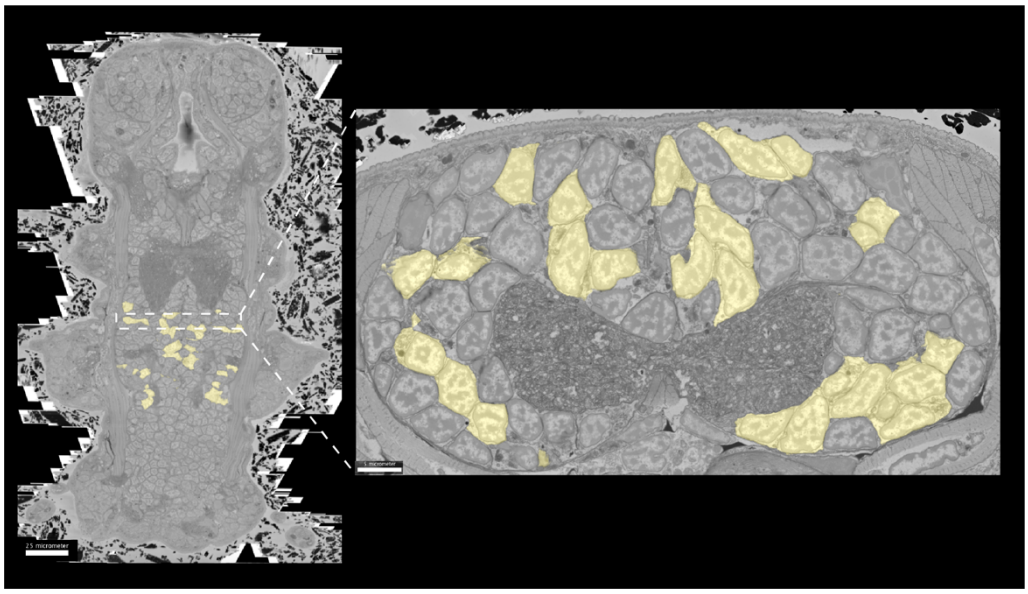


Figure II-26. Segmented soma of the commissural neurons of the second segment
The left panel shows a ventral view of the PlatyBrowser. The right panel is a transversal view at the level of the dashed lines. In both cases, the segmented commissural neurons are highlighted in yellow.

Valentyna then provided me with the overlap values (Definition 1) of the 125 commissural neurons. To analyse these values, I started by making the distinction between broadly- and sparsely-expressed genes, because the distribution of their values is different (Figure II-3). I binarized the overlap values using 0.5 as a threshold for broadly-expressed genes and 0.4 for sparse genes. Overlap values below the threshold were transformed to 0 and I considered them as not being expressed. Overlap values equal or above the threshold were transformed to 1 and I considered them as being expressed.

I found that six genes are expressed by more than 50% of the commissural cells and they are all neuronal markers. Given that no TF from the ProSPr atlas is expressed by every commissural neuron, it is possible that there is no master CoRC that commits neurons to a commissural fate. Alternatively, it is possible that the TFs that make up this hypothetical master CoRC are not part of the ProSPr atlas at this moment.

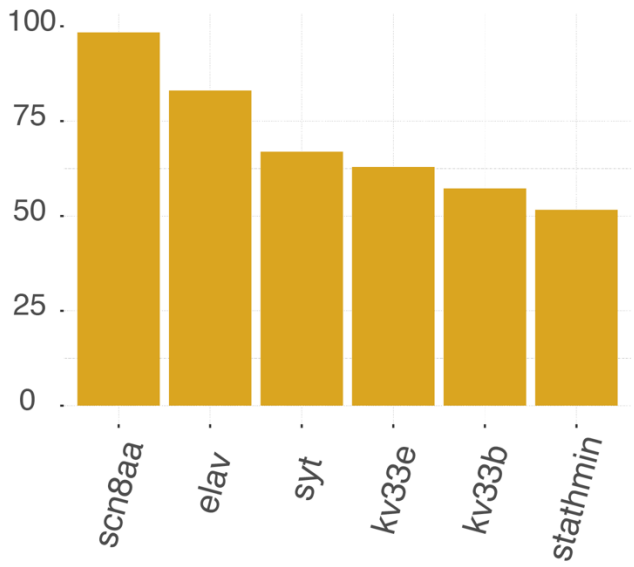


Figure II-27. Genes expressed in >50% of the commissural neurons

This bar plot shows the genes that are expressed in more than half of the segmented nuclei of the commissural neurons. The height of each bar indicates the percentage of cells that express this gene. I decided that a gene was expressed if it its overlap value in the segmented nucleus was equal or higher than 0.4 or 0.5 for sparse and broad genes, respectively. All of these genes are markers of mature neurons. Scn8aa is a sodium voltage-gated channel. *Elav* is a conserved RNA-binding protein with a role in neuronal differentiation. *Syt* is synaptotagmin, a protein involved in the regulation of membrane trafficking at the synapse. Kv33e and kv33b are voltage-gated potassium channels. *Stathmin* is a protein involved in the regulation of the microtubule filament system.

4.4 Second Segment Commissural Neurons Belong To Several Distinct Molecular Types

I used the gene-expression profile of the segmented nuclei to study how the commissural neurons relate to the other neurons of the second segment and of the VNC as a whole. The VNC had been automatically segmented (Introduction, section 6.1.3). I used this is a base to manually select all the nuclei that belong to the second segment VNC (Figure II-28). The VNC has approximately 2100 segmented nuclei, approximately 1200 of which are in the second segment. I used the same threshold on the overlap values as described in section 4.3 of the Results.

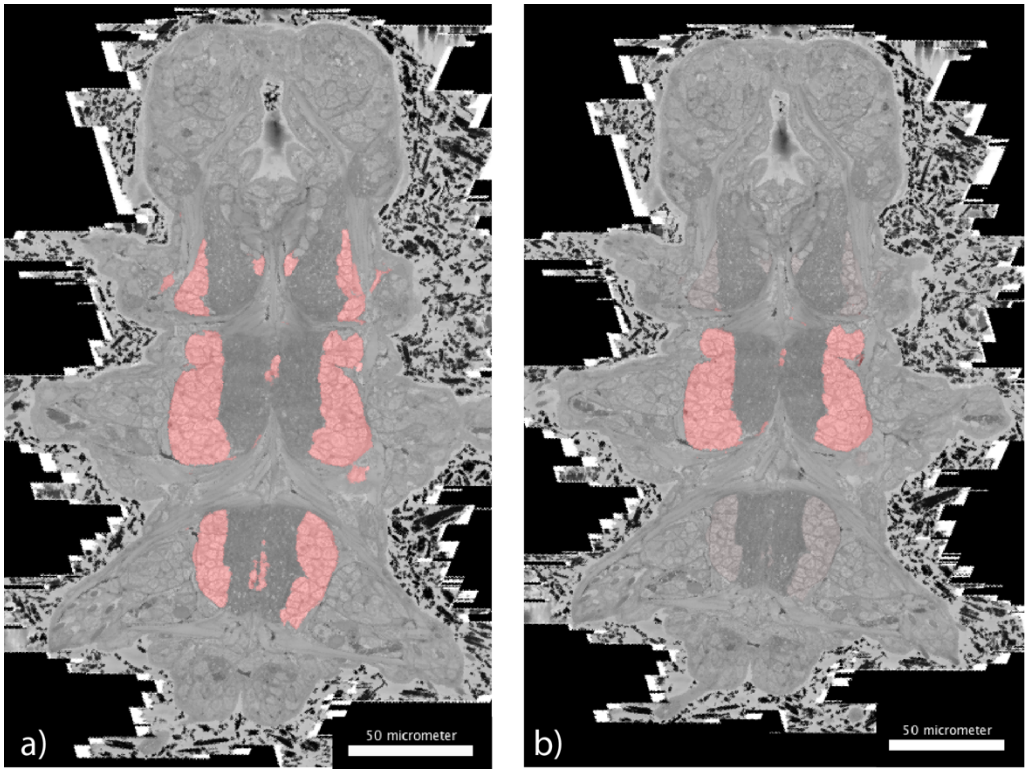


Figure II-28. Segmented cells of the VNC and second segment
Ventral views of the SBF-SEM dataset. The segmented cells of the VNC (a) and second segment (b) are coloured red (see Text). Scalebar: 50 µm

I computed a uniform manifold approximation and projection (UMAP[171]) of the VNC, highlighting the commissural neurons of the second segment (Figure II-29b). I used the R package 'umap' with the default configuration, which uses the Euclidean metric. This package is an R implementation of the Python package 'umap', created by the authors of the algorithm.

For reference, I also show a UMAP of all the segmented nuclei with the VNC highlighted (Figure II-29a). This UMAP was computed by Kimberly Meechan using the overlap in volume between every segmented nucleus and every genes in the PlatyBrowser. The code to generate it is available as a snakemake workflow at: https://github.com/platybrowser/platybrowser-backend/tree/master/analysis/gene_clustering. I coloured the segmented VNC cells black (red in Figure II-28a).

While the VNC forms a distinct group when plotted in a UMAP with all the cells (Figure II-29a, UMAP plot kindly provided by Kimberly Meechan [150]), the commissural neurons do not all cluster together when plotted in a UMAP with only the VNC cells (Figure II-29b).

Upon closer inspection, I realized that while the commissural neurons do not cluster all together, they are also not homogeneously distributed. This leads us to hypothesise that there are several molecularly distinct commissural cell types.

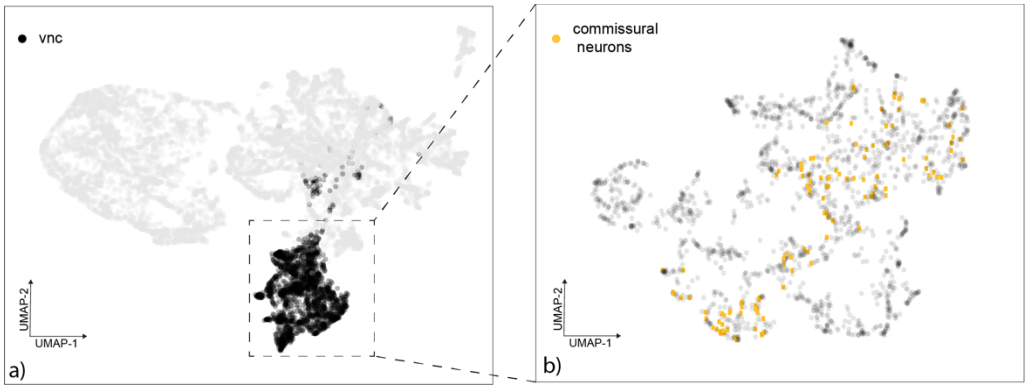


Figure II-29. Overview of commissural neurons of the second segment

Panel a is a UMAP of all of the segmented nuclei of the SBF-SEM dataset, computed by Kimberly Meechan and published in Ref. [150]. It was calculated using the overlap in volume between the segmented nuclei and the expression domains of every gene in the PlatyBrowser. The cells of the VNC, identified manually on the SBF-SEM dataset, are coloured black. Panel b is a UMAP of the cells of the VNC. Since I wanted to identify cell types within the VNC, I calculated this UMAP using the overlap of expression of TFs only. The commissural neurons that I identified are coloured gold. The middle-left panel shows a fron*tal* plane of the SBF-SEM dataset with the anterior towards the top. The segmented commissural neurons are coloured gold. The middle-right panel is a transverse view at the level of the second segment where many of the commissural neurons are visible. The bottom-left panel includes the neuronal traces; the trace of every neuron is represented with a different colour. The bottom-right panel is a bar plot that shows the six genes that were expressed by more than 50% of the commissural neurons. The y-axis is the percentage of cells.

5 Molecular Identity of Commissural Cell Types

Combining the commissural neurons from our biased (Results, section 3) and unbiased (Results, section 4) searches I compiled a catalogue of commissural neurons of *Platynereis dumerilii*.

Our catalogue comprises 133 cells, 125 identified in the unbiased search of the second segment, 5 identified in the first segment from the biased search for p0-like neurons, and 3 identified in the pygidium from the biased search for V0-like neurons.

To analyse how the commissural neurons relate to each other I determined their TF-expression profile. I extracted the overlap between TF expression and the segmented nuclear volume (overlap values) of the commissural neurons using the PlatyBrowser.

To assign gene expression to neurons, I took into account the distinction between broadly- and sparsely-expressed genes (see above) (Figure II-3). I kept only the TFs and binarized the overlap values using 0.5 as a threshold for broadly-expressed genes and 0.4 for sparse genes. Overlap values below the threshold were transformed to 0. Overlap values equal or above the threshold were transformed to 1.

I used a variety of methods for this analysis, including distance-based methods like hierarchical clustering and UMAP plots, as well as parsimony-based methods. Distance-based methods have the advantage of being quick and I was able to try several of them in my preliminary analysis (Supplementary figure 7 -Supplementary figure 13).

The paired-domain homeobox transcription factor *Phox2* was found to define the dorsal region of the VNC, called dorsal ganglionic mass (DGM). *Pax6*, also a paired-domain homeobox transcription factor, defines in combination with other TFs the different regions of the ventral region of the VNC, the ventral ganglionic mass (VGM). The NK homeobox transcription factor *Nk6* defines the medial-most region of the VGM[56].

The commissural neurons in the second segment are, therefore, primarily distinguished by their regional identity, belonging to either the *Phox2*+, *Pax6*$^{+}$, or *Nk6*+ subregions of the nerve cord.

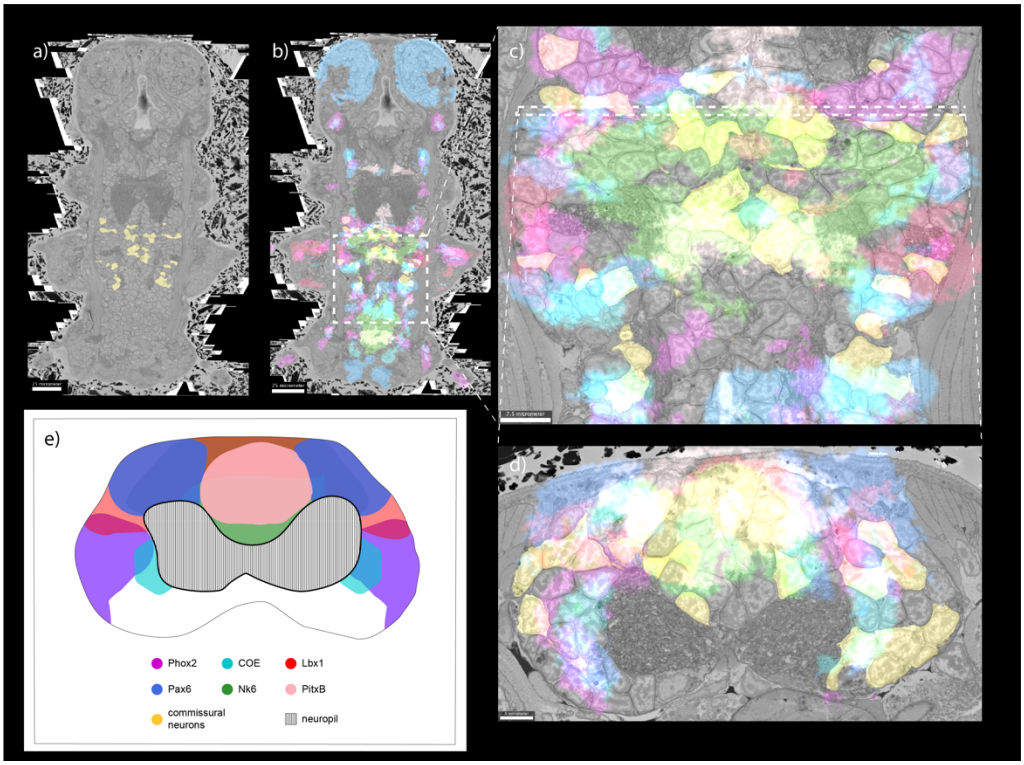


Figure II-30. TFs driving the clustering of the commissural neurons

a) Ventral view of the *Platynereis* SBF-SEM dataset with the segmented commissural neurons highlighted in yellow. b) I overlapped the gene-expression patterns on *Phox2*, *COE*, *Lbx1*, *Pax6*, *Nk6*, and *PitxB*. These TFs were identified as driving much of the clustering of the commissural neurons.c) This inlay zooms in the second segment to appreciate better the gene-expression patterns. The most medially-expressed TF is *PitxB*, followed by *Nk6*. *Lbx1*, *Pax6*, *COE*, and *Phox2* are expressed more laterally. d) A cross-section of c) is shown to visualize the dorso-ventral differences in gene expression. e) Summary diagram of the expression patterns at this level of the VNC in cross-section. Scale bars: 25 µm for a) and b), 7.5 µm for c, and 5 µm for d.

Distance-based methods are quick because they reduce the dissimilarity between cells to a single number, losing the information about the TFs that they express. The clusters that are found through these methods have no information about the TFs that the cells express; for this, one must return to the original data.

Parsimony-based methods keep the TF information throughout the whole process. Briefly, instead of calculating a distance between two cells, trees are built where each cell occupies an endpoint (or leaf) of the tree. The minimum number of changes in TF expression that is required to explain a given tree is calculated; we call this the score of a tree. The cells are then reordered into new leaf positions and the score for the new tree is calculated. The tree with the lowest score is said to be the most parsimonious.

The current consensus favours this tree as the best hypothesis about the relatedness of the cells, because it minimizes their evolutionary distance. Parsimony-based methods also allow one to reconstruct the characters of ancestors at any place of the tree, something that is not possible with distance-based methods. But since the number

of trees that can be built even with a modest number of taxa is very big, parsimony-based methods are very slow and are often unable to assess every possible tree.

After preliminary explorations of our data with distance-based methods, I proceeded to switch to parsimony-based methods (see Materials & Methods, section IV.15). Given the number of cells that make up our dataset, it was impossible to make an exhaustive assessment of every possible tree. Instead, I performed a heuristic search of the most parsimonious tree and found more than one tree with the lowest score. I built a consensus tree with them and used it in the subsequent analysis (Figure II-31).

I followed a cladistic approach to interrogate the phylogenetic affinities of the 133 commissural neurons for which I identified a segmented nucleus and for which we had overlap values of gene expression.

The characters I used to build these trees are the TFs of ProSPr (Supplementary table 2). The character states that I used were: not expressed (0), unknown (?), expressed (1). I used the TF-expression overlap values to attribute a state to each TF (see Materials and Methods, section IV.15 for more details).

I designed a heuristic search for the most parsimonious tree (see Materials and Methods, section IV.15), and kept 10000 of the best-scoring trees to build a consensus tree. The consensus tree only includes clusters that appear in more than 50% of the trees (Figure II-31).

The differences that we found from one method to another were minor, and this allowed us to build a general picture of the groups into which our cells clustered. We identified the paired-domain homeobox transcription factors *Phox2*, *Pax6*, and *PitxB*, the NK homeobox factors *Nk6* and *Lbx1*, and the *bHLH* transcription factor *COE* as TFs that seem to drive much of the clustering. Interestingly, many of these TFs had been identified as major regional specification genes of the 6 dpf *Platynereis* VNC[56].

We found 7 clusters that we believe are cell types and are well supported and positively defined by their TF-expression profile (Figure II-49). They belong to any of three groups. The first group is defined by the expression of *Phox2* and *Hox*, and contains two cell types: one expresses *Brn3* and the other *Isl*.

The second group is defined by *COE* and *Pax6* and contains two cell types: one expresses *Lhx1/5* and *Eve,* and the other *Dbx1*. These two cell types are made up of the cells that we identified during our biased approach. Refer to the section 1 of the Results for an analysis of these cell types.

The third group is defined by *Nk6* and *PitxB* and contains three cell types: the first expresses *Maf*, the second *Tal,* and the third *Irx6*.

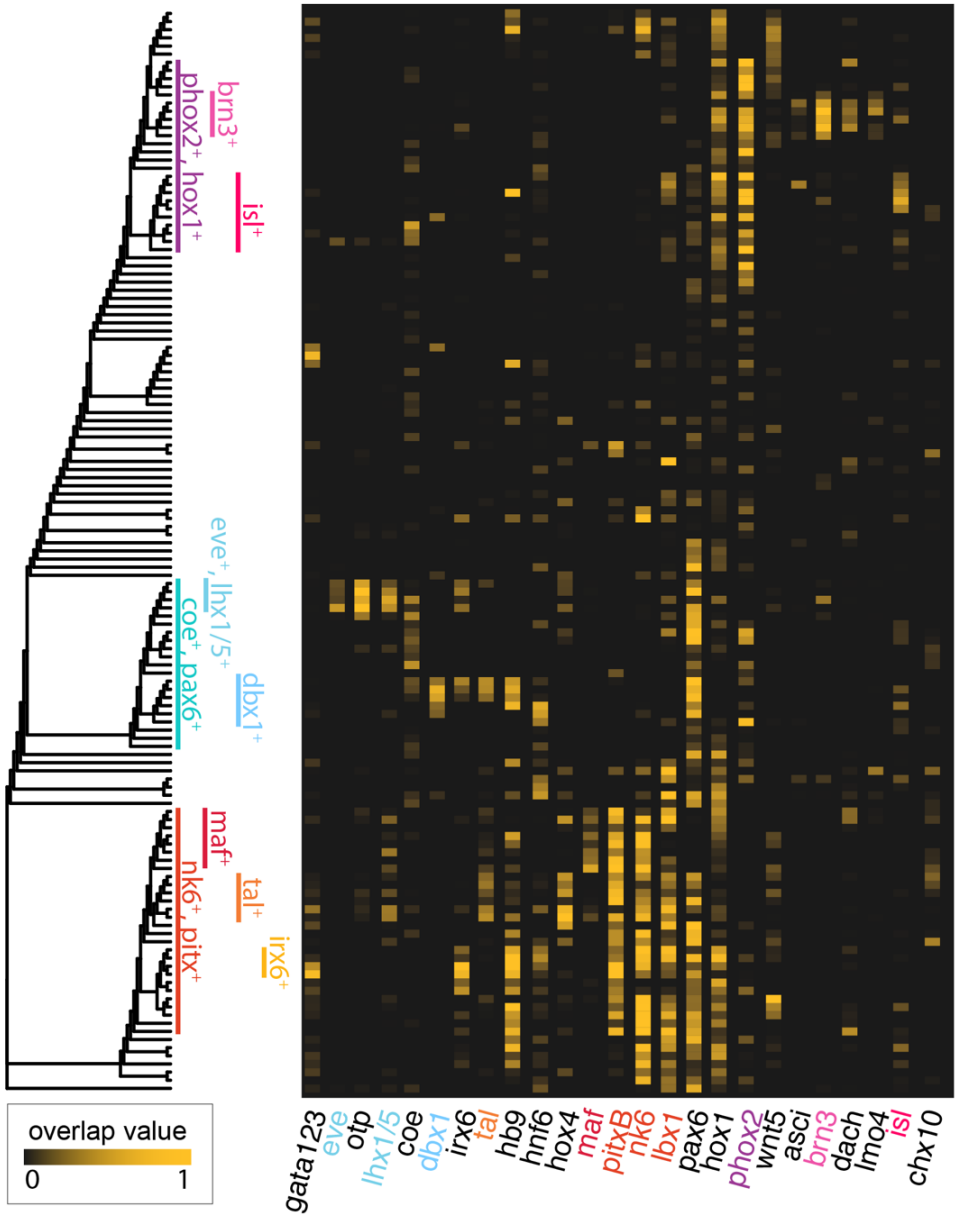


Figure II-31. Maximum-parsimony consensus tree of all commissural neurons and a heatmap of TF overlap values

This is the maximum-parsimony consensus tree of the 133 commissural neurons for which I identified a segmented nucleus. The colour scale shows the raw overlap values and not the character state values that were used for the parsimony tree. Based on their TF-expression profile we identified 7 cell types in 3 groups. The first group consists of a paraphyletic cluster that expresses *Phox2* and *Hox*1. We found two cell types within this group: the first expresses *Brn3* and the second *Isl*. The second group is a monophyletic cluster that expresses *COE* and *Pax6*. We found two cell types within this group: one expresses *Eve* and *Lhx1/5* and the other expresses *Dbx1*. These two clusters consist of the exact same cells that I identified in our biased search and are shown in sections 3.1.1.5 3.2.2 of the Results. The third group is a monophyletic cluster that expresses *Nk6* and *PitxB*. We identified three cell types within this group: the first is characterised by the expression of *Maf*, the second by *Tal,* and the third by *Irx6*.

5.1 *Phox2*, *Hox1* Group

This group consists of 27 neurons that express *Phox2* and *Hox*1, but do not share the expression of many more TFs (Figure II-31). I highlighted the segmented cells of this group in the PlatyBrowser and overlaid the expression domains of *Phox2* (Figure II-32). The soma are mostly distributed in the lateral-dorsal region of the VNC, where *Phox2* is expressed. They are all in the second segment.

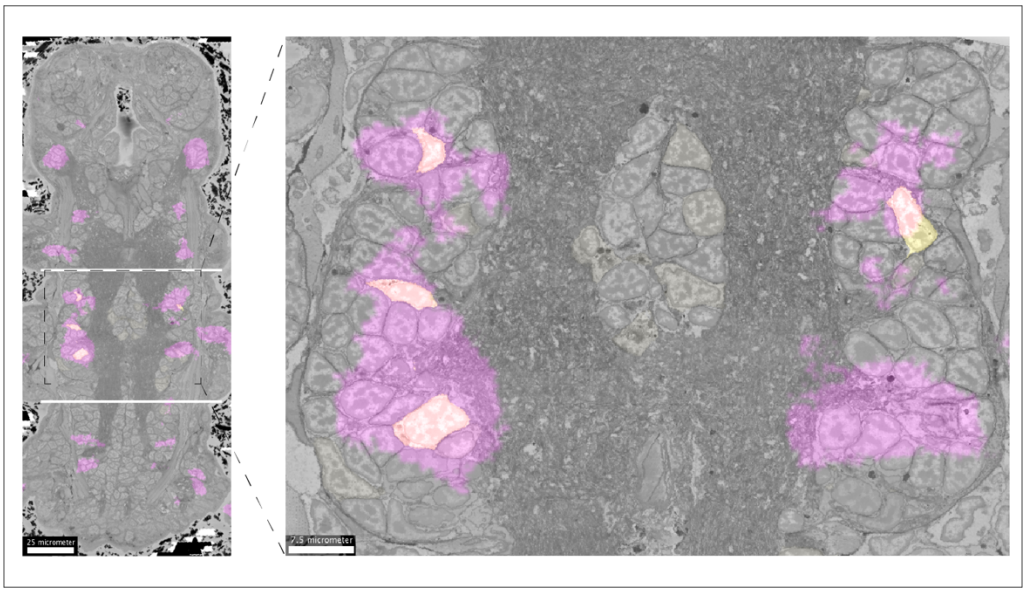


Figure II-32. Segmented cells of cells of the *Phox2*+, *Hox*1+ group of commissural neurons
Ventral views of the SBF-SEM dataset. The expression pattern of *Phox2* is shown in violet, and the segmented cells of the commissural neurons of the *Phox2*+, *Hox*1^{+} group are highlighted in yellow. Scale bar 25 µm for left panel and 7.5 µm for right panel. I added two horizon*tal* bars to the left panel to delineate the second segment.

To further characterise these cells molecularly, I extracted the overlap of effector gene expression for every segmented nucleus of this group (overlap values) using the PlatyBrowser. Again, I distinguished between broadly- and sparsely-expressed effector genes (Figure II-3). Finally, I binarized the overlap values using 0.5 as a threshold for broadly-expressed genes and 0.4 for sparse genes. I considered overlap values below the threshold to indicate that this gene is not expressed in this segmented nucleus. I considered overlap values equal or above the threshold to indicate that this gene is expressed in this segmented nucleus.

They are mature, based on their expression of the sodium and potassium voltage-gated channels Scn8aa, Kv33b, Kv33e, and Hcn1. More than half of them are cholinergic. The *Brn3*+ cell type that we identified in this group has a lower proportion of cholinergic neurons, totalling about 25% (data not shown). It does not vary with respect to the other effector genes. The *Isl*$^{+}$ cell type does not differ in the expression of effector genes from the rest of the neurons in the *Phox2*+*, Hox1*+ group.

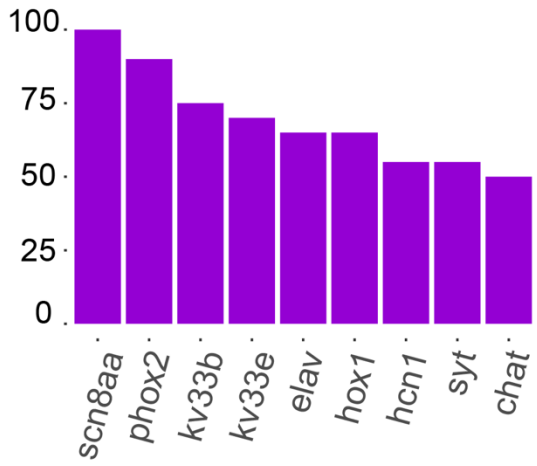


Figure II-33. Genes expressed in more than half of the neurons of the *Phox2*, *Hox*1 group
This bar plot shows the genes that are expressed in more than half of the segmented nuclei of the *Phox2*+, *Hox*1+ group of commissural neurons. The height of each bar indicates the percentage of cells that express this gene. I decided that a gene was expressed if it its overlap value in the segmented nucleus was equal or higher than 0.4 or 0.5 for sparse and broad genes, respectively. Many of these genes are markers of differentiated neurons. Scn8aa is a sodium voltage-gated channel. Kv33e and kv33b are voltage-gated potassium channels. Elav is a conserved RNA-binding protein with a role in neuronal differentiation. Hcn1 is a hyperpolarization-activated potassium ion channel. Syt is synaptotagmin, a protein involved in the regulation of membrane trafficking at the synapse. Chat is choline acetyltransferase, an enzyme that catalyses the synthesis of acetylcholine.

I classified more than half of these commissural axons as projecting laterally, meaning that the whole trace is within one segment (Table II-4). Most of the remaining axons project anteriorly, meaning that they cross the anterior segmental boundary. Many of the ascending axons belong to the *Brn3*+ and *Isl*+ cell types (Figure II-34).

A minority projects posteriorly, crossing the posterior segmental boundary. I also observed bidirectionally-projecting axons that bifurcate into an anteriorly- and a posteriorly-projecting segment (Table II-4).

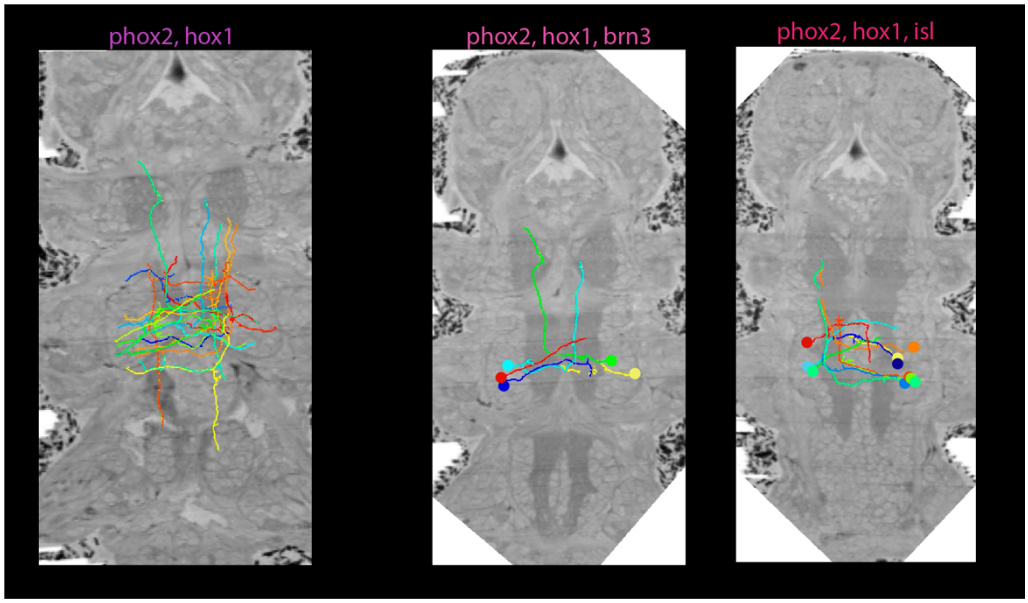


Figure II-34. Traces of the *Phox2*+, *Hox*1^{+} group and its cell types

Ventral view of the traces of the commissural neurons of the *Phox2*+, *Hox*1^{+} group. The left panel shows the traces of the 27 neurons that make up this group with coloured lines. The middle panel shows the traces of 5 neurons that make up the *Phox2*+, *Hox*1^{+}, *Brn3*$^{+}$ cell type. The right panel shows the traces of the 10 neurons that make up the *Phox2*+, *Hox*1^{+}, isl^{+} cell type. The soma positions are indicated with coloured circles in the middle and left panels. In each panel a single plane of the SBF-SEM dataset is visible, but belongs to a slightly more dorsal position so as not to interfere with the commissural traces.

I caution that I was not able to trace every neuron completely because of the membrane-fusion artefact (Figure II-19a). There could be neurons that I did not trace far enough to realize that they ascend or descend. In this case they would be wrongly classified as lateral.

Table II-4 Axonal projection of the commissural neurons of the *Phox2*+, *Hox*1^{+} group

The axonal projection of the 27 commissural neurons that make up the *Phox2*, *Hox*1 group is summarized in this table. Lateral: the whole trace is within one segment. Anterior: crosses the anterior segmental boundary. Posterior: crosses the posterior segmental boundary. Bidirectional: axons bifurcate into an anteriorly- and a posteriorly-projecting segment.

Anterior	10
Bidirectional	2
Lateral	13
Posterior	2

We were able to identify two smaller clusters within this group that we hypothesise to represent cell types.

5.1.1 THE *PHOX2*+, *HOX*1^{+}, *BRN3*$^{+}$ ASCENDING COMMISSURAL NEURONS

This group consists of 5 neurons (Figure II-34). Two of them (green and cyan) look like bilateral pairs, given the similar location of their soma and axonal projection. They both cross the midline and project anteriorly into the first segment, possibly into the head. Their axons run medially in the ventral nerve cord neuropil. The remaining three axons could not be traced very far.

The TFs *Dach*, *AscI*, and *Lmo4* are additional markers. Using the PlatyBrowser I searched for every segmented nucleus that coexpresses *Phox2*, *Hox*1, *Brn3*, *Dach*, *AscI*, and *Lmo4*. When plotted in a UMAP, we find them in 2 clusters (Figure II-35): one in the VNC and one in the lateral ectoderm (Supplementary figure 14). The neurons in the VNC correspond to the commissural neurons that I traced. The cells in the lateral ectoderm are part of the peripheral nervous system.

Figure II-35. UMAP of all cells highlighting the cells that express *Phox2*, *Hox*1, *Brn3*, *Dach*, asci, and lmo4

UMAP of all of the segmented nuclei of the SBF-SEM dataset, computed by Kimberly Meechan and published in Ref. [150]. It was calculated using the overlap in volume between the segmented nuclei and the expression domains of every gene in the PlatyBrowser. I coloured the nuclei that express *Phox2*, *Hox*1, *Brn3*, *Dach*, *AscI*, and *Lmo4*. I decided that a gene was expressed if it its overlap value in the segmented nucleus was equal or higher than 0.4 or 0.5 for sparse and broad genes, respectively.

5.1.2 THE *PHOX2*+, *HOX1*$^{+}$, *ISL*$^{+}$ ASCENDING COMMISSURAL NEURONS

This group consists of 10 neurons. Most of them project anteriorly after crossing the midline, but one (red in Figure II-34) projects posteriorly. Many of these neurons could not be traced very far and it is difficult to establish based on their projection whether they might represent bilaterally symmetric pairs of neurons.

Lbx1 and *Hnf6* are additional markers of this cell type. Using the PlatyBrowser I searched for every segmented nucleus that coexpresses *Phox2*, *Hox*1, *Isl*, *Lbx1*, and *Hnf6*. When plotted in a UMAP, we find them in 3 clusters (Figure II-36): two in the VNC and one in the lateral ectoderm (Supplementary figure 14).

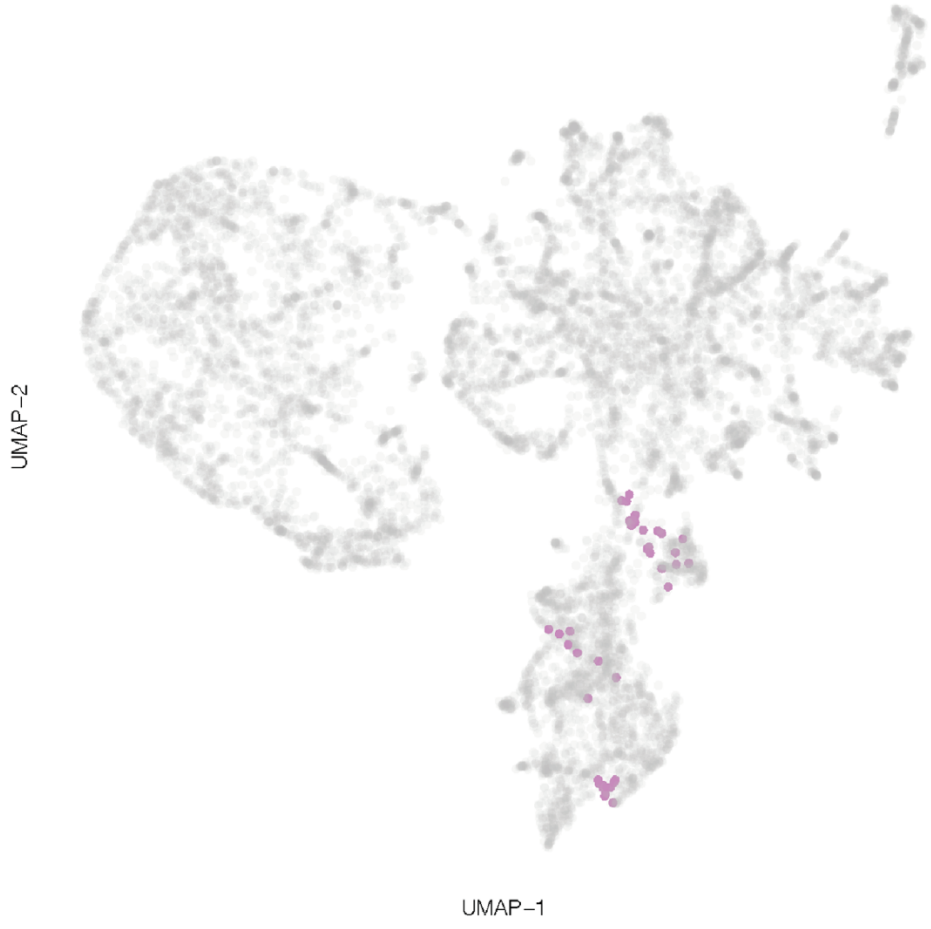


Figure II-36. UMAP of all cells highlighting the cells that express *Phox2*, *Hox*1, Isl, *Lbx1*, and *Hnf6*

UMAP of all of the segmented nuclei of the SBF-SEM dataset, computed by Kimberly Meechan and published in Ref. [150]. It was calculated using the overlap in volume between the segmented nuclei and the expression domains of every gene in the PlatyBrowser. I coloured the nuclei that express *Phox2*, *Hox*1, Isl, *Lbx1*, and *Hnf6*. I decided that a gene was expressed if it its overlap value in the segmented nucleus was equal or higher than 0.4 or 0.5 for sparse and broad genes, respectively.

5.1.3 Developmental Origin Of The *Brn3*$^+$ And *Isl*$^+$ Commissural Cell Types Of The *Phox2*+, *Hox*1$^+$ Group

I used the ProSPr atlases to explore the possible development of these two cell types at earlier larval stages (Figure II-37). I found a region at 48 hpf where *Phox2*, *Hox*1, *Brn3*, and *Isl* are coexpressed. It is within the lateral expression domain of the homeodomain transcription factor *Msx* (Figure I-6) of the mediolateral patterning of the VNC. A subset of them also expresses the *bHLH* transcription factor *Neurogenin* (*Ngn*), suggesting that at least some of these cells are neuronal progenitors at 48 hpf (Figure II-37).

Regions with the same expression domain can be found at 72 hpf, but they no longer express *Ngn*. They are also located in the lateral VNC and are plausibly the daughters of the progenitors at 48 hpf.

At 6 dpf, we find cells in the VNC and in the lateral ectoderm with the full TF signature of each of the two identified cell types (Figure II-37). The commissural neurons that I identified correspond to the neurons in the VNC. We suspect that the cells in the lateral ectoderm are neurons of the peripheral nervous system (PNS), since *Msx* has been shown to specify neural progenitors of both the CNS and PNS across bilaterians[172].

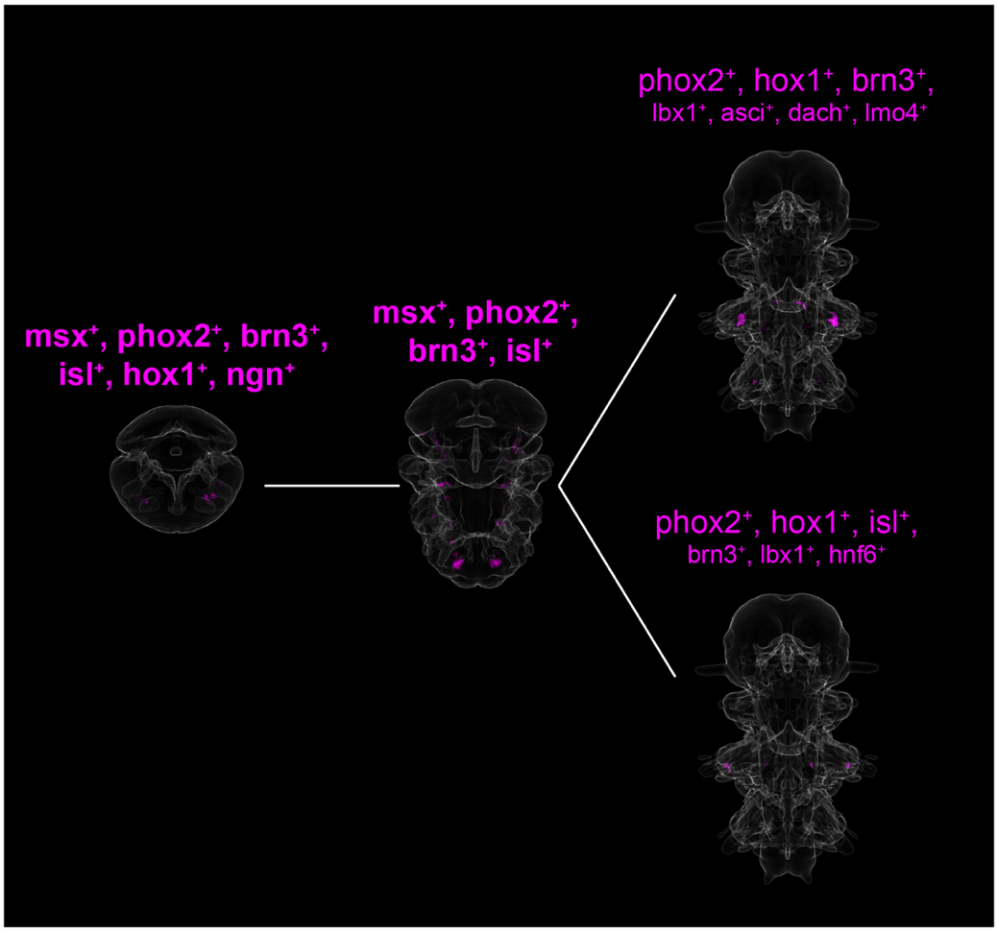


Figure II-37. *Brn3*⁺ and *Isl*⁺ commissural and peripheral cell types at 3 stages of larval development

These coexpression panels were generated using the ProSPr atlases of gene expression at 48 hpf (left), 72 hpf (middle), and 6 dpf (right). The text above the 48 hpf and 72 hpf panels lists the TFs that are coexpressed. The text above the panels at 6 dpf shows the coexpressed TFs in the first line. The second line lists the TFs that are additionally expressed in the corresponding commissural cell types.

5.1.4 *Phox2* Is Expressed In Commissural Neurons At 48 Hpf

I also combined a WMISH against *Phox2* with immunostaining against acetylated tubulin (see Materials and Methods, section IV.8). I determined that *Phox2*+ neurons at 48 hpf project commissural axons across the midline (Figure II-38). We cannot be certain, nevertheless, that these are the same neurons that are also *Brn3*⁺ and *Isl*⁺. A subset of the *Brn3*+, *Isl*+ cells might be neuronal progenitors based on their expression of *Ngn*. Yet 48 hpf is a time of intense differentiation in the VNC and the differentiation state of the trunk is not homogeneous in the antero-posterior axis (see Introduction, section I.3).

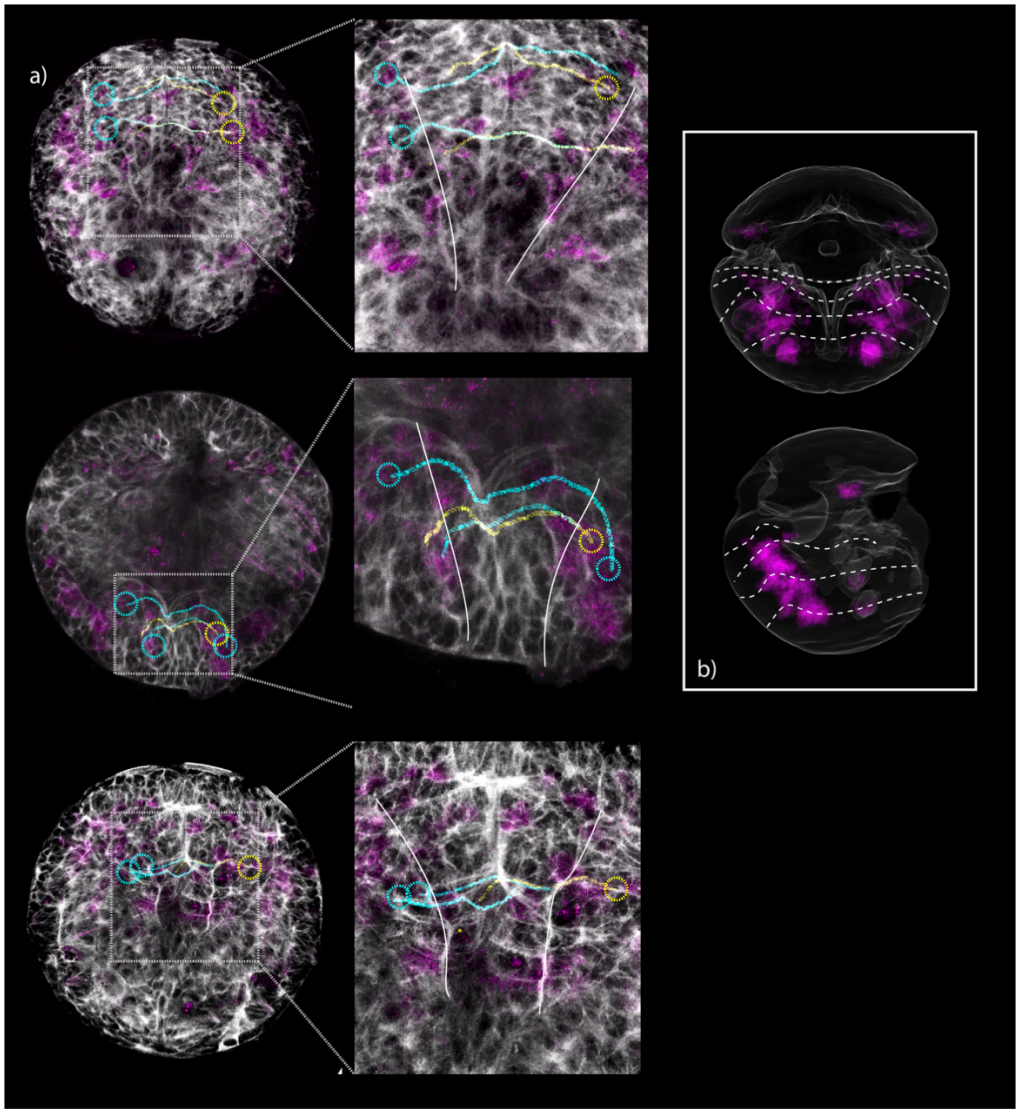


Figure II-38. Commissural axons of *Phox2*+ neurons at 48 hpf

a) Three larvae at 48 hpf that underwent a WMISH against *Phox2* and an immunostaining against acetylated tubulin. In each case, a panel zooming in on the midline is shown where two white lines delineate the nerve cords. The signal from the WMISH is in magenta and from the acetylated tubulin in grey. I traced the axons of the neurons that express *Phox2* (see Materials and Methods, section IV.10) and found that some of them were commissural. The traces of these commissural axons are shown in cyan or yellow. The dashed circles indicate the location of the somas of the traced axons. b) ProSPr expression panel of *Phox2* at 48 hpf. The dashed lines indicate the anterior boundaries of segments 1, 2, 3 and 4 (see Materials and Methods, section IV.14).

5.2 *COE*, *Pax6* Group

We identified two cell types in this group: one characterised by the expression of *Dbx1* and one by *Eve* and *Lhx1/5*. They correspond to the *Pax6*$^{+}$, *Dbx1*$^{+}$, *Irx6*$^{+}$ cells of section 3.1.1.5 and to the *Eve*+, *Lhx1/5*+ neurons of section 3.2.2, respectively. Please refer to these sections for more results concerning these cells. In this section I will only refer to what we learn about these cell types from the cell type tree of Figure II-31.

The tree reveals that these two cell types are more closely related to each other than to most other commissural neurons. All of my attempts to relate them to a common progenitor domain failed (data not shown). Rather, I believe that they rise from two different progenitor domains, as described in Figure II-16 and Figure II-21.

We hypothesised in section 3.2.2 that the *Lhx1/5*+, *Eve*+ commissural neurons that we identified belong to two different cell types, one in the pygidium and the other one in the second segment. Nevertheless, we find in the cell type tree that they all cluster together. If they do belong to two different cell types, they must be very closely related cell types.

5.3 *Nk6*, *PitxB* Group

This group consists of 28 neurons that express *Nk6* and *PitxB*. I highlighted the segmented cells of this group in the PlatyBrowser and overlaid the expression domains of *Nk6* and *PitxB* (Figure II-39). The soma are mostly distributed in the ventro-medial region of the VNC. They are all in the second segment.

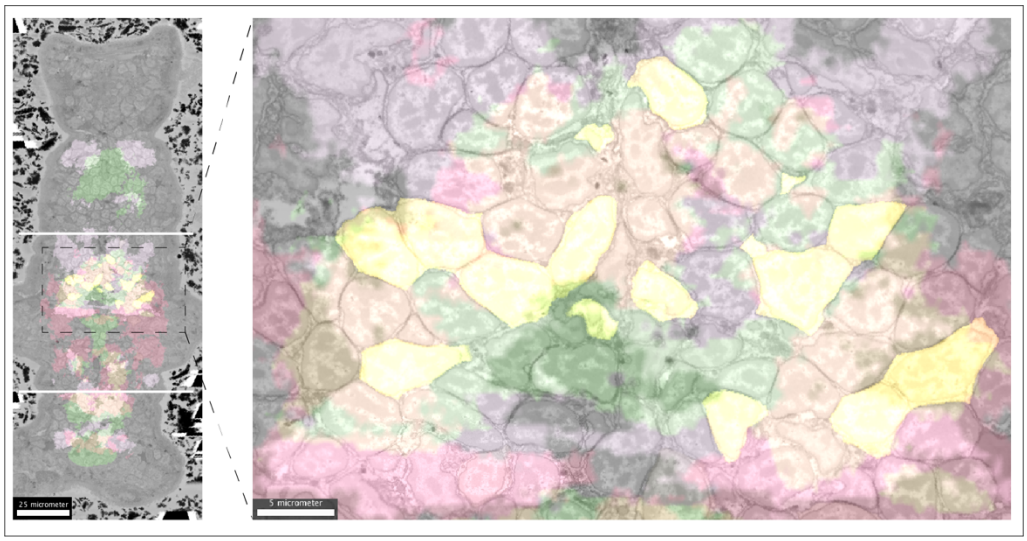


Figure II-39. Segmented cells of cells of the *Nk6*+, *PitxB*$^{+}$ group of commissural neurons

Ventral views of the SBF-SEM dataset. The expression pattern of *Nk6* is shown in green, of *PitxB* in pale pink, and of *Lbx1* in red. The segmented cells of the commissural neurons of the *Nk6*+, *PitxB*$^{+}$ group are highlighted in yellow. Scale bar 25 µm for left panel and 5 µm for right panel. I added to horizon*tal* bars to the left pan to delimit the second segment. All of the cells in this group are in the second segment.

To characterise these cells molecularly, I extracted the overlap of gene expression for every segmented nucleus of this group (overlap values) using the PlatyBrowser. I distinguished between broadly- and sparsely-expressed genes (Figure II-3). Finally, I binarized the overlap values using 0.5 as a threshold for broadly-expressed genes and 0.4 for sparse genes. I considered overlap values below the threshold to indicate that this gene is not expressed in this segmented nucleus. I considered overlap values equal or above the threshold to indicate that this gene is expressed in this segmented nucleus.

They are all mature neurons based on their expression of the sodium voltage-gated channel scn8aa. In contrast to the other groups, these cells express the glutamate receptor grik3, meaning that they are post-synaptic to glutamatergic neurons. We can also observe that most of these cells divide between 42 and 48hpf, and about half of them express *Pax6*.

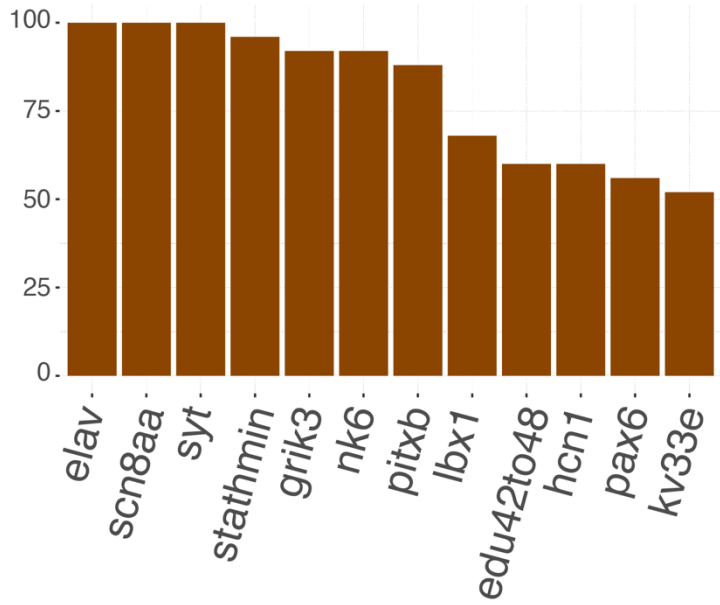


Figure II-40. Genes expressed in more than half of the neurons of the *Nk6*, *PitxB* group

This bar plot shows the genes that are expressed in more than half of the segmented nuclei of the *Nk62*+, *PitxB*+ group of commissural neurons. The height of each bar indicates the percentage of cells that express this gene. I decided that a gene was expressed if it its overlap value in the segmented nucleus was equal or higher than 0.4 or 0.5 for sparse and broad genes, respectively. Many of these genes are markers of differentiated neurons. Elav is a conserved RNA-binding protein with a role in neuronal differentiation. Scn8aa is a sodium voltage-gated channel. Syt is synaptotagmin, a protein involved in the regulation of membrane trafficking at the synapse. Stathmin is a protein involved in the regulation of the microtubule filament system. Grik3 is a glutamate receptor.Hcn1 and Kv33e is a voltage-gated potassium channels. Edu42to48 marks the nuclei that divided between 42 and 48 hpf, as determined by Edu pulses. *Nk6*, *PitxB*, *Lbx1*, and *Pax6* are TFs.

I classified almost half of these commissural axons as projecting laterally, meaning that the whole trace is within one segment (Table II-5). Most of the remaining axons project posteriorly, meaning that they cross the posterior segmental boundary. I also observed anterior- and bidirectional axons. Anterior axons cross the anterior segmental boundary. Bidirectionally axons bifurcate into an anteriorly- and a posteriorly-projecting segment.

I caution that I was not able to trace every neuron completely because of the membrane-fusion artefact (Figure II-19a). There could be neurons that I did not trace far enough to realize that they ascend or descend. In this case they would be wrongly classified as lateral.

Table II-5. Axonal projection of the commissural neurons of the *Nk6+*, *PitxB*$^{+}$ group

The axonal projection of the 28 commissural neurons that make up the *Nk6*, *PitxB* group is summarized in this table. Lateral: the whole trace is within one segment. Anterior: crosses the anterior segmental boundary. Posterior: crosses the posterior segmental boundary. Bidirectional: axons bifurcate into an anteriorly- and a posteriorly-projecting segment.

Anterior	5
Bidirectional	3
Lateral	12
Posterior	8

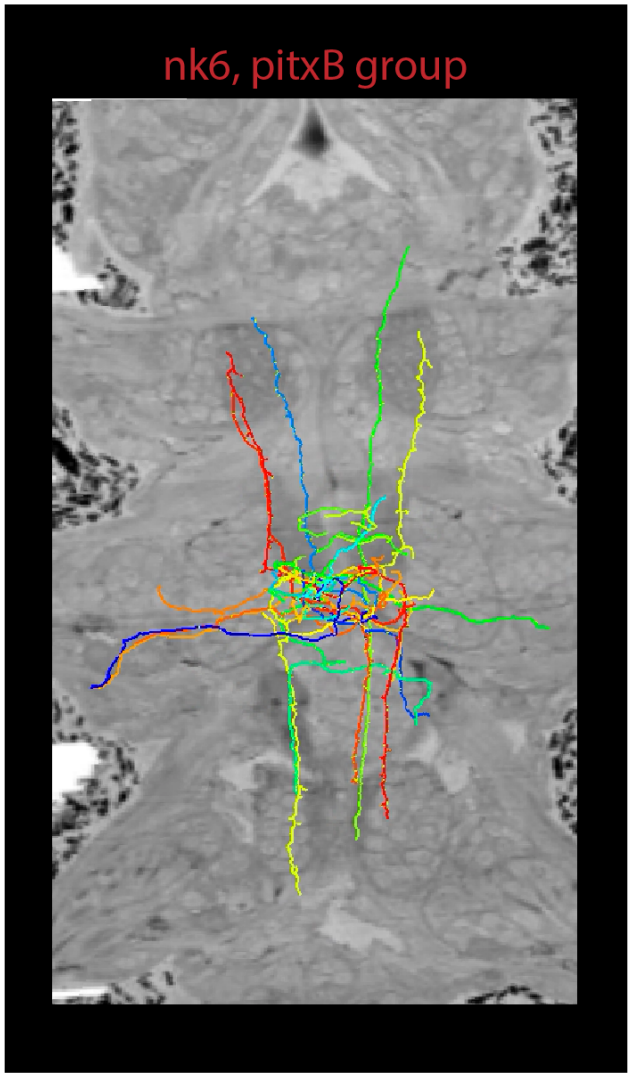


Figure II-41. Traces of the *Nk6+*, *PitxB*$^{+}$ group

Ventral view of the traces of the commissural neurons of the the *Nk6+*, *PitxB*$^{+}$ group. The traces of the 28 neurons that make up this group are shown with coloured lines. A single plane of the SBF-SEM dataset is visible, but belongs to a slightly more dorsal position so as not to interfere with the commissural traces.

We were able to identify three smaller clusters within this group that we hypothesise represent cell types.

5.3.1 THE *NK6+*, *PITXB*$^{+}$, *MAF+* COMMISSURAL NEURONS

This subgroup consists of 8 neurons. They express *Maf* in addition to *Nk6* and *PitxB*, and do not express *Pax6*. Cells with this fingerprint are found in 1 clusters in the VNC when plotted in the UMAP of the whole body (Figure II-42). Ipsilateral neurons are found interspersed among the commissural ones, suggesting this cell type might have a mixed ipsilateral and commissural phenotype.

Figure II-42. UMAP of all cells highlighting the cells that express *Nk6*, *PitxB*, and *Maf* and do not express *Pax6*

UMAP of all of the segmented nuclei of the SBF-SEM dataset, computed by Kimberly Meechan and published in Ref. [150]. It was calculated using the overlap in volume between the segmented nuclei and the expression domains of every gene in the PlatyBrowser. I coloured the nuclei that express *Nk6*, *PitxB*, and *Maf* and do not express *Pax6*. I decided that a gene was expressed if it its overlap value in the segmented nucleus was equal or higher than 0.4 or 0.5 for sparse and broad genes, respectively.

These cells are located more medially than the other neurons of this group. I did not find clear bilateral pairs of cells based on the axonal projection among these neurons. One neuron (orange trace in Figure II-43a) projects anteriorly into the first segment after crossing the midline. Two cells (green and cyan traces in Figure II-43a), project posteriorly into the third segment after crossing the midline. Two other cells (blue and green traces in Figure II-43a) project distally into the parapodia.

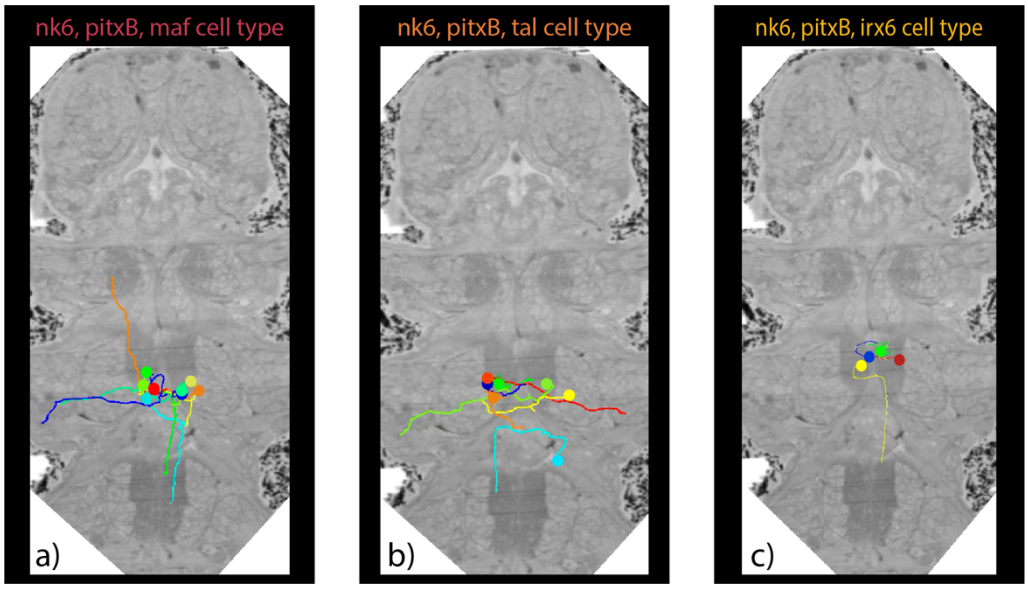


Figure II-43. Traces of the three *Nk6+*, *PitxB*$^{+}$ cell types

The left panel shows the traces of 8 neurons that make up the *Nk6+*, *PitxB*$^{+}$, *Maf+* cell type. The middle panel shows the traces of the 7 neurons that make up the *Nk6+*, *PitxB*$^{+}$, *Tal*$^{+}$ cell type. The right panel shows the traces of the 4 neurons that make up the *Nk6+*, *PitxB*$^{+}$, *Irx6*$^{+}$ cell type. The soma positions are indicated with coloured circles in the middle and left panels. In each panel a single plane of the SBF-SEM dataset is visible, but belongs to a slightly more dorsal position so as not to interfere with the commissural traces.

5.3.2 THE *NK6*+, *PITXB*$^+$, *TAL*$^+$ COMMISSURAL NEURONS

This subgroup consists of 7 neurons. They express *Lbx1*, and *Hox*4 in addition to *Nk6*, *PitxB* and *Tal*. Most of the cells with this fingerprint cluster together in the VNC when plotted in the UMAP of the whole body (Figure II-44). They are interspersed with ipsilateral neurons and very close to the cluster from the *Maf*+ cell type (Figure II-42), indicating that they are closely related.

Two cells have a similar projection pattern (red cell and the green cell whose soma is on the right side of the midline in Figure II-43b). They both project distally into the parapodia. One cell (cyan trace in Figure II-43b) starts projecting anteriorly before turning towards the midline, crossing it, and then projecting posteriorly. The remaining axons could not be traced very far.

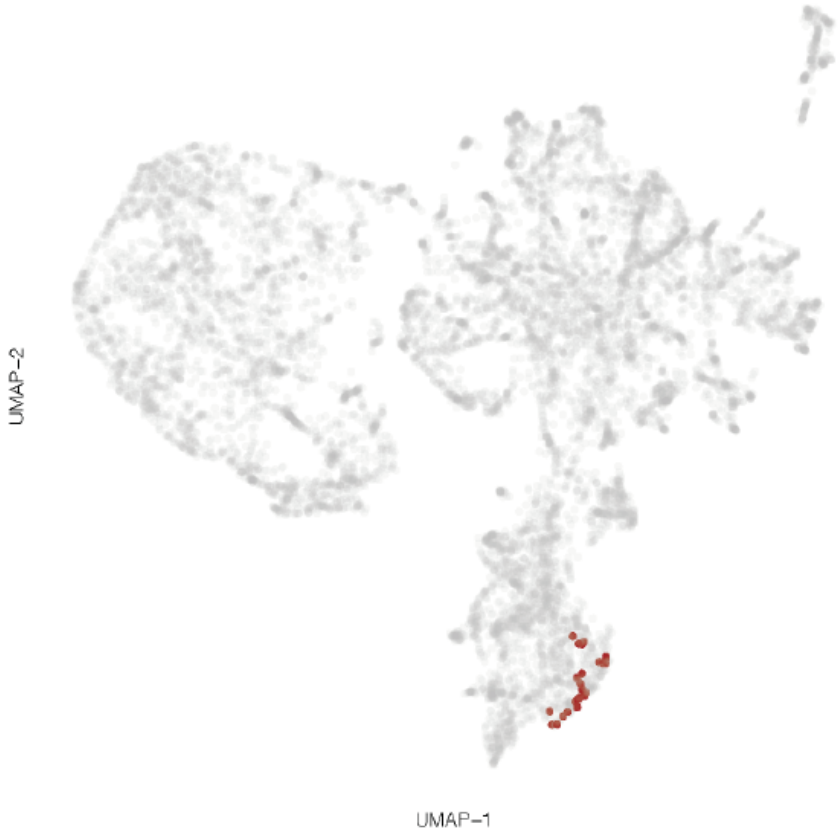


Figure II-44. UMAP of all cells highlighting the cells that express *Nk6*, *PitxB*, *Tal*, *Lbx1*, and *Hox*4.

UMAP of all of the segmented nuclei of the SBF-SEM dataset, computed by Kimberly Meechan and published in Ref. [150]. It was calculated using the overlap in volume between the segmented nuclei and the expression domains of every gene in the PlatyBrowser. I coloured the nuclei that express *Nk6*, *PitxB*, *Tal*, *Lbx1*, and *Hox*4. I decided that a gene was expressed if it its overlap value in the segmented nucleus was equal or higher than 0.4 or 0.5 for sparse and broad genes, respectively.

5.3.3 THE *NK6+*, *PITXB*$^+$, *IRX6*$^+$ COMMISSURAL NEURONS

This subgroup consists of 4 neurons. In addition to *Nk6*, *PitxB*, and *Irx6* they express *Lbx1*, *Hb9*, and *Hnf6*. The cells with this fingerprint cluster together in the VNC when plotted in the UMAP of the whole body. This cluster is very close to the other two *Nk6*+, *PitxB*$^+$ clusters we described, and are also interspersed with ipsilateral cells.

Two of these neurons have a similar projection: they project dorsally before decussating towards the midline (blue and green traces in Figure II-43c). One other cell (yellow trace in Figure II-43c) projects posteriorly into the third segment. The fourth cell (red trace in Figure II-43c) could not be traced very far, but it projects directly towards the midline.

Figure II-45. UMAP of all cells highlighting the cells that express *Nk6*, *PitxB*, *Irx6*, *Lbx1*, *Hb9*, and hnf6.

UMAP of all of the segmented nuclei of the SBF-SEM dataset, computed by Kimberly Meechan and published in Ref. [150]. It was calculated using the overlap in volume between the segmented nuclei and the expression domains of every gene in the PlatyBrowser. I coloured the nuclei that express *Nk6*, *PitxB*, *Irx6*, *Lbx1*, *Hb9*, and hnf6. I decided that a gene was expressed if it its overlap value in the segmented nucleus was equal or higher than 0.4 or 0.5 for sparse and broad genes, respectively.

The existence of this cell type had already been hypothesised [56]. This paper identifies a hypothetical cell type with the same TF signature and location in the VNC. The authors analysed virtual cells, which are groups of neighbouring supervoxels of the ProSPr atlas that have the size, internal gene-expression consistency, and bilateral symmetry that is expected of actual cells of the 6 dpf Platynereis larva (refer to the paper for more details). We, on the other hand, used the PlatyBrowser, which uses the overlap in gene-expression between the 6 dpf ProSPr atlas and the segmented nuclei of the SBF-SEM. We interpret this as a partially-independent confirmation that these cells represent a cell type.

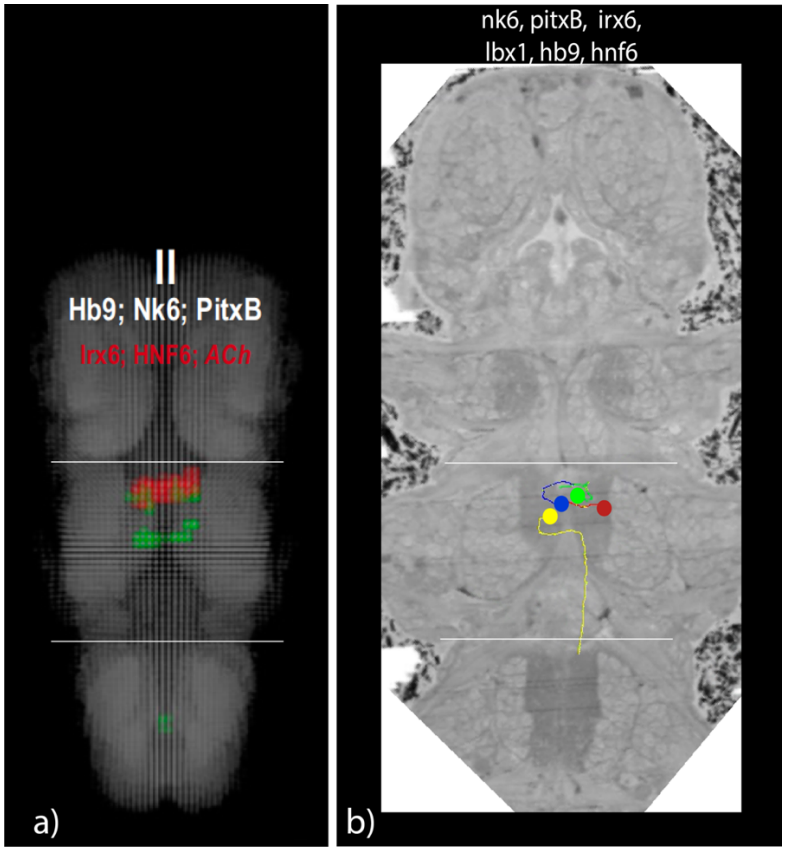


Figure II-46. Putative cell type identified in Ref. [56] (a) and this study (b)

a) This panel shows the region of the DAPI reference of the 6 dpf ProSPr atlas that corresponds to the VNC in grey. The authors of Ref. [56] analysed the gene-expression patterns in the VNC using the 6 dpf ProSPr atlas and identified a putative cell type that expresses *Hb9*, *Nk6*, *PitxB*, *Irx6*, *Hnf6*, and *ACh*. b) We found a cell type defined by the same TFs and located in the same position of the VNC. We used the PlatyBrowser, which uses the overlap in gene-expression between the 6 dpf ProSPr atlas and the segmented nuclei of the SBF-SEM. I added to horizontal bars in each panel to delimit the second segment and serve as reference points between both panels.

Regarding these three groups, we can say that based on the UMAP plots and their shared expression of TFs, that they are very closely related VNC neurons. They cluster with ipsilateral cells too, suggesting that these cell types have a mixed ipsilateral and commissural phenotype.

5.3.4 Developmental Origin Of The *Maf*+, *Tal*$^+$, And *Irx6*$^+$ Commissural Cell Types Of The *Nk6*+, *Pitxb*$^+$ Group

I used the ProSPr atlases to explore the development of these three cell types (Figure II-47Figure II-37). We found a region at 48 hpf where *Nk6* and *PitxB* are coexpressed. It is within the medial *Nk6* domain (Figure I-6) of the mediolateral patterning of the VNC. They also express neurogenin, indicating that they are progenitors.

At 72 hpf we can already identify regions of coexpression of the TFs that define each of the cell types I found at 6 dpf. These cells no longer express neurogenin. They are located in the medial VNC, like the progenitors we found at 48 hpf, and could be their daughters.

At 6 dpf, we find cells in the medial VNC that express the full TF signature of each of the three identified cell types. We suspect that they represent a subset of those identified at 72 hpf (Figure II-47Figure II-37).

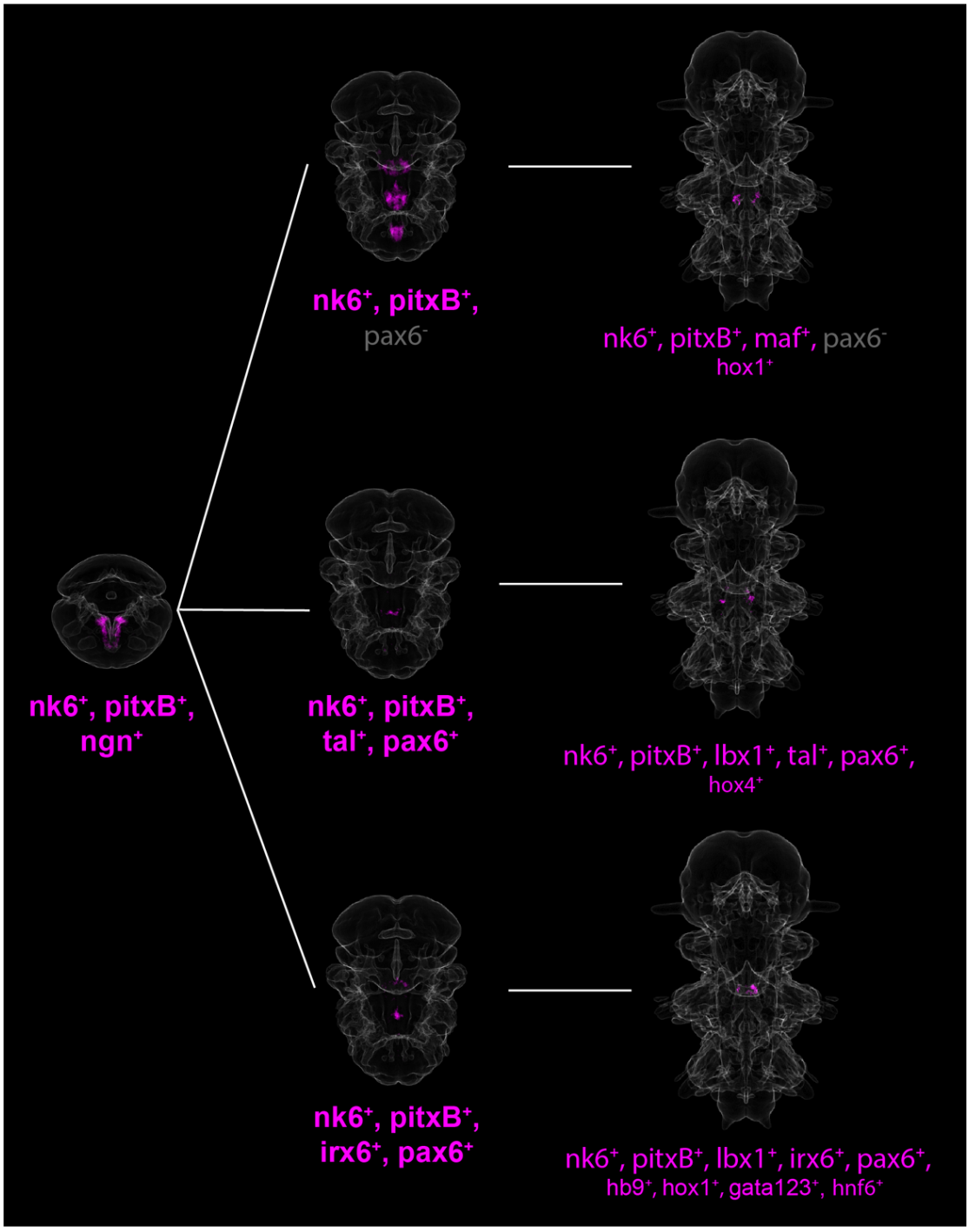


Figure II-47. Putative development of the *Maf*+, tal$^+$, *Irx6*$^+$ commissural cell types of the *Nk6*+, *PitxB*$^+$ group

These coexpression panels were generated using the ProSPr atlases of gene expression at 48 hpf (left), 72 hpf (middle), and 6 dpf (right). The text below the 48 hpf and 72 hpf panels lists the TFs that are coexpressed. The text below the panels at 6 dpf shows the coexpressed TFs in the first line. The second line lists the TFs that are additionally expressed in the corresponding commissural cell types but are not visualized. In the two panels where *Pax6* is coloured grey, it is explicitly not expressed in the coexpression region.

5.3.5 *Nk6* Is Expressed In Commissural Neurons At 48 Hpf

I also combined a WMISH against *Nk6* with immunostaining against acetylated tubulin (see Materials and Methods, section IV.8). I determined that *Nk6*$^{+}$ neurons at 48 hpf project commissural axons across the midline (Figure II-48). We cannot be certain, nevertheless, that these are the same neurons that are also *PitxB*$^{+}$.

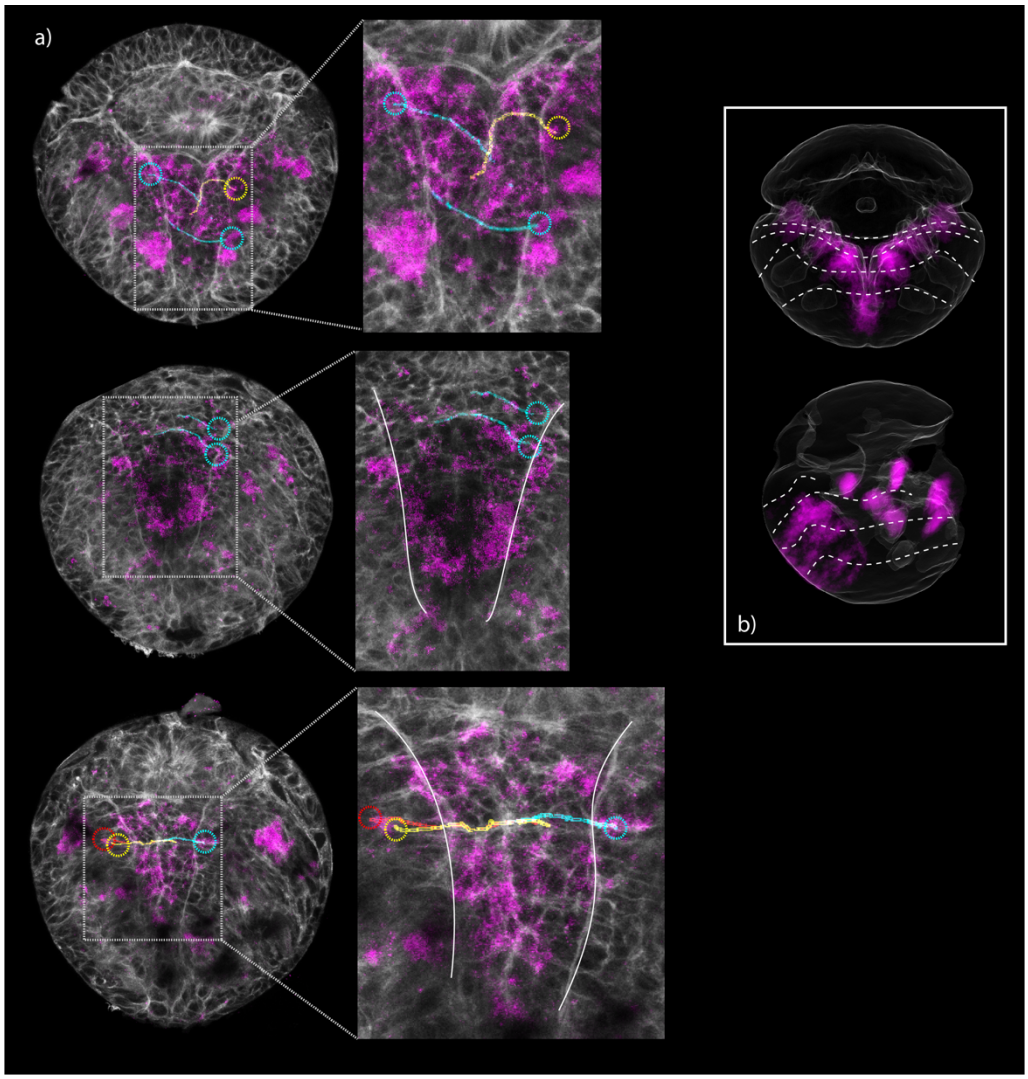
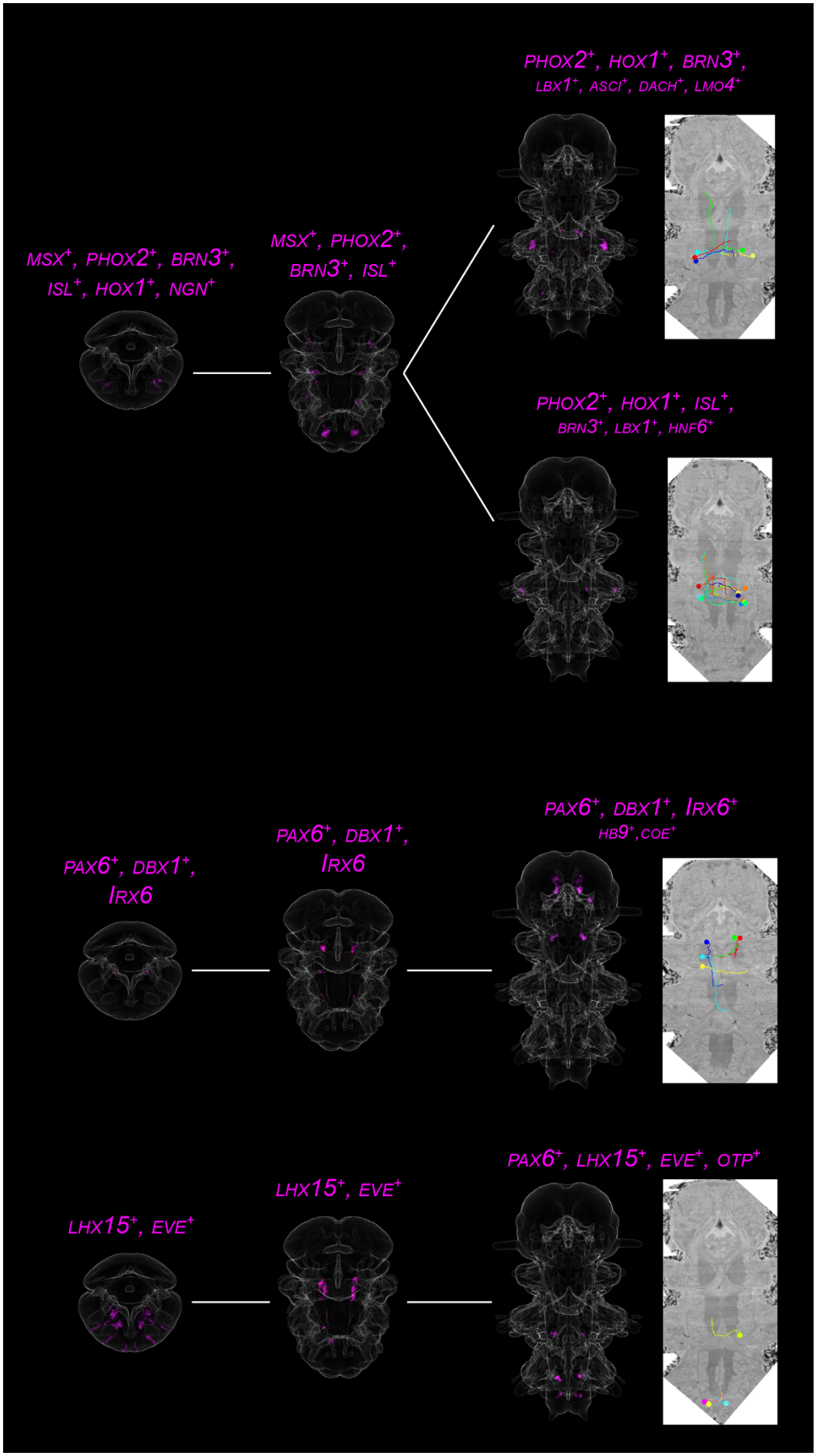


Figure II-48. Commissural axons of *Nk6*+ neurons at 48 hpf

a) Three larvae at 48 hpf that underwent a WMISH against *Nk6* and an immunostaining against acetylated tubulin. In each case, a panel zooming in on the midline is shown where two white lines delineate the nerve cords. The signal from the WMISH is in magenta and from the acetylated tubulin in grey. I traced the axons of the neurons that express *Nk6* (see Materials and Methods, section IV.10) and found that some of them were commissural. The traces of these commissural axons are shown in cyan or yellow. The dashed circles indicate the location of the somas of the traced axons. b) ProSPr expression panel of *Nk6* at 48 hpf. The dashed lines indicate the anterior boundaries of segments 1, 2, 3 and 4 (see Materials and Methods, section IV.14).


MSX+, PHOX2+, BRN3+, ISL+, HOX1+, NGN+
MSX+, PHOX2+, BRN3+, ISL+
PHOX2+, HOX1+, BRN3+, LBX1+, ASCI+, DACH+, LMO4+
PHOX2+, HOX1+, ISL+, BRN3+, LBX1+, HNF6+
PAX6+, DBX1+, IRX6
PAX6+, DBX1+, IRX6
PAX6+, DBX1+, IRX6+ HB9+,COE+
LHX15+, EVE+
LHX15+, EVE+
PAX6+, LHX15+, EVE+, OTP+

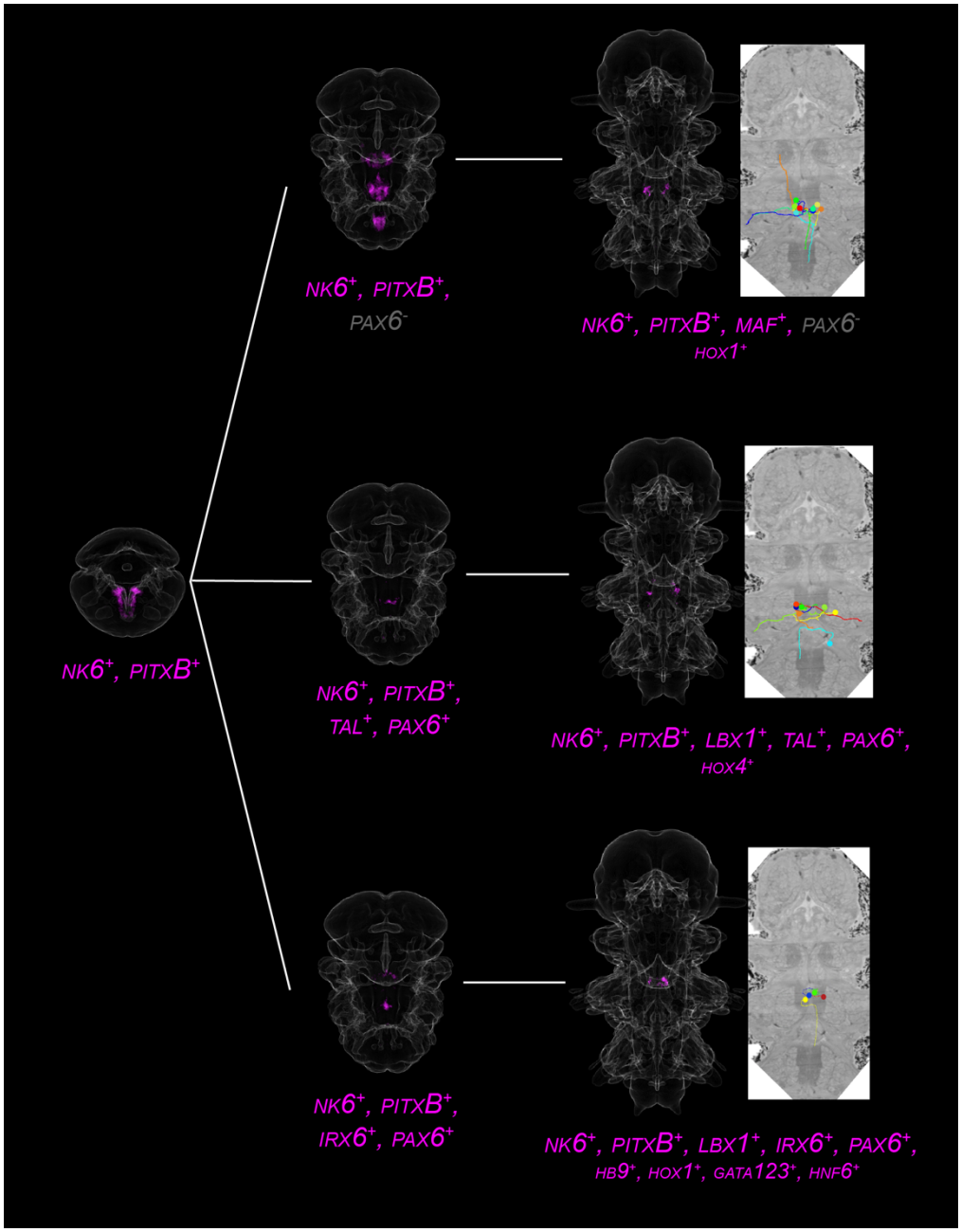


Figure II-49. Overview of the commissural cell types we identified in this thesis

The coexpression panels were generated with the ProSPr atlases of 48 hpf (left), 72 hpf (middle), and 6 dpf (right). They show the putative development of the 7 commissural cell types we identified. The coexpression of the TFs that characterise each cell type is shown in magenta. For 48 and 72 hpf, the magenta text indicates the TFs whose coexpression is shown in the panel. For 6 dpf, the magenta text in the first line indicates the TFs whose coexpression is shown in the panel and the grey text, the TFs whose expression domain does not overlap with this. The text in the second line indicates additional TFs that identify these cell types, but whose coexpression is not shown in the panels. Next to the 6 dpf ProSPr panels, we show the traces of the commissural neurons that are part of each cell type. A single plane of the SBF-SEM dataset is visible, but belongs to a slightly more dorsal position so as not to interfere with the commissural traces. All views are ventral.

III. DISCUSSION

1 Commissural Phenotype Evolved Independently In *Drosophila* Using Conserved *Neo/DCC–Netrin* And *Slit-Robo* Cues

I did not find a TF signature universal to all commissural neurons in *P. dumerilii*, suggesting that they belong to several different cell types. We identified 7 commissural cell types and hypothesise that they are generated from 3 progenitor domains. We do not know which proportion of neuronal cell types in the VNC are commissural, but I speculate that it is a minority because < 15% of the neurons in the second segment are commissural.

In the mouse neural tube, 7 out of 11 progenitor domains give rise to commissural neurons -some of them also give rise to ipsilateral ones-, and 11 of 22 postmitotic cell types contain commissural neurons[67]. In *Drosophila*, most of the neurons and neuronal cell types of the VNC are commissural and arise from all but four NBs [56,73–75].

This leads us to believe that in *Drosophila*, many cell types acquired the ability to cross the midline independently. Curiously, commissural neurons in *Drosophila* cross the midline using fewer axon guidance molecules than vertebrates do. The Neo/DCC – Netrin system, and well as the Slit – Robo system, are conserved systems. We show that Neo/DCC plays a role in commissural axon guidance in *P. dumeilii*. If Urbilateria had commissural neurons, we could hypothesise that they used these conserved cues.

2 Neuronal Cell Types In The Platynereis VNC Can Have Mixed Ipsilateral And Commissural Phenotypes

The observation that in *Platynereis* commissural neurons are not homogeneously distributed throughout the VNC (Figure II-29), indicates that not all cell types in the VNC generate commissural neurons. In fact, we could identify local concentrations of commissural neurons that were surrounded by ipsilateral neurons. On the other hand, we found clusters that seem to be made up of ipsilateral neurons exclusively (Figure II-29).

We deduce that commissural neurons arise from some cell types of the VNC that also give rise to ipsilateral neurons, and that some cell types are made up of ipsilateral neurons only. The *Pax6*$^{+}$, *Eve*+, *Lhx1/5*$^{+}$, Dbx^{-}, Otp^{+} cell type might be an exception since we did not find ipsilateral neurons with this signature, so it might represent a purely commissural cell type.

We must, however, caution that many neurons in the VNC express too few TFs to be meaningfully classified into cell types. Adding more TFs to the 6dpf ProSPr will allow us to better resolve the cell types in the VNC, both commissural and ipsilateral.

3 Three Of The Identified Commissural Cell Types In *Platynereis* May Be Conserved In Bilateria

The BNS is patterned into different neurogenic regions by the conserved expression of transcription factors across Bilateria and Cnidaria. The conservation of the cells that arise from these regions is a question that is only starting to be answered. In the introduction I summarized some of the cells that are hypothesised to be conserved. In this thesis I present evidence that three commissural cell types in *Platynereis dumerilii* are shared with other bilaterians.

The first of these cell types is made up of the *Lhx1/5*$^{+}$, *Eve*+, ascending commissural neurons, which I suggest arise from the *Pax6* domain of the BNS. I hypothesise that they are homologous to the vertebrate V0 and the insect EL commissural neurons, based on the shared expression of homologous TFs, a similar axonal projection pattern, and a developmental origin in the *Pax6* domain of the BNS.

The second and third of these cell types are the ones that I suggest arise from a *Phox2*+, *Brn3*$^{+}$, *Isl*$^{+}$, Msx^{+} domain. In the vertebrates, putative homologs have been suggested in the hindbrain[56]. I suggest they are homologous to the dlLb and dl5 commissural neurons of the neural tube, based on their TF expression pattern, developmental origin in the *Msx* domain of the BNS, and their commissural phenotype.

4 Towards An Understanding Of The Blastoporal Closing In Urbilateria

Whether the blastopore was closed in Urbilateria is open to debate. Commissural neurons are, by definition, neurons that cross the midline, and the midline is a structure that arises when the blastopore closes. A closed blastopore is, therefore, a prerequisite for commissural neurons to evolve.

Evidence showing that commissural neurons are conserved in Deuterostomes and Protostomes can be interpreted as meaning that the blastopore was closed in Urbilateria, the last common ancestor of these two clades. But the fact that in the *Drosophila* VNC most of the neurons are commissural and that almost every NB generates commissural neurons, indicates that the commissural phenotype is easily

evolvable, at least in the *Drosophila* lineage. Under these circumstances it is possible that homologous commissural neurons in Vertebrates and *Drosophila* indeed evolved from an ancestral neuronal cell type in Urbilateria, but that this ancestral cell type was ipsilateral and acquired a commissural phenotype convergently in both clades.

The even-skipped$^+$ commissural neurons have been hypothesised to be a conserved commissural cell type across Bilateria (see Introduction, section I.4). They have been found in a number of Ecdysozoan and Deuterostome species. But it is not clear whether in all of the species studied the commissural phenotype seems to have evolved multiple times or not. This is a crucial question that affects how we can interpret the conservation of commissural cell types in Bilateria.

Platynereis dumerilii belongs to the Lophotrochozoan clade, in which the *even-skipped*$^+$ cells have not been found, nor looked for before, to the best of my knowledge. We found that in the second segment of the 6 dpf larva, under 15% of all neurons are commissural. Using their TF-expression profile, we were able to categorize 70/172 (164 in the second segment plus the 8 that we found in the first segment and the pygidium) of the commissural neuron into three groups depending on their developmental origin at 48 hpf: the first arises from a *Phox2*+, *Brn3*$^+$, *Isl*$^+$, *Msx*$^+$ domain, the second from a *Pax6*$^+$, *Dbx1*$^+$, *Eve*+, *Lhx1/5*$^+$ domain, and the third one from a *Nk6*+, *PitxB*$^+$ domain.

We hypothesise that two of these groups represent homologs of vertebrate cell types. The first group has a similar molecular fingerprint and developmental origin in the lateral-most, msx$^+$, domain of the BNS as the dlLb and dl5 commissural neurons of the vertebrate neural tube. The second group has a similar fingerprint and developmental origin in the medial, *Pax6*$^+$, domain of the BNS as the V0 commissural neurons of the vertebrate neural tube. This second group harbours the even-skipped$^+$, anteriorly-projecting commissural neurons that have been hypothesised to descend from an ancestral commissural cell type in Urbilateria. We have not found putative vertebrate homologs of the third group.

Given that in *Platynereis* a minority of the neurons are commissural, and that 44/70 that we could ascribe to a cell type seem to be conserved in vertebrates, our data are consistent with the hypothesis that these commissural cell types are conserved not only with respect to their molecular fingerprint but also to their commissural phenotype. From this interpretation, it would follow that the blastopore in Urbilateria was closed.

We caution that this is not the only possible interpretation of our data. 26 of the commissural neurons that we ascribed to a cell type do not seem to have a homolog in vertebrates.

5 Future Directions

There are around 100 commissural neurons in our catalogue that we could not classify into cell types. The incorporation of additional markers into the ProSPr atlas will help us resolve more clearly the cell type identity of the commissural cell types. It is possible that after doing this, the only commissural neurons in *Platynereis* that appear to be conserved are the ones we already described, representing a minority. Finding, for example, that homologous cell types are commissural in *Platynereis* and ipsilateral in vertebrates would support the hypothesis that conserved axonal projection might be convergent.

To further test the homology of the commissural cell types across species one could study the regulatory relationships between TFs and the commissural phenotype. The concept of a CoRC (Introduction, section I.1.1) can be used for this. In both *Platynereis* and vertebrates, some commissural and ipsilateral neurons are closely related. In each of these cases, one can analyse in detail their CoRCs to determine the difference between the commissural and ipsilateral one. One can then attempt to reconstruct the CoRC changes that occurred in the course of the ipsilateral-commissural transition, and compare them between species. It should already be possible to do this in vertebrates. In *Platynereis*, a first step would be to determine if the *Eve*$^{+}$ commissural neurons of *Platynereis* have a closely related ipsilateral cell type.

Finally, the comparative study of commissural neurons should be expanded to other species, particularly Lophotrochozoans. This would provide insight into the evolution of the commissural neurons within Lophotrochozoa, and strengthen our comparisons to other phyla.

IV. MATERIALS AND METHODS

1 PLATYNEREIS DUMERILII CULTURE

The *P. dumerilii* larvae used in this study were obtained by crossing male and female adult worms from the breeding culture at the EMBL, Heidelberg. This culture is maintained following the protocol described in[56].

Briefly, the fertilized eggs are kept in plastic cups. After 24 hours, their jelly is removed. If these larvae are to be used in experiments, they kept in an incubator (Type KB53, Binder, Tuttlingen, Germany) at 18°C with a light/darkness cycle of 16/8h. These are the standard conditions for *P. dumerilii* larval growth and are used to time the developmental stages. If the larvae are to be used to replenish the culture, they are kept in rooms climatized at either 16°C or 18°C with a light/darkness cycle of 16/8h.

Given that the maturation of the worms is controlled by the moonlight cycle, and in order to have a constant supply of sexually-mature worms, the larvae are distributed into one of three rooms with partially overlapping moon cycles. These rooms have a baseline light/darkness cycle of 16/8h in addition to an artificial moonlight cycle. The moonlight is mimicked by a lightbulb that is kept on during the daily 8h of darkness. The cycle consists of a one-week phase when the lightbulb is kept on, followed by a three-week phase when it is kept off.

During the first 3 months, they are fed with *Tetraselmis*. After that, with a mixture of *Tetraselmis* and *Dunaliella tertiolecta*. Once the worms stop feeding, they are transferred to plastic tray until their sex can be visually determined. Once this is possible, they are transferred to trays with other worms of the same sex until they are ready to mate. Females and males are brought together in a plastic cup when they are ready to mate. The fertilization is external and the worms die shortly after mating.

2 REAGENTS AND SOLUTIONS

- 10x PBS: 18.6 mM monohydrate NaH2PO4 + 84.1 mM dihydrate Na2HPO4 + 1,75 mM NaCl. Mix the phosphates and adjust to pH 7.4. Then add NaCl.
- PTW: Dilute PBS to 1x, treat with DEPC and autoclave. Add Tween to 0.1%.
- 20x SSC pH 7.0: 3M NaCl + 0.3M Dinatrium citrate. Adjust pH to 7.0, treat with DEPC and autoclave
- 20x SSC pH 4.5: 3M NaCl + 0.3M Dinatrium citrate. Adjust pH to 4.5, treat with DEPC and autoclave
- SSCT: 1x SSC + 0.1% Tween-20

- Hybmix: 50% formamide + 5xSSC + 50 ug/ml heparin, 0.1%Tween-20, 5 mg/ml torula yeast RNA. Store at -20°C.
- MAB: 0.1M Maleic Acid + 0.15M NaCl
- MABT: MAB + 0.1% Tween-20
- Staining buffer 7.5: 100 mM TrisCl pH 7.5, 100mM NaCl, 0.1%Tween-20
- Staining buffer 7.5 + MgCl2: Staining buffer 7.5 + 50mM MgCl2 (prepare fresh every time)
- Staining buffer 9.5: 100 mM TrisCl pH 9.5, 100mM NaCl, 0.1%Tween-20, 50mM MgCl2 (prepare fresh every time)
- Staining solution: Staining buffer 9.5 + 4.5µl/ml NBT + 3.5µl/ml BCIP

3 Antibodies And Other Fluorescent Reagents

- Monoclonal Anti-Tubulin, Acetylated antibody produced in mouse. Sigma Aldrich. Catalog number: T6793.2ML
- DAPI (4′,6-Diamidino-2-phenylindole dihydrochloride). Sigma. Catolog number: D8417-5MG
- Cy5 Donkey, Anti-Mouse IgG (H+L). Jackson ImmunoResearch. Catolog number: 715-175-150

4 Fixation of *Platynereis* Larvae

- Collect larvae in meshes (100 mm mesh size, NITEX, Gebr. Stallmann).
- Relax muscles using a solution of 50% MgCl2 – 50% NSW. This step ensures that the larvae are all fixed in the same position.
- Fix in 4% PFA in PTW for 2h.
- Rinse 5x3min in PTW
- Store the larvae at 4°C for up to a week.
- Alternatively, dehydrate in increasing concentrations of MeOH in PTW (5min 25% MeOH, 5min 50% MeOH, 5min 75% MeOH, 5min 100% MeOH) and store in 100% MeOH at -20°C. Note: if phalloidin will be used, use EtOH instead of MeOH.

5 IMMUNOSTAINING PROTOCOL

Rehydration

- Collect fixed larvae in meshes.
- If the larvae are dehydrated, rehydrate them in decreasing concentrations of MeOH in PTW (5min 75% MeOH, 5min 50% MeOH, 5min 25% MeOH) or EtOH in PTW.
- Rinse 2x5min in PTW

Proteinase K digestion

- Treat larvae with proteinase K in PTW (100 ug/ul). For larvae younger than 48hpf, treat for 1min; between 48-72hpf, for 2min; older than 72hpf, for 3min.
- Stop the proteinase K treatment by washing 2x2min in glycine 2 mg/ml in PTW.
- Rinse 5x3min in PTW

Post-fixation

- Fix in 4% PFA in PTW for 20min
- Rinse 5x3min in PTW

Primary antibody binding

- Transfer larvae to Eppendorf tubes and replace the PTW with 5% sheep serum in PTW. Place on thermomixer at <450 rpm for 1h.
- Replace solution with 50-200 µL of primary antibody diluted in 5% sheep serum in PTW. Incubate overnight at 4°C in a thermomixer at <450 rpm.
- Rinse 5x3min in PTW at room temperature in a thermomixer at <450 rpm.

Secondary antibody binding

- Replace PTW with 5% sheep serum in PTW. Place on thermomixer at <450 rpm for 1h.
- Replace solution with 50-200 µL of secondary antibody diluted in 5% sheep serum in PTW. DAPI and phalloidin can be added at this step as well. Incubate overnight at 4°C or at room temperature for 2h in a thermomixer at

<450 rpm. From this step onwards, keep larvae in the dark to avoid bleaching of the fluorophores.

- Rinse 5x3min in PTW at room temperature in a thermomixer at <450 rpm.
- Proceed immediately to imaging.

6 RNA Probe Design For In-Situ Hybridisation

6.1 Ephrin

Ephrin is a guidance molecule involved in the guidance of commissural neurons in the vertebrate neural tube[161]. We wanted to study its expression in the *Platynereis* ProSPr atlases and designed an RNA probe for WMISH.

We started out by searching the *Platynereis* phylome for the trees that contained Ephrin sequences (Figure IV-1). We found one sequence of *Platynereis* that was homologous to all the Ephrin proteins of vertebrates. We refer to the *Platynereis* gene as simply Ephrin.

Next, we searched the *Platynereis* transcriptome for the transcript sequence associated to the protein sequence from the Ephrin gene tree. I designed the following PCR primers to amplify the transcript's sequence using IDT's PrimerQuest:

- Ephrin PCR primer 1 (5'-3') GGTGGTGGTATCATCCACATTA
- Ephrin PCR primer 2 (5'-3') CATCGTCAAGACCCTCGATAAA

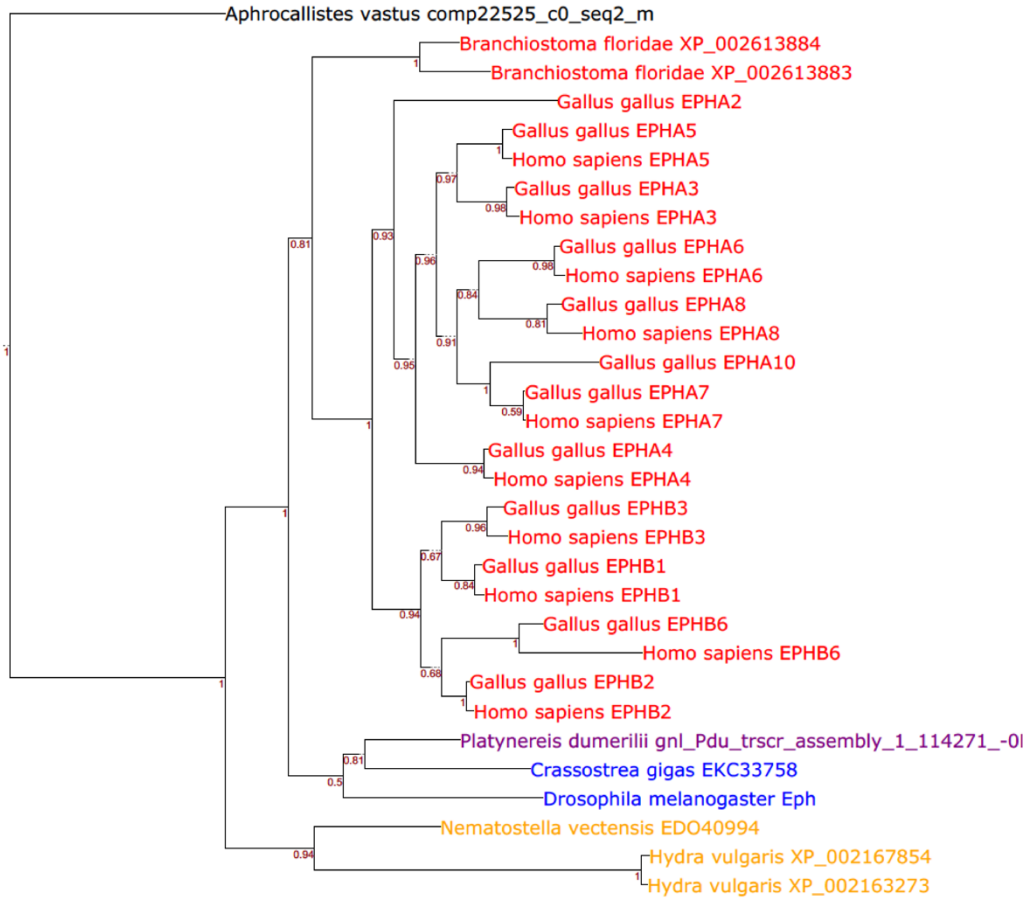


Figure IV-1. Ephrin gene tree from the *Platynereis* phylome

The *Platynereis* phylome was generated by Jacob Musser and Jaime Huerta-Cepas. They concatenated all existing *Platynereis* transcriptomes and inferred protein sequences with the program Transdecoder using default settings. The homologs of these inferred protein sequences were identified in approximately 50 metazoan species through blast. Finally, they generated gene trees using every inferred protein sequence as a seed. The methods for orthology assessment, tree building, and parsing are modifications of Ref. [174,175]. Deuterostomes are colored red, Protostomes except *Platynereis dumerilii* are coloured blue, Cnidarians are yellow, and the outgroup is black. Bootstrap values are shown at each node.

6.2 GLYPICAN 1/2/4/6

Glypicans are membrane-associated proteins that make up the core of heparan sulfate proteoglycans. Glypicans have been found to be implicated in the axonal guidance of commissural neurons[162,163]. We wanted to study its expression in the *Platynereis* ProSPr atlases and designed an RNA probe for WMISH.

We started out by searching the *Platynereis* phylome for the trees that contained Glypican sequences (Figure IV-2). We found two sequences of *Platynereis* (purple) that were homologous to the Glypican proteins GPC1, GPC2, GPC4, and GPC6 of vertebrates.

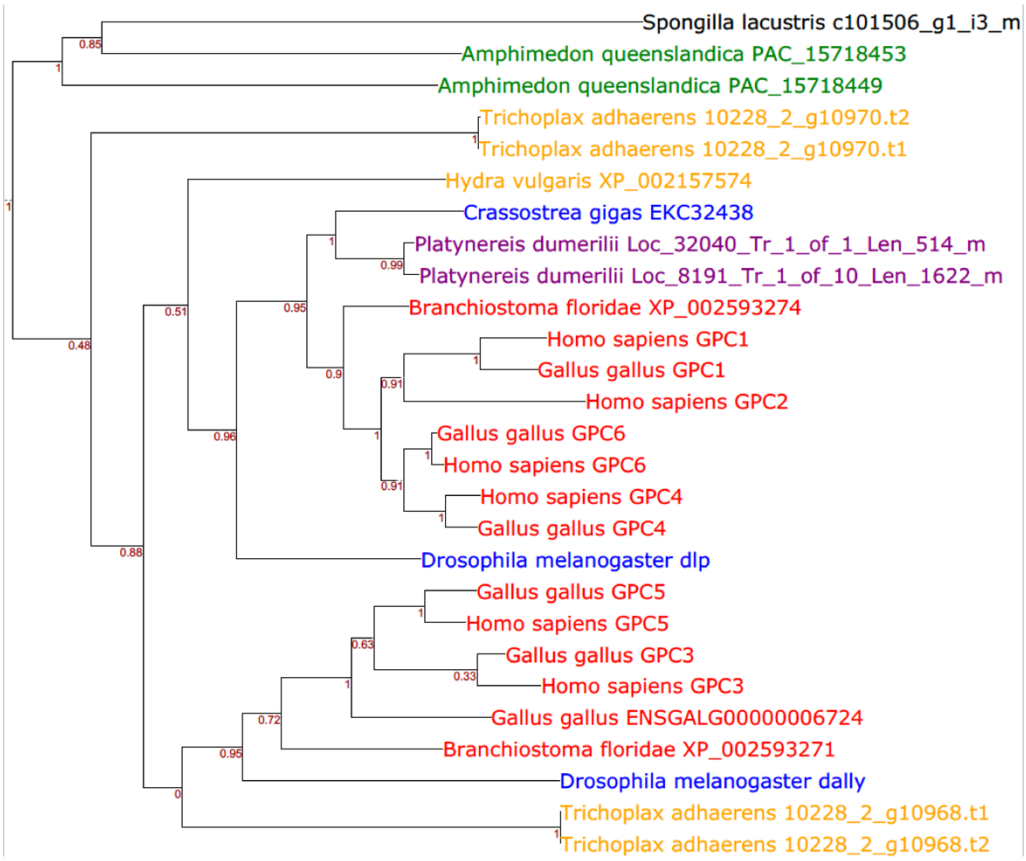


Figure IV-2. Glypican gene tree from the *Platynereis* phylome

The *Platynereis* phylome was generated by Jacob Musser and Jaime Huerta-Cepas. They concatenated all existing *Platynereis* transcriptomes and inferred protein sequences with the program Transdecoder using default settings. The homologs of these inferred protein sequences were identified in approximately 50 metazoan species through blast. Finally, they generated gene trees using every inferred protein sequence as a seed. The methods for orthology assessment, tree building, and parsing are modifications of Ref. [174,175]. Deuterostomes are colored red, Protostomes except *Platynereis dumerilii* are coloured blue, Cnidarians are yellow, Porifera are green, and the outgroup is black. Bootstrap values are shown at each node.

Next, we searched the *Platynereis* transcriptome for the transcript sequences associated to the protein sequences from the Glypican gene tree. The two *Platynereis* transcripts had an 86% sequence similarity as assessed by BLAST and we made the assumption that they were two transcripts of the same gene. We refer to the *Platynereis* gene as Glypican 1/2/4/6, shortened Gly1/2/4/6.

I designed the following PCR primers to amplify the transcript's sequence using IDT's PrimerQuest:

- Gly1/2/4/6 PCR primer 1 (5'-3') GGCAGGAATTCGACAAGACTAT
- Gly1/2/4/6 PCR primer 2 (5'-3') GAGGGCTTTCTCACACTCTTT

7 WHOLE-MOUNT IN-SITU HYBRIDISATION (WMISH)

WMISH was done following the protocol described in Ref. [176] with the following modification: larvae were acetylated after the proteinase K treatment by incubating in 1% triethanolamine (TAE) for 5 min, 1% TAE 0.2% acetic anhydride for 5 min, and 1% TAE 0.4% acetic anhydride for 5 min.

8 WMISH PROTOCOL FOLLOWED BY IMMUNOSTAINING

The WMISH protocol was modified slightly: the acetylation step was entirely skipped. After the WMISH was completed, the immunostaining protocol followed starting from the primary antibody binding step.

9 CONFOCAL IMAGING

Mounting

- Transfer larvae to meshes and immerse them in increasing concentrations of TDE in PTW: >10min in 33% TDE, >10min in 67% TDE, >10min in 99%, leave in fresh 99% TDE until imaging. If larvae are stained with phalloidin, use 2.5mg/ml DABCO in 87% glycerol instead of increasing concentrations of TDE for >1h.
- Mount larvae on slides using tape layers on each side to create space for the coverslip. For larvae <72hpf, use 2 layers on one side and 1 layer on the other side. For larvae >72hpf, use 3 layers on one side and 2 layers on the other.

Imaging

A Leica TCS SP8 upright microscope was used for confocal imaging. The NBT/BCIP precipitate of WMISH was detected with a 633nm laser in reflection mode [177]. When combined with immunostaining, it was detected with a 514nm laser in reflection mode, as the 633nm laser was used to detect the fluorescent signal of Cy5 in the secondary antibody used in conjunction with the primary antibody against acetylated tubulin.

10 Neuronal Tracing On Confocal Stacks Using Simple Neurite Tracer

Single axons can be traced on confocal stacks with the Simple Neurite Tracer[178] plug in of Fiji. We found that this method worked on 48 hpf and 50 hpf larvae that were stained against acetylated tubulin using the immunostaining protocol we describe in section 5.

We used this to count the commissural axons in the functional analysis of the genes *Dbx1* and Neo/DCC and in combination with WMISH to determine whether the genes *Dbx1*, *Phox2*, and *Nk6* are expressed in commissural neurons.

To trace an axon with the Simple Neurite Tracer we started by selecting a seed to trace an axon. This can be, for example, the axon hillock of a neuron in the VNC with signal from the WMISH (arrowhead in Figure IV-3b). Next, we add nodes along the axon which will be joined by a line which constitutes the trace.

We can navigate the confocal stack different view panes (XY, YZ, or XZ) to trace the different segments of an axon. In Figure II-6c the left panels show the initial segment of an axon in red in the XY and XZ view panes. The right panels show a later segment of the same axon and illustrates how and axon that is difficult to trace in the XY pane might be more easily followed in the XZ pane.

In some cases it was impossible to distinguish the path of an axon, especially in regions with high axonal density. We discarded the traces in this case.

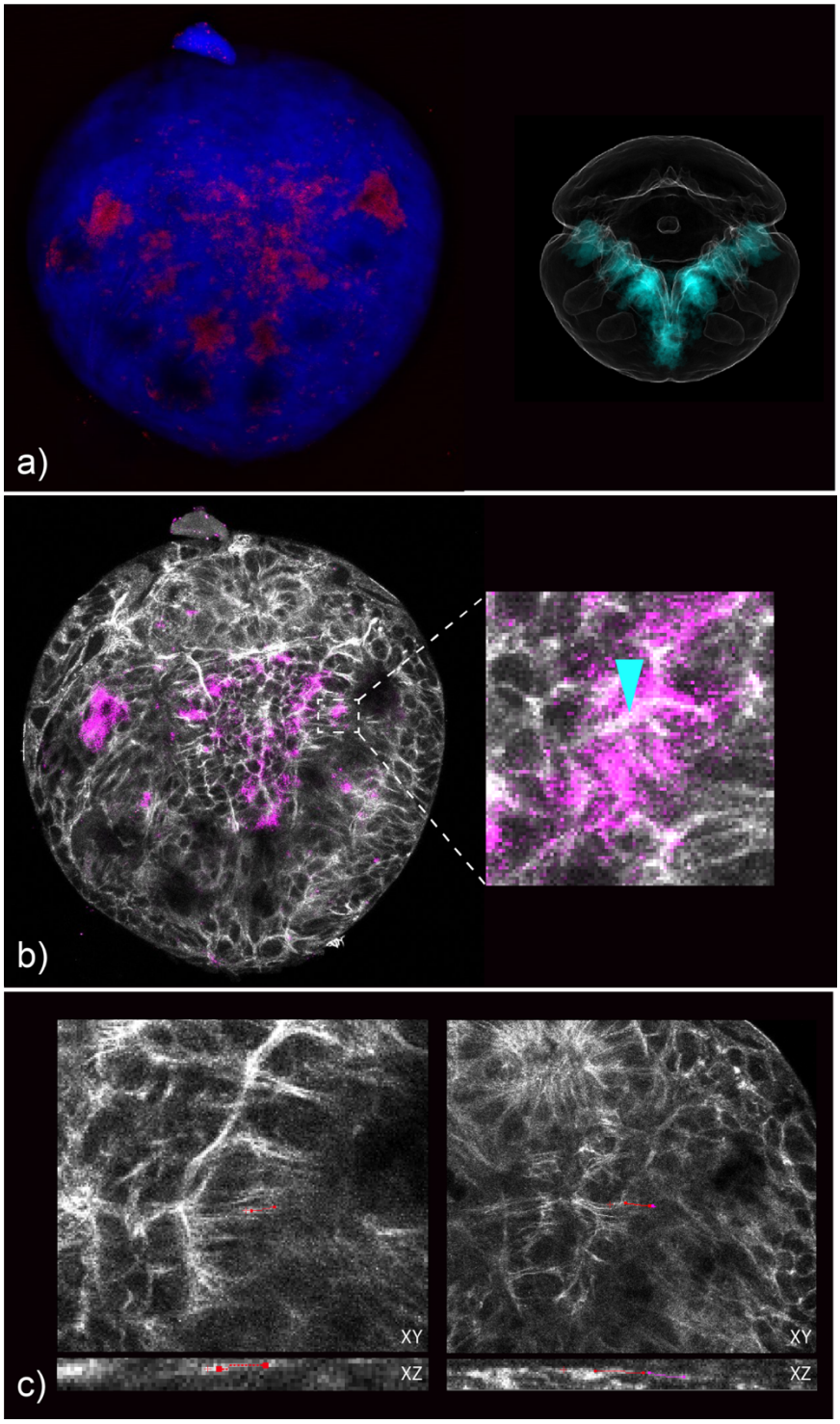


Figure IV-3. Tracing axons on confocal stacks with Simple Neurite Tracer (SNT)

a) Left panel: Z-projection of a confocal stack of a 48 hpf larva showing the signal from the DAPI channel in blue and the signal from the WMISH channel in red. Right panel: Expression pattern of *Nk6* from the 48 hpf ProSPr atlas as a reference. **b)** Z-projection of a confocal stack of the same larva as in a), showing the signal from channel of the immunostaining against acetylated tubulin channel in grey and the signal from the WMISH channel in magenta. The inlay magnifies an *Nk6+* cell whose axon we traced with the SPN. The cyan arrowhead points to the axonal hillock of this neurons that we used as a seeding point for our trace. **c)** Screenshot of the XY and XZ navigation viewports of SNT. Left: two nodes along the axon are shown in red. Right: 4 nodes of this trace are shown. Deciding on the position of the left-most node was impossible in the XY viewport, but the XY provided a perspective that allowed us to decide where to place this node.

11 CAS9-MEDIATED MUTATION OF *NEO/DCC*

We started by blasting the protein sequence of the *Zebrafish* neo1a gene (Uniprot entry Q8JI27_DANRE) to the *Platynereis* proteome v11 using the software CLC. Next, we found the gene tree in the *Platynereis* phylome corresponding to the first hit of the blast (Figure IV-4). Three *Platynereis* protein sequences cluster together and are the sister group to the vertebrate Neo1 and DCC genes. The *Drosophila* frazzled (fre) gene is the sister of all the other protein sequences shown.

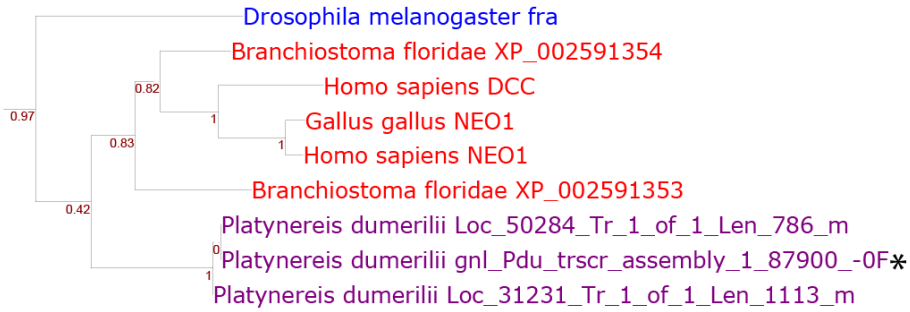


Figure IV-4 – Neo/DCC gene tree from the *Platynereis* phylome

The *Platynereis* phylome was generated by Jacob Musser and Jaime Huerta-Cepas. They concatenated all existing *Platynereis* transcriptomes and inferred protein sequences with the program Transdecoder using default settings. The homologs of these inferred protein sequences were identified in approximately 50 metazoan species through blast. Finally, they generated gene trees using every inferred protein sequence as a seed. The methods for orthology assessment, tree building, and parsing are modifications of Ref. [174,175]. Deuterostomes are colored red, Protostomes except *Platynereis dumerilii* are coloured blue. The first hit in the *Platynereis* proteome from the blast of the *Zebrafish* neo1a protein sequence is marked with an asterisk.

We blasted the three *Platynereis* transcripts that correspond to the protein sequences shown in the gene tree against all genomic sequences of *Platynereis* using the Blast tool of the Arendt Lab (http://4dx.embl.de/platy/) with the default settings. I found that they all map to the same scaffold and deduced that they are coded in the same genomic locus.

We therefore named this gene Neo/DCC, to reflect that it is a one to many paralog of the vertebrate Neo and DCC genes.

11.1 TARGET SITE IDENTIFICATION

We used the expasy translating tool (http://web.expasy.org/translate/) to translate the protein sequence of the longest of the three transcripts. We kept the longest translated sequence and searched for structural domains using the CATH tool (http://www.cathdb.info) and found that the netrin receptor was coded in the first exon (Figure IV-5).

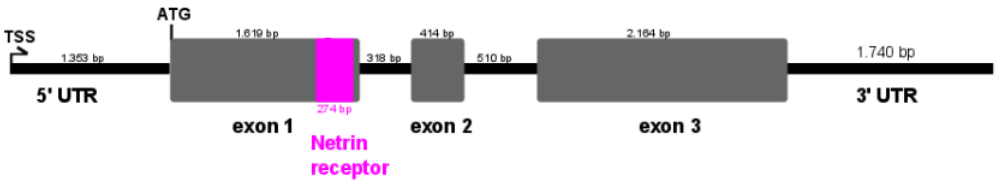


Figure IV-5. Neo/DCC locus. Reproduced from Ref. [157]
This representation of the *Platynereis* Neo/DCC locus was done by Ines Kübler. We determined that Neo/DCC has three exons by comparing the Neo/DCC transcript and the genome. We found that the Neo/DCC transcript mapped to a single place

We identified three Cas9 target sequences for the Netrin receptor of Neo/DCC using the sgRNA scorer 2.0 tool[179]. The tool predicts the efficiency of the target site as a cleavage site for Cas9. We selected three sequences that were predicted to be the most efficient.

We cloned these target sequences into the Addgene DR274 plasmid (Plasmid # 42250)[180] following the cloning strategy described in the publication. This plasmid can be used to transcribe a mature sgRNA which contains the target sequence as well as the guide RNA scaffold that is necessary for the Cas9 protein to be functional.

Table IV-1. Oligos ordered for the target sequence of Neo/DCC
We identified three target sites in the Netrin receptor of Neo/DCC for which we designed sgRNAs. The two oligos that we ordered for each target sequence are shown. In bold are the nucleotides that are necessary for the cloning into the DR274 sequence and that are independent of the target site.

Target site (5'-3') Pam underlined	Oligo 1 (5'-3')	Oligo 2 (5'-3')
GATGGCCGCTGCAAGGTCCCTGG	**TA**GATGGCCGCTGCAAGGTCCC	**AAAC**GGGACCTTGCAGCGGCCA
GGATAGAGTCACAAATCTGGTGG	**TA**GGATAGAGTCACAAATCTGG	**AAAC**CCAGATTTGTGACTCTAT
AGGAAGGATCCACCAGAGTGAGG	**TA**AGGAAGGATCCACCAGAGTG	**AAAC**CACTCTGGTGGATCCTTC

11.2 Cloning Of The Target Sequence Into The DR274 Plasmid

We annealed the oligos by heating 3 µL of each oligo pair (100 uM) in T4 ligase buffer for 10 min at 96°C in a PCR machine. After the 10 min, we stopped heating but we left the machine closed and allowed it to cool down for 3 h.

We phosphorylated the annealed oligos with T4 polynucleotide kinase at 37°C overnight and heat inactivated it by heating it for 10 min at 65° C. In parallel, we cut the DR274 vector with BsaI by heating it for 5h at 37°C. We ligated 0.1 µL of the annealed and

phosphorylated oligos into 70 ng of the cut plasmid, using T4 DNA ligase and incubating overnight at 16° C.

We used 5 µL of the ligation product to transform 100 µL of chemically competent DH10-ß cells and plated on LB medium with Kanamycin. The plates were incubated at 37° C overnight. We picked individual clones to inoculate 5 ml of LB-Kanamycin growing medium that were grown overnight at 37°C while shaken.

We extracted the cloned plasmid from these cultures using the Qiagen Mini-prep kit, following the manufacturer's protocol and sequenced them by Sanger sequencing (Eurofins Genomics) with the M13 forward primer. We found plasmids where each of the three target sequences had been inserted.

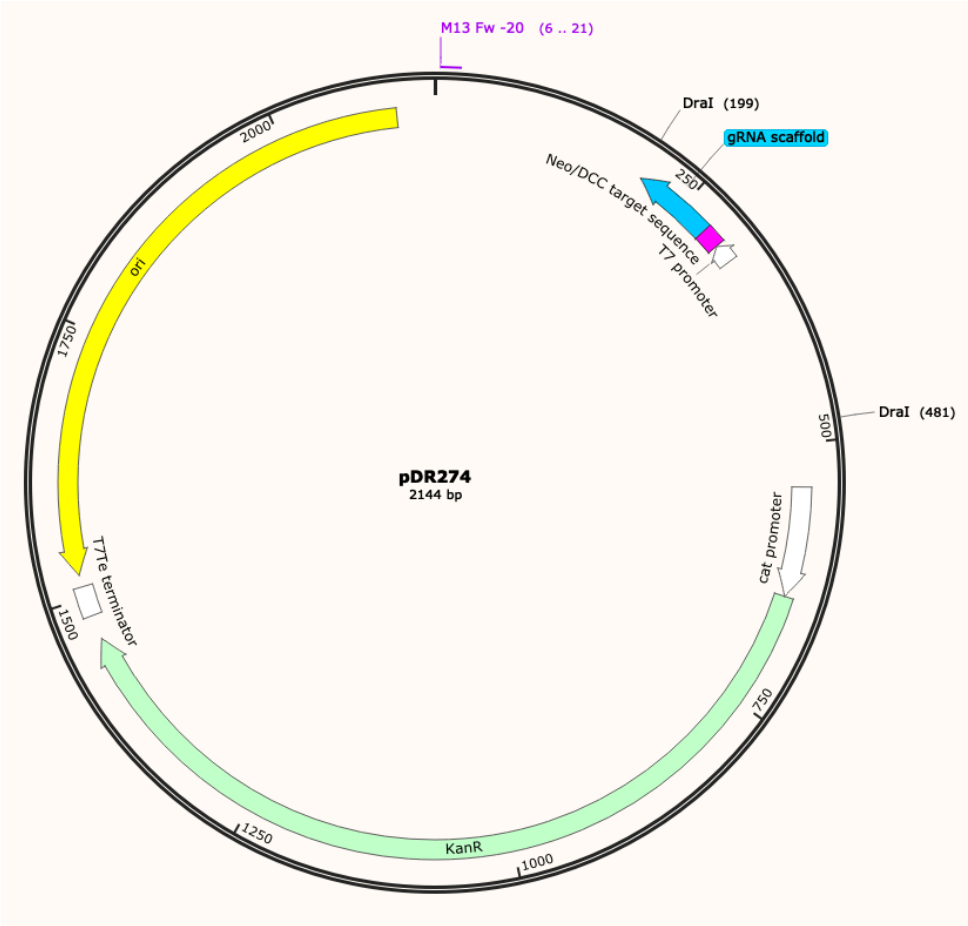


Figure IV-6. Plasmid map of the cloned DR274 with a cloned Neo/DCC target sequence
This screenshot shows the plasmid map and was generated using the software Snapgene.

11.3 In-Vitro Transcription Of sgRNA

We cut the cloned plasmids with DraI by heating at 37° C overnight. This cut the plasmid into two fragments: one contains the cat promoter and is 1862 bp long, the other one contains the Neo/DCC target sequence and is 282 bp long. We separated these bands through an electrophoresis on agarose gel, and then extracted the shorter fragment using the Qiqgen Gel Extraction Kit, following the manufacturer's protocol.

The sgRNA was transcribed from the purified short fragment with the T7 MEGAshort script kit following the manufacturer's protocol.

11.4 Micro-Injection Of *Platynereis* Zygotes

Zygotes were microinjected with an injection cocktail consisting of equal parts of sgRNA (50 ng/ul), Cas9 mRNA (1ug/ul), and dextran using a FemtoJet express Microinjector. The needles were pulled with a Sutter needle puller with the following settings: pressure: 500, heat: 530, pull: 40, velocity: 130.

Fertilized eggs were incubated at 18 °C for approximately 1 h. The jelly around the eggs was removed by washes with filtered natural sea water (FNSW) and treated with ProteinaseK (48 ug/ml) for 40 seconds. After rinsing them with abundant FNSW, they were transferred to a 2% agarose mould for microinjection. Zygotes were injected until the first cell division.At this point they were transferred to a 6-well plate with FNSW and kept at 18° C.

At approximately 48 hpf, the larvae were inspected under a fluorescent microscope to detect the injected larvae through the fluorescence of dextran in the RFP1 channel.

11.5 Mutant Genotyping

We extracted the genomic DNA (gDNA) of the injected larvae. We pipetted the larvae in 8.6 µL of FNSW and mixed it with 1 µL of 10x Kappa extraction buffer and 0.4 µL of Kappa express extraction enzyme.

We amplified the region surrounding the cutting site for Cas9 through PCR using the following oligos.

Neo/DCC genotyping oligo 1 (5'-3')	AGTTTCTGTGGACTCCTTCGA
Neo/DCC genotyping oligo 2 (5'-3')	CGGCTTTGGTCTTCACAGAG

We mixed 0.46 µL of DMSO, 3 µL of GC Phusion PCR buffer, 0.75 µL of each oligo, 0.3 µL dNTPs, 0.75 µL of gDNA, and 0.1 µL of Phusion polymerase. The PCR was programmed with 35 amplification rounds with 15 sec of denaturation at 98° C, 30

second of annealing at 62.7 °C, and 2 min of extension at 72° C. This was followed by a final extension step of 15 min at 72°C. The product was purified with the QIA PCR purification kit.

An A-tail was added to the PCR product by mixing 17 µL of the purified amplicon, 2 µL of Taq buffer, 0.5 of dATP (100 µM), and 0.5 µL of Taq polymerase. This mixture was incubated for 30 min at 72° C and immediately ligated into the TOPO vector using the TOPO cloning kit.

The ligation product was used to transform chemically competent XL1ß cells and plated them in LB gowing medium with ampicillin and kanamycin, and 40 µL of X-Gal and 46 µL of IPTG for blue/white selection of the colonies. We incubated the plates at 37° C overnight. White colonies were inoculated in 5 ml of liquid LB medium with ampicillin (1:1000) and left to grow overnight at 37° C while shaking.

We extracted the plasmid from this culture using the Quiagen mini-prep kit following the manufacturer's protocol. The purified plasmids and sent them to sequence with the M13 rev primer.

Figure IV-7. Genotyping strategy for the Neo/DCC mutants. Reproduced from Ref. [157]
This is part of the sequence of TOPO vector into which an amplicon for the genotyping of Neo/DCC was inserted (red). The three target sequences for the sgRNAs that we used are in green. The red sequence can be amplified by the oligos found at its extremities (purple arrows), and correspond to the Neo/DCC genotyping oligos 1 and 2. The M13 rev primer was used for sequencing.

Finally, we aligned the sequences of the larvae with the T-coffee multiple sequence alignment tool of EMBL-EBI (Figure IV-8). We detected a 68bp deletion in the sequence corresponding to the injected larva when compared to the WT sequence.

Figure IV-8. Alignment of the sequences from mutant and WT larvae. Reproduced from Ref. [157]

These sequences were aligned with the T-coffee multiple sequence alignment tool of EMBL-EBI. The red/yellow/green colours code how well the sequences align. The target sequence of the sgRNA used in this injection is highlighted in blue. The sequence that corresponds to the injected larva (Neo10) has a 68 bp deletion upstream of the target sequence.

12 Reanalysis Of The Commissural Count In *Dbx1* And Neo/Dcc Knock-Out Larvae

We reanalysed the results from the knock-out experiments of *Dbx1* and neo/dcc published in Martinez-Vergara, 2017a and Kübler, 2018 respectively. Instead of using an automated approach to quantify the commissural axons in 50hpf larvae, we manually counted the commissural axons.

Refer to these publications and to section for details on the methods. Briefly, larvae were grown at 18°C and fixed at 50hpf, when the ventral nerve cords are furthest apart and the commissural axons most visible. The commissural axons were visualised through immunostaining against acetylated tubulin (see Materials and Methods, section 5).

The larvae were grouped into three categories according to the treatment they received:

- Wild type (WT): uninjected control larvae

- Knock out (KO): zygotes were microinjected with Cas9 mRNA, a sgRNA directed to the either *Dbx1* or Neo/DCC, and dextran
- Control (Ctrl): zygotes were microinjected with Cas9 mRNA and dextran. In the neo/dcc experiments, a scrambled sgRNA with no target in the *Platynereis* genome was added to the injection cocktail. This sequence was obtained by scrambling a sequence from the SoxE gene. Its sequence is AGTTGTACGGGAGGCGTCTCGGC and a blast against the *Platynereis* genomic sequences using the Arendt Lab blast has not hits.

We tested for normal distribution with the Shapiro Wilk test. Homoscedasticity with the Bartlet Test. To assess the effect knocking-out a gene, we compared the number of commissures in the VNC of mutant, control, and WT larvae. We used the ANOVA test followed by Tukey's honest significance test.

13 Neuronal Tracing On The *Platynereis* SBF-SEM Dataset Using PyKnossos

The SBF-SEM Platynereis dataset used in this study is the same one described in [150]. Briefly, immobilized 6 dpf Platynereis larvae were immobilized and fixed following [151], starting directly with the cacodylate buffer incubation of step 1.

An image stack from an entire 6 dpf Platynereis larva (300 x 150 x 150 um) was acquired. This consisted of 11731 slices, 25nm thick. They were acquired with a pixel size of 10x10 nm^2. To alleviate the undesired effects of the electron beam on the sample, it was removed from the Epon resin before it was cured and transferred into an epoxy glue with silver particles, also known as EE embedding [152]. The images were acquired using low-voltage SEM, which produces images of an outstanding quality.

The manual tracing of the neurons was done using the PyKnossos software. One starts by setting a seeding point for a neuron. This can be a segmented nucleus, for example, or a commissural axon as it crosses the midline (Figure II-23). To trace a neuron, one places nodes along the axon, which become connected with a line (Figure II-18a). It is also possible to annotate synapses (Figure II-18c).

Tracing whole neurons was not always possible because the membranes of several axons sometimes fused, an artefact that was likely caused during the fixation step (Figure II-19a). The dataset also has some alignment problems (Figure II-19b).

14 Identification Of Segments In The 48 Hpf Expression Panels

While at 6 dpf and 72 hpf the segments are easily identifiable thanks to morphological features such as the parapodia, at 48 hpf it is more difficult to distinguish the segments. We used the TF *engrailed* (En) to delineate the segments, since it is expressed in segmental stiped whose anterior border abuts on the boundary between segments (Figure IV-9, Ref. [181]).

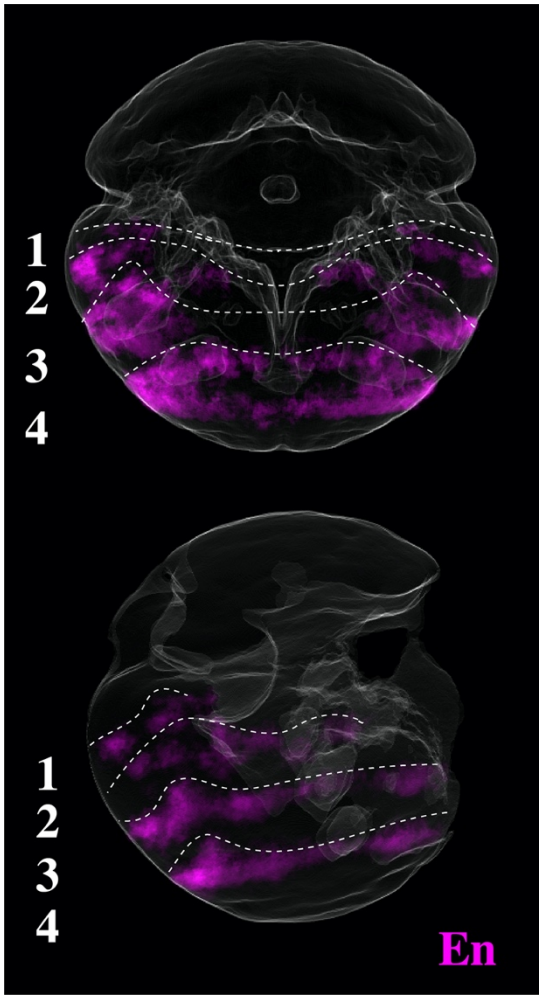


Figure IV-9. Engrailed expression to delineate segments in the 48 hpf larva

15 Phylogenetic Analyses Of The Commissural Neurons

For the phylogenetic analysis of the commissural neurons we used the gene-overlap data of transcription factors only, since these are hypothesised to carry to most amount

of evolutionary information. There are 55 transcription factors expressed in the 133 commissural neurons for which we could identify the segmented nucleus.

We used the overlap of their expression pattern with the segmented nuclei of the commissural neurons (overlap value) to classify them into any of three states: not expressed, unknown expression, expressed.

Overlap values under 0.1 comprise the vast majority (Figure II-2) and we consider that they indicate that a gene is not expressed in that cell. We next classified the TFs into broadly-expressed (broad) and sparsely-expressed (sparse) because they do not have the same distribution (Figure II-3). We used 0.4 and 0.5 of overlap value as thresholds for sparse and broad genes, respectively. It is difficult to decide if overlap values between 0.1 and the respective threshold indicate that the TF is expressed and decided we would treat these values as unknown expression. Finally, we consider overlap values equal or above the respective thresholds as indicating that the TF is expressed.

More than 100 hours of heuristic tree searching under different parameters revealed that there are many equally parsimonious trees, which is to be expected given the low number of characters. To find a consensus tree of all of these equally-parsimonious tree, we performed a heuristic search with 1000 replicates

During each replicate up to 10 initial trees are iteratively rearranged using the tree bisection-reconnection (TBR) algorithm with steepest descent, which we found to be the quickest at finding parsimonious trees in our dataset. An initial tree is obtained by a series of steps in which the taxa are added one by one. The order in which the taxa are added is random, but only the 10 most parsimonious trees are kept at each step. In other words, at every step a taxon chosen at random will be added to the growing initial tree at every position. The 10 most parsimonious trees are kept for the next step, when the next taxon is added. The algorithm continues until all the taxa have been added.

At each repetition the TBR algorithm finds the most parsimonious rearrangement of the initial trees. This tree is compared against all trees in memory and if its score is the same or lower than the most parsimonious trees in memory, it will be kept. If the tree found in a repetition is less parsimonious than the trees already in memory, it will be discarded.

We decided to limit the number of trees that we saved from each initial tree to 100. Nevertheless, if none of the trees during this search match or beat the maximum score, they are all discarded.

The 1000 replicates resulted in more than 30000 equally-parsimonious trees. We calculated the consensus tree using PAUP, keeping the clades present in more than 50% of the trees (Figure II-31)

V. ANNEX

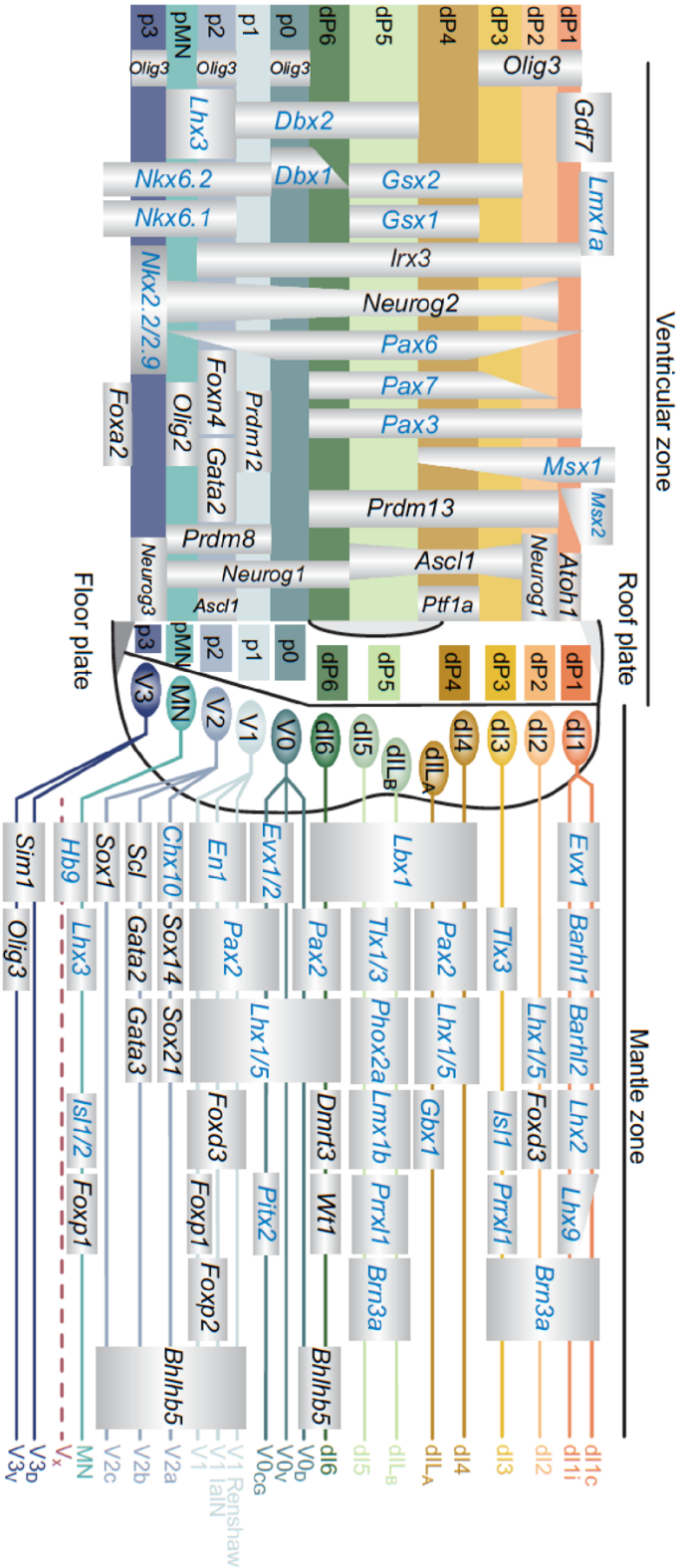
Broadly-expressed genes	*chat, edu42to48, elav, grik3, Hb9, Hox4, kv33b, kv33e, Nk6, Pax6, Phox2b, rab3, scn8aa, stathmin, syt, tmtc3*
Sparsely-expressed genes	*ache, allcr1, anpra, ascb, asci, ash, asicalpha, barh1, beta3, boule-like, brn124, Brn3a, cal2, calexcitin2, calmodulin, catl, ccvd, cepu1, chx10, cnga, co1a1, coe, Dach, Dbx1, edu3to4at6dpf, edu4to5at6dpf, edu5to6at6dpf, Eve, fezf, for, foxn4, fvri, fxl21, gad, gata123, gcb, gcd, gdpd1, glt1, glyt, gnrh, gpcr209, gpcr210, gppcr21, grm7, gsx, hand, hcn1, hnf6, Hox1, Hox7, irx2546, Irx6, isl, jkip3, kanl3, Lbx1b, leucin-rich, Lhx1/5, lhx2, lhx3, lhx6, lmo4, lmx1, loc5285ct, loc8913nt, maf, mecom, mhcl4, mrlc2, msx, mthfsd, nachr, ndus1, neog, netrin, neurod, ngb, Ngn, nmdar, noe1, ntrps, nucleolin, odo2, olm2a, otp, paraxis, patched, pcdh15, pde9, pea3, pgam, phc2, pikachu, pitxb, ppib, prdm8, prox, psmf1, robo, ryr2, sema2, sert, sim1, slit, sox2, sox4, sp8, synaptopodin, syt12, syt7, Tal, tbh, th, timp, tlx, trk, trph, trpv, trpv5, twist, unc22, uncx, upp, usp9x, vacht, vat1l, vegfr, vglut, wnt16, wnt5*

Supplementary table 1. Broadly- and sparsely-expressed genes

This table lists the genes that are expressed in the VNC and classifies them into broadly- and sparsely-expressed genes

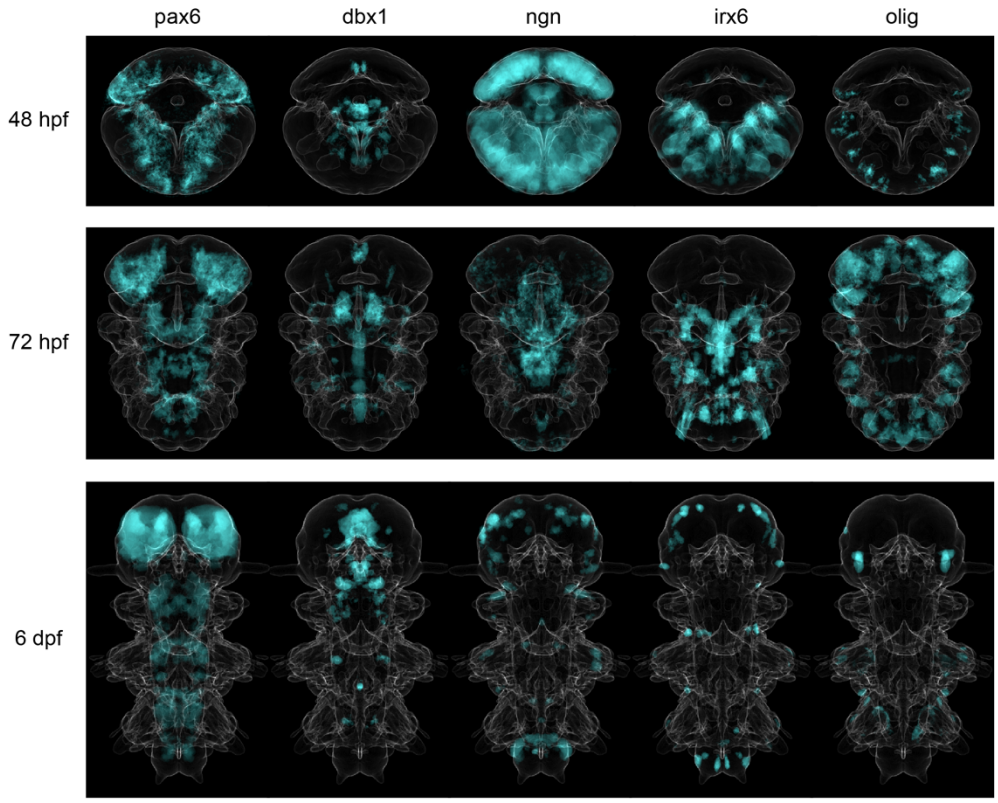
Supplementary table 2. Transcription factors of the 6 dpf ProSPr atlas

ap2, arx, ascb, asci, ascii, ash, ato2, barh1, beta3, bmpr1, brn123, *Brn3*a, bsx, chx10, coe, *Dach*, *Dbx1*, delta, dlx, *Eve,* fezf, foxd3, foxn4, gata123, gsx, hand, *Hb9*, hnf6, *Hox*7, *Hox*1, *Hox*4, hr38, *Irx6*, irx2546, isl, *Lbx1*b, *Lhx1/5*, lhx2, lhx3, lhx6, lmo4, lmx1, maf, mecom, mitf, msx, munc22, muncb, neurod, *Ngn*, nk21, nk22, *Nk6*, olig, otp, paraxis, pax258, *Pax6*, pea3, phc2, *Phox2*b, pitxb, prdm8, prox, ptf1, rx, sim1, six12, six4, sox2, sox4, soxb12, soxb2, sp8, *Tal,* tbx20, tlx, wnt16, wnt5, twist, wing



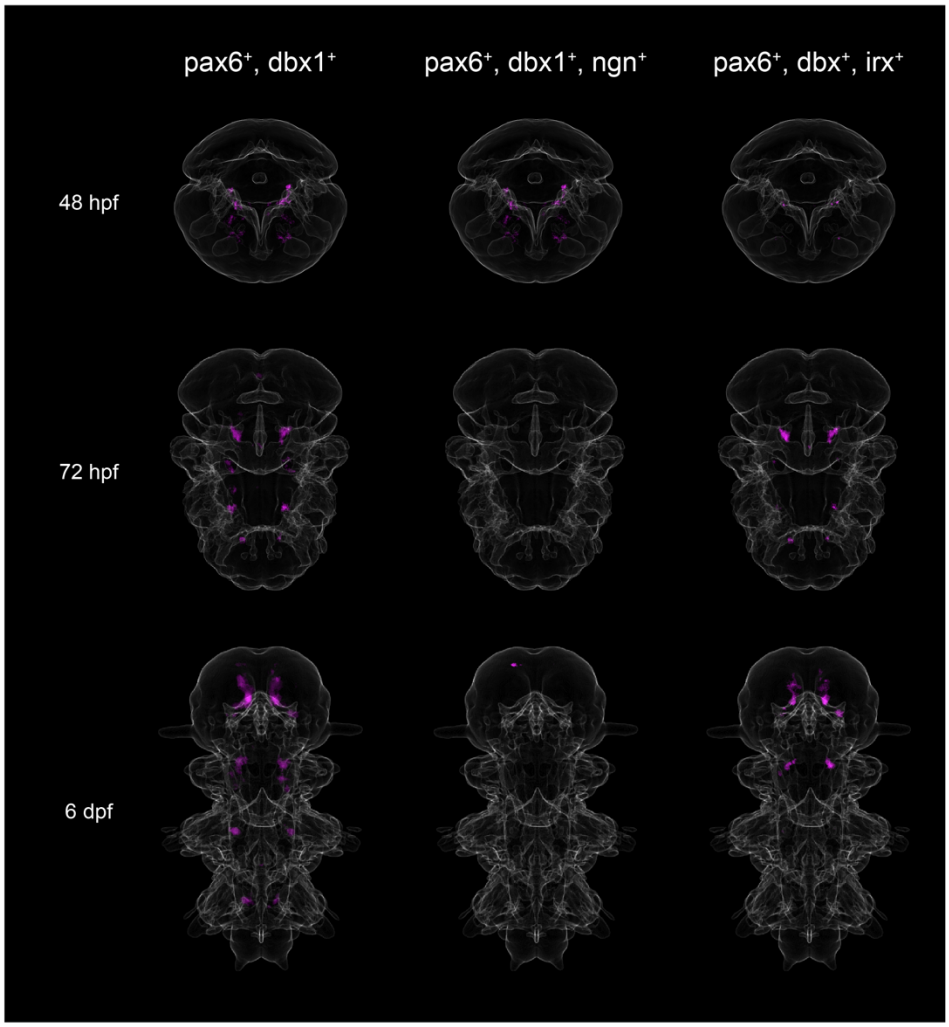
Supplementary figure 1. Gene expression in the mouse VNC at the ventricular and mantle zones. Reproduced from Ref. [158].

I reproduce this illustration as reference for the TFs that are expressed in the mouse neural tube. The ventricular zone harbours progenitors and the mantle zone, differentiated neurons.



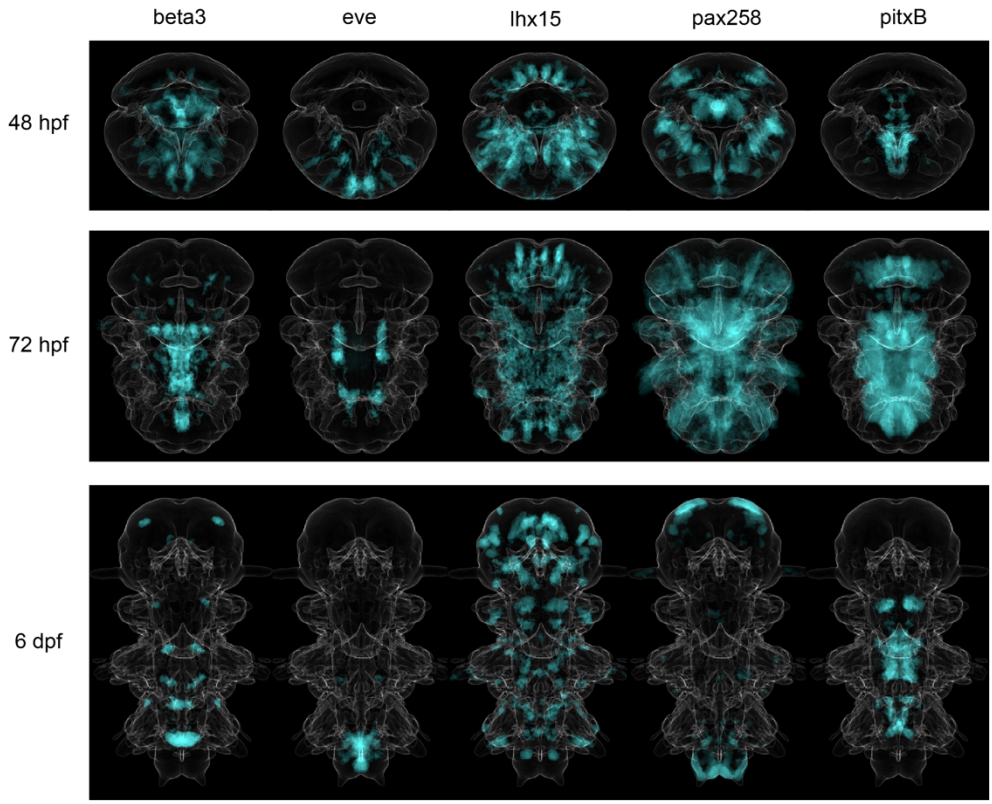
Supplementary figure 2. Expression patterns of the markers of the vertebrate p0 cell type

These expression panels were generated using the ProSPr atlases of gene expression at 48 hpf, 72 hpf, and 6 dpf. The show the expression domains of Pax6, Dbx1, Ngn, Irx6, and Olig, markers of the p0 cell type of the vertebrate neural tube.



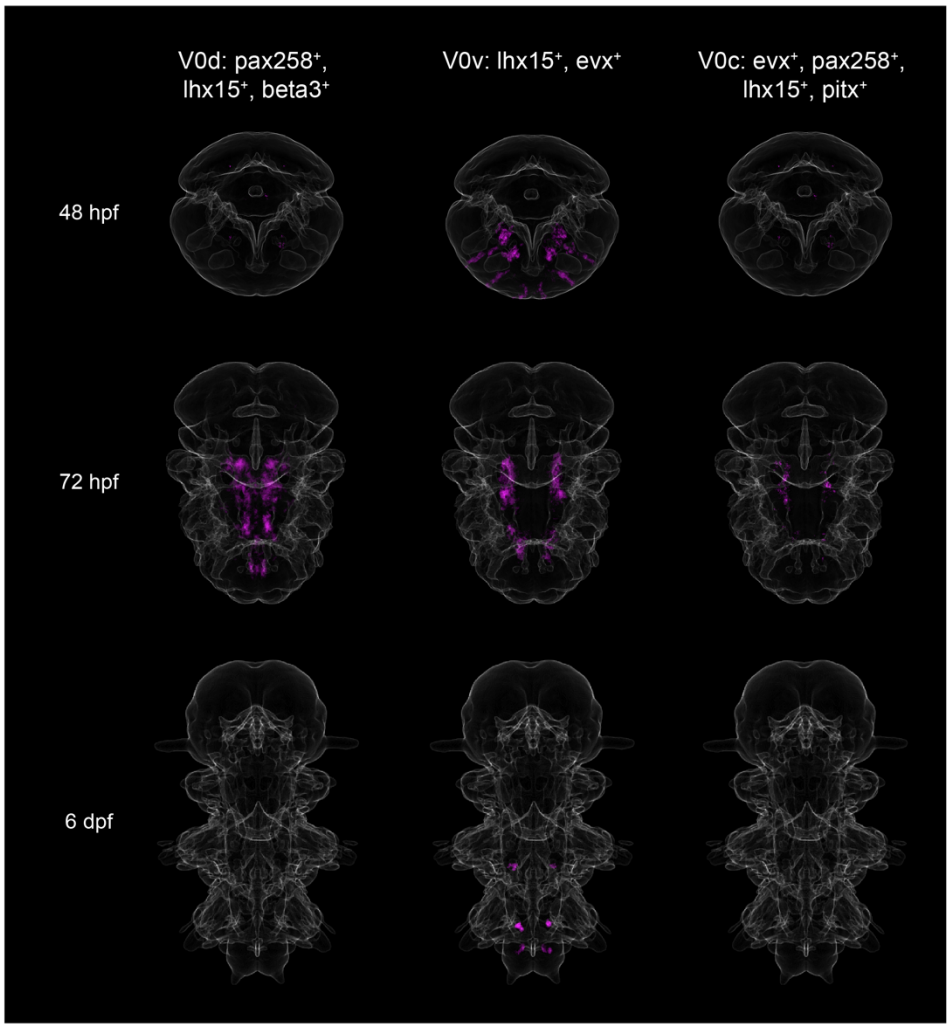
Supplementary figure 3 Coexpression panels of the markers of the vertebrate p0 cell type

These expression panels were generated using the ProSPr atlases of gene expression at 48 hpf, 72 hpf, and 6 dpf. The show the coexpression of *Pax6, Dbx1, Ngn,* and *Irx6* in different combinations.



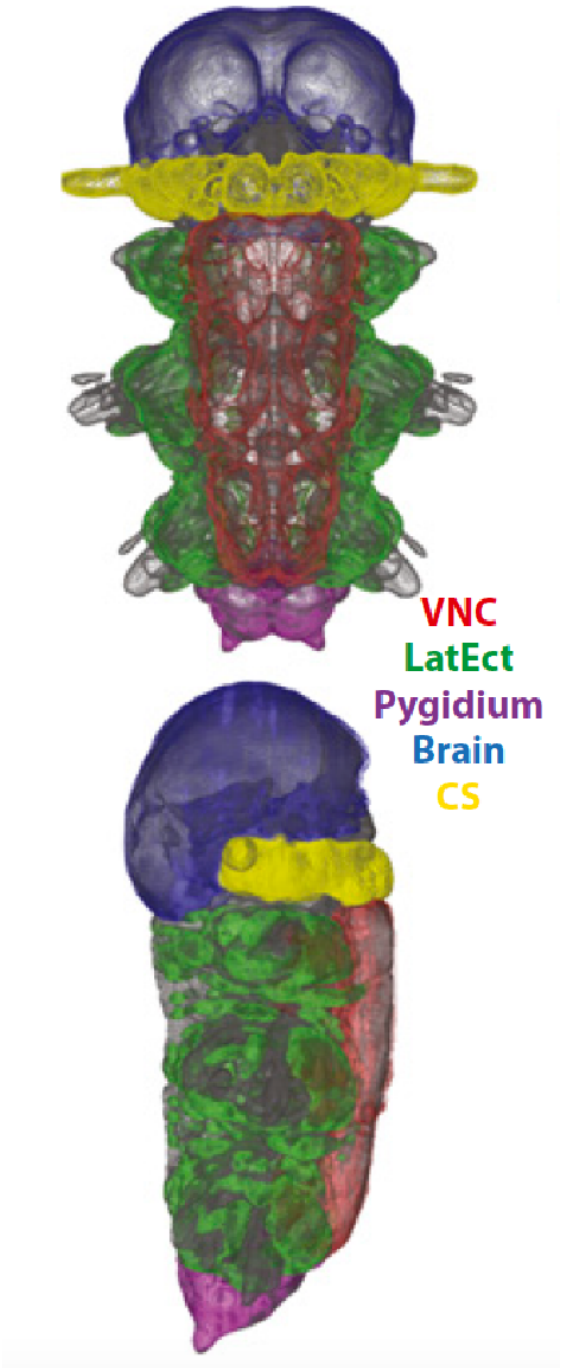
Supplementary figure 4. Expression patterns of the markers of the vertebrate V0 cell types

These expression panels were generated using the ProSPr atlases of gene expression at 48 hpf, 72 hpf, and 6 dpf. The show the expression domains of *Beta3, Eve, Lhx1/5, Pax2/5/8,* and *PitxB*, markers of the V0 cell types of the vertebrate neural tube.



Supplementary figure 5. Coexpression panels of the markers of the vertebrate V0 cell types

These expression panels were generated using the ProSPr atlases of gene expression at 48 hpf, 72 hpf, and 6 dpf. The show the coexpression of *Pax2/5/8, Lhx1/5, Beta3,* and *Pitx* in different combinations.



Supplementary figure 6. Manually segmented tissues on the 6 dpf ProSPr atlas. Reproduced from Ref. [56]

The top panel shows a ventral view and the bottom panel shows a side view with the ventral side on the right. Some tissues were manually segmented by Hernando Martinez and are shown in different colours. Our study focuses on the VNC.

Supplementary table 3. . Quantification of the commissural axons in the *Dbx1* and Neo/DCC mutants, controls, and WTs

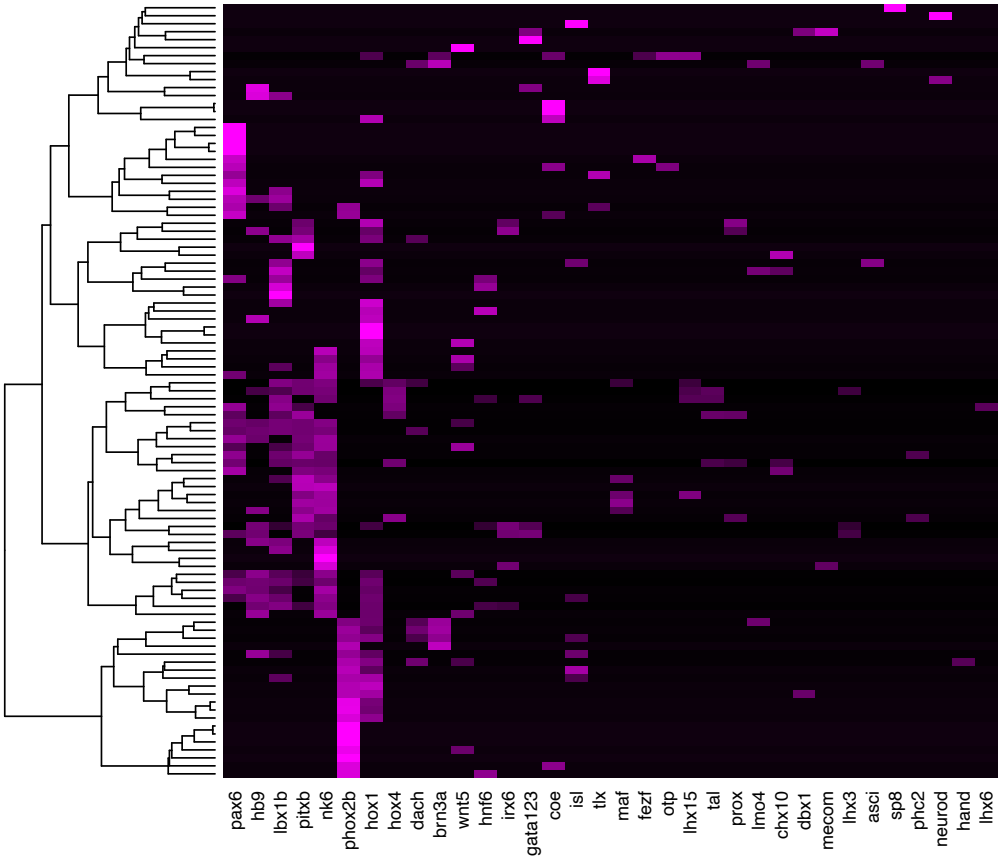
The number of commissural axons for every larva used in the functional analysis of *Dbx1* and Neo/DCC is presented in this table. The larvae are classified according the treatment they received into three categories: control (Ctrl), mutant (KO), or wildtype (WT) (see Materials and Methods, section IV.12). The batch column shows which larvae are siblings; all larvae with the same batch name are siblings.

Larva ID	# Commissural axons	Treatment	Batch
2_I_Ctrl	29	Ctrl	Neo/DCC_I
4_I_Ctrl	24	Ctrl	Neo/DCC_I
7_I_Ctrl	23	Ctrl	Neo/DCC_I
1_I_KO	21	KO	Neo/DCC_I
2_I_KO	14	KO	Neo/DCC_I
5_I_KO	15	KO	Neo/DCC_I
7_I_KO	18	KO	Neo/DCC_I
9_I_KO	21	KO	Neo/DCC_I
10_I_KO	19	KO	Neo/DCC_I
11_I_KO	16	KO	Neo/DCC_I
14_I_KO	23	KO	Neo/DCC_I
17_I_KO	15	KO	Neo/DCC_I
19_I_KO	20	KO	Neo/DCC_I
1_I_WT	22	WT	Neo/DCC_I
3_I_WT	29	WT	Neo/DCC_I
4_I_WT	28	WT	Neo/DCC_I
5_I_WT	28	WT	Neo/DCC_I
1_II_Ctrl	32	Ctrl	Neo/DCC_II
3_II_Ctrl	33	Ctrl	Neo/DCC_II
4_II_Ctrl	26	Ctrl	Neo/DCC_II
6_II_Ctrl	26	Ctrl	Neo/DCC_II
7_II_Ctrl	29	Ctrl	Neo/DCC_II
8_II_Ctrl	23	Ctrl	Neo/DCC_II
10_II_Ctrl	21	Ctrl	Neo/DCC_II
11_II_Ctrl	31	Ctrl	Neo/DCC_II
12_II_Ctrl	26	Ctrl	Neo/DCC_II
13_II_Ctrl	25	Ctrl	Neo/DCC_II
14_II_Ctrl	22	Ctrl	Neo/DCC_II
1_II_KO	5	KO	Neo/DCC_II
2_II_KO	29	KO	Neo/DCC_II
3_II_KO	19	KO	Neo/DCC_II
4_II_KO	21	KO	Neo/DCC_II
5_II_KO	24	KO	Neo/DCC_II
7_II_KO	25	KO	Neo/DCC_II
8_II_KO	29	KO	Neo/DCC_II

9_II_KO	23	KO	Neo/DCC_II
10_II_KO	25	KO	Neo/DCC_II
11_II_KO	12	KO	Neo/DCC_II
12_II_KO	21	KO	Neo/DCC_II
1_II_WT	40	WT	Neo/DCC_II
2_II_WT	18	WT	Neo/DCC_II
3_II_WT	32	WT	Neo/DCC_II
4_II_WT	36	WT	Neo/DCC_II
5_II_WT	41	WT	Neo/DCC_II
7_II_WT	17	WT	Neo/DCC_II
8_II_WT	45	WT	Neo/DCC_II
9_II_WT	45	WT	Neo/DCC_II
10_II_WT	28	WT	Neo/DCC_II
11_II_WT	33	WT	Neo/DCC_II
12_II_WT	38	WT	Neo/DCC_II
1_III_Ctrl	26	Ctrl	Neo/DCC_III
2_III_Ctrl	27	Ctrl	Neo/DCC_III
3_III_Ctrl	21	Ctrl	Neo/DCC_III
4_III_Ctrl	17	Ctrl	Neo/DCC_III
5_III_Ctrl	29	Ctrl	Neo/DCC_III
7_III_Ctrl	25	Ctrl	Neo/DCC_III
8_III_Ctrl	36	Ctrl	Neo/DCC_III
11_III_Ctrl	29	Ctrl	Neo/DCC_III
16_III_Ctrl	23	Ctrl	Neo/DCC_III
19_III_Ctrl	29	Ctrl	Neo/DCC_III
20_III_Ctrl	25	Ctrl	Neo/DCC_III
21_III_Ctrl	33	Ctrl	Neo/DCC_III
22_III_Ctrl	16	Ctrl	Neo/DCC_III
23_III_Ctrl	22	Ctrl	Neo/DCC_III
25_III_Ctrl	34	Ctrl	Neo/DCC_III
28_III_Ctrl	33	Ctrl	Neo/DCC_III
29_III_Ctrl	26	Ctrl	Neo/DCC_III
30_III_Ctrl	23	Ctrl	Neo/DCC_III
1_III_KO	26	KO	Neo/DCC_III
2_III_KO	29	KO	Neo/DCC_III
3_III_KO	17	KO	Neo/DCC_III
4_III_KO	27	KO	Neo/DCC_III
5_III_KO	20	KO	Neo/DCC_III
7_III_KO	6	KO	Neo/DCC_III
8_III_KO	29	KO	Neo/DCC_III
10_III_KO	24	KO	Neo/DCC_III

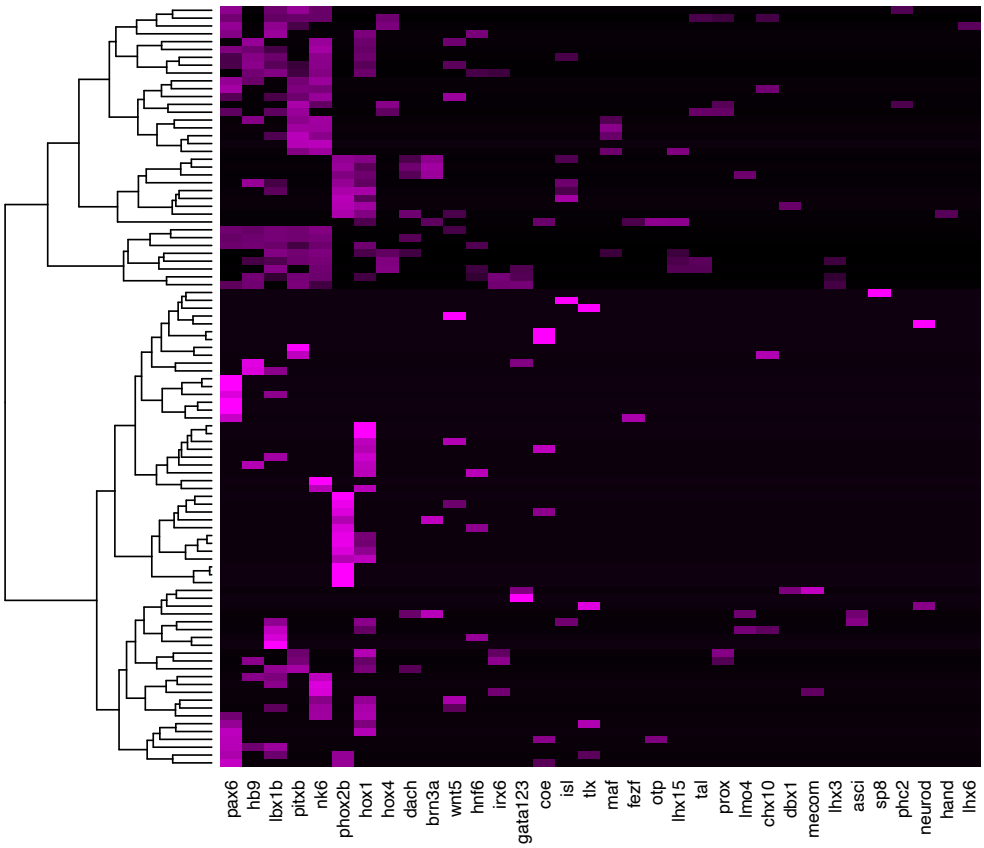
11_III_KO	28	KO	Neo/DCC_III
12_III_KO	17	KO	Neo/DCC_III
14_III_KO	27	KO	Neo/DCC_III
15_III_KO	28	KO	Neo/DCC_III
16_III_KO	15	KO	Neo/DCC_III
18_III_KO	20	KO	Neo/DCC_III
19_III_KO	17	KO	Neo/DCC_III
20_III_KO	17	KO	Neo/DCC_III
21_III_KO	24	KO	Neo/DCC_III
22_III_KO	19	KO	Neo/DCC_III
23_III_KO	23	KO	Neo/DCC_III
25_III_KO	23	KO	Neo/DCC_III
1_III_WT	37	WT	Neo/DCC_III
2_III_WT	30	WT	Neo/DCC_III
4_III_WT	41	WT	Neo/DCC_III
5_III_WT	37	WT	Neo/DCC_III
6_III_WT	24	WT	Neo/DCC_III
7_III_WT	33	WT	Neo/DCC_III
8_III_WT	39	WT	Neo/DCC_III
9_III_WT	28	WT	Neo/DCC_III
10_III_WT	38	WT	Neo/DCC_III
11_III_WT	41	WT	Neo/DCC_III
12_III_WT	31	WT	Neo/DCC_III
21_Ctrl_I	43	Ctrl	*Dbx1*_I
20_Ctrl_I	40	Ctrl	*Dbx1*_I
19_Ctrl_I	45	Ctrl	*Dbx1*_I
18_Ctrl_I	45	Ctrl	*Dbx1*_I
17_Ctrl_I	36	Ctrl	*Dbx1*_I
16_Ctrl_I	49	Ctrl	*Dbx1*_I
10_Ctrl_I	55	Ctrl	*Dbx1*_I
7_Ctrl_I	50	Ctrl	*Dbx1*_I
5_Ctrl_I	43	Ctrl	*Dbx1*_I
3_Ctrl_I	34	Ctrl	*Dbx1*_I
13_KO_I	32	KO	*Dbx1*_I
11_KO_I	32	KO	*Dbx1*_I
9_KO_I	25	KO	*Dbx1*_I
5_KO_I	32	KO	*Dbx1*_I
2_KO_I	24	KO	*Dbx1*_I
25_KO_I	50	KO	*Dbx1*_I
23_KO_I	27	KO	*Dbx1*_I
21_KO_I	34	KO	*Dbx1*_I

18_KO_I	35	KO	*Dbx1*_I
15_KO_I	28	KO	*Dbx1*_I
14_KO_II	40	KO	*Dbx1*_II
11_KO_II	31	KO	*Dbx1*_II
8_KO_II	26	KO	*Dbx1*_II
6_KO_II	24	KO	*Dbx1*_II
4_KO_II	34	KO	*Dbx1*_II
14_WT_I	40	WT	*Dbx1*_I
12_WT_I	54	WT	*Dbx1*_I
8_WT_I	54	WT	*Dbx1*_I
6_WT_I	45	WT	*Dbx1*_I
4_WT_I	59	WT	*Dbx1*_I
18_WT_II	59	WT	*Dbx1*_II
16_WT_II	54	WT	*Dbx1*_II
10_WT_II	48	WT	*Dbx1*_II
8_WT_II	53	WT	*Dbx1*_II
14_WT_III	44	WT	*Dbx1*_III
12_WT_III	45	WT	*Dbx1*_III
10_WT_III	43	WT	*Dbx1*_III
8_WT_III	36	WT	*Dbx1*_III
6_WT_III	43	WT	*Dbx1*_III
30_WT_IV	44	WT	*Dbx1*_IV
28_WT_IV	42	WT	*Dbx1*_IV
24_WT_IV	34	WT	*Dbx1*_IV
22_WT_IV	41	WT	*Dbx1*_IV
20_WT_IV	49	WT	*Dbx1*_IV
13_WT_IV	43	WT	*Dbx1*_IV
11_WT_IV	32	WT	*Dbx1*_IV
8_WT_IV	48	WT	*Dbx1*_IV
5_WT_IV	53	WT	*Dbx1*_IV
2_WT_IV	47	WT	*Dbx1*_IV



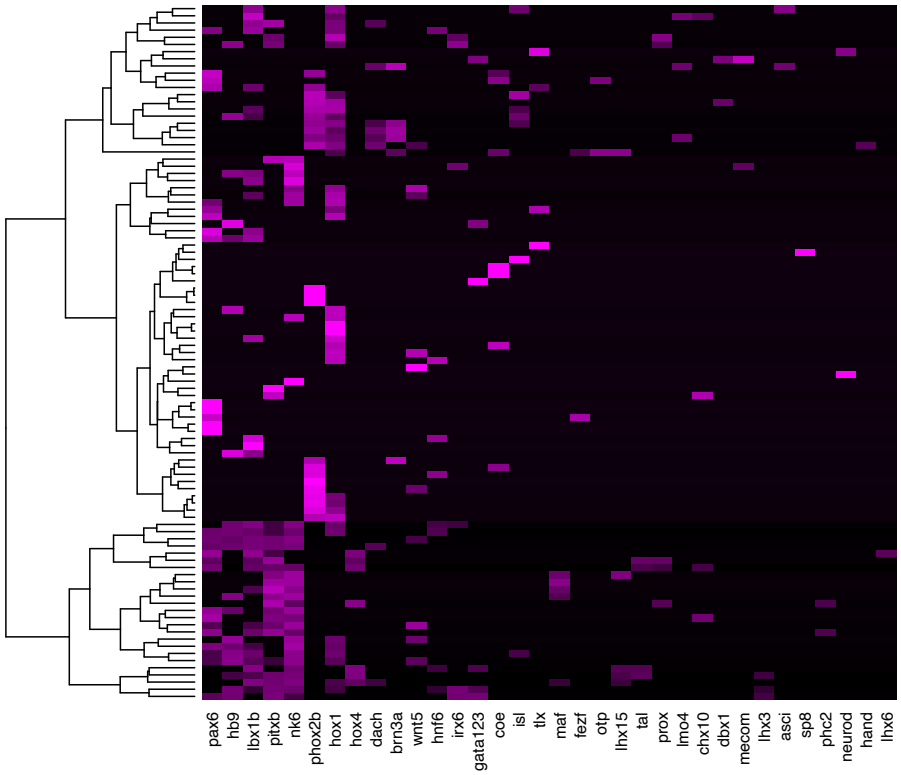
Supplementary figure 7. Clustering of the commissural neurons based on their TF-expression profile – Jaccard

We extracted the overlap between TF expression and the segmented nuclear volume (overlap values) of the commissural neurons using the PlatyBrowser. Since overlap values under 0.1 comprise the vast majority (Figure II-2) and we consider that they indicate that a gene is not expressed in that cell, we transformed every overlap value below 0.1 to 0. The clustering was computed using the Jaccard distance measure. The scale is a gradient between black (overlap value = 0) and magenta (overlap value = 1). Graph created with the R package 'ComplexHeatmap'.



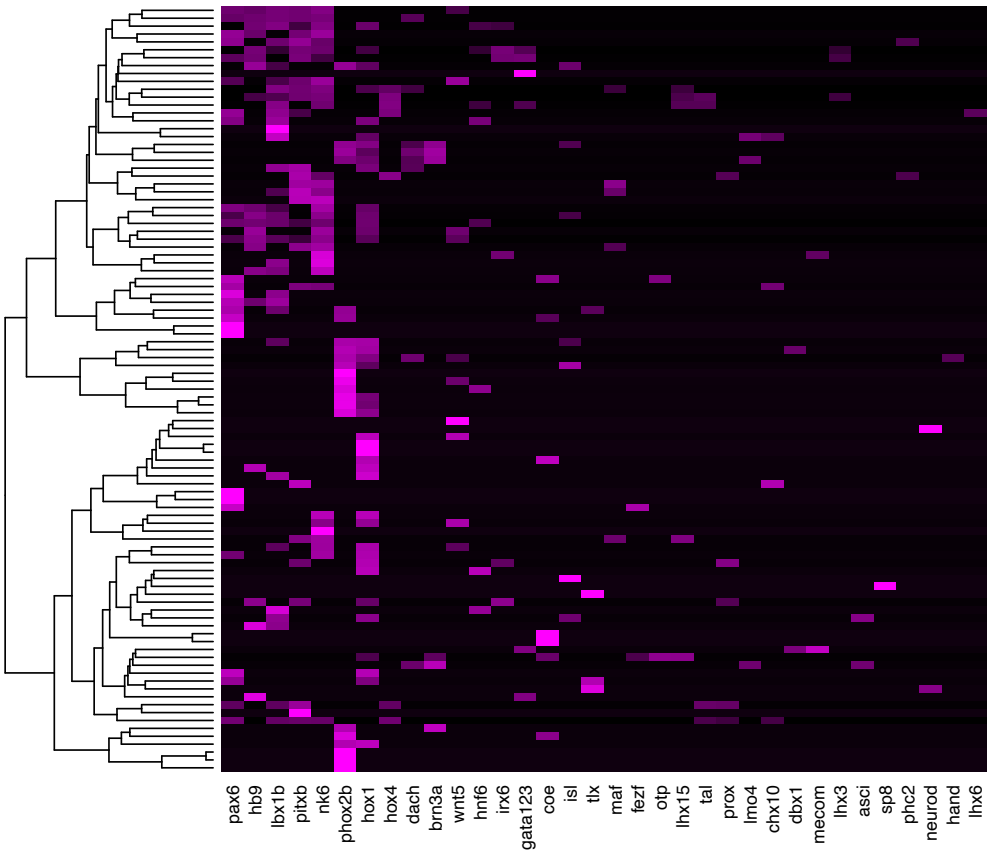
Supplementary figure 8. Clustering of the commissural neurons based on their TF-expression profile – Euclidean

We extracted the overlap between TF expression and the segmented nuclear volume (overlap values) of the commissural neurons using the PlatyBrowser. Since overlap values under 0.1 comprise the vast majority (Figure II-2) and we consider that they indicate that a gene is not expressed in that cell, we transformed every overlap value below 0.1 to 0. The clustering was computed using the Euclidean distance measure. The scale is a gradient between black (overlap value = 0) and magenta (overlap value = 1). Graph created with the R package 'ComplexHeatmap'.



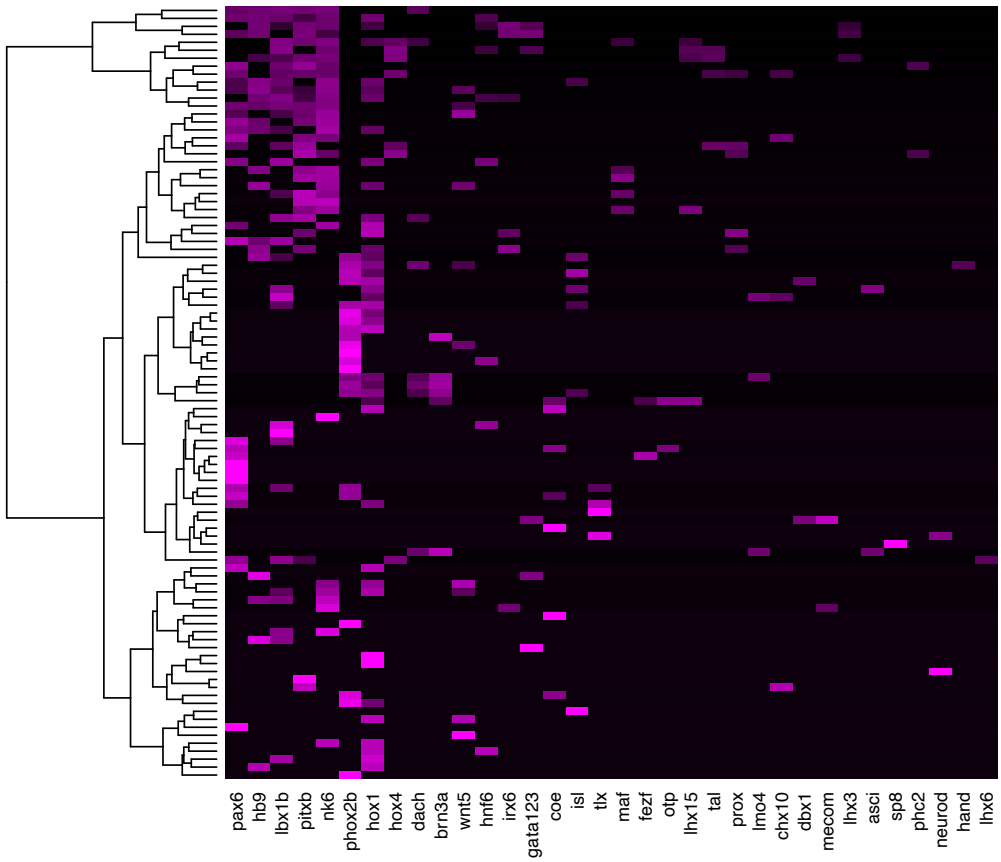
Supplementary figure 9. Clustering of the commissural neurons based on their TF-expression profile – Manhattan

We extracted the overlap between TF expression and the segmented nuclear volume (overlap values) of the commissural neurons using the PlatyBrowser. Since overlap values under 0.1 comprise the vast majority (Figure II-2) and we consider that they indicate that a gene is not expressed in that cell, we transformed every overlap value below 0.1 to 0. The clustering was computed using the Manhattan distance measure. The scale is a gradient between black (overlap value = 0) and magenta (overlap value = 1). Graph created with the R package 'ComplexHeatmap'.



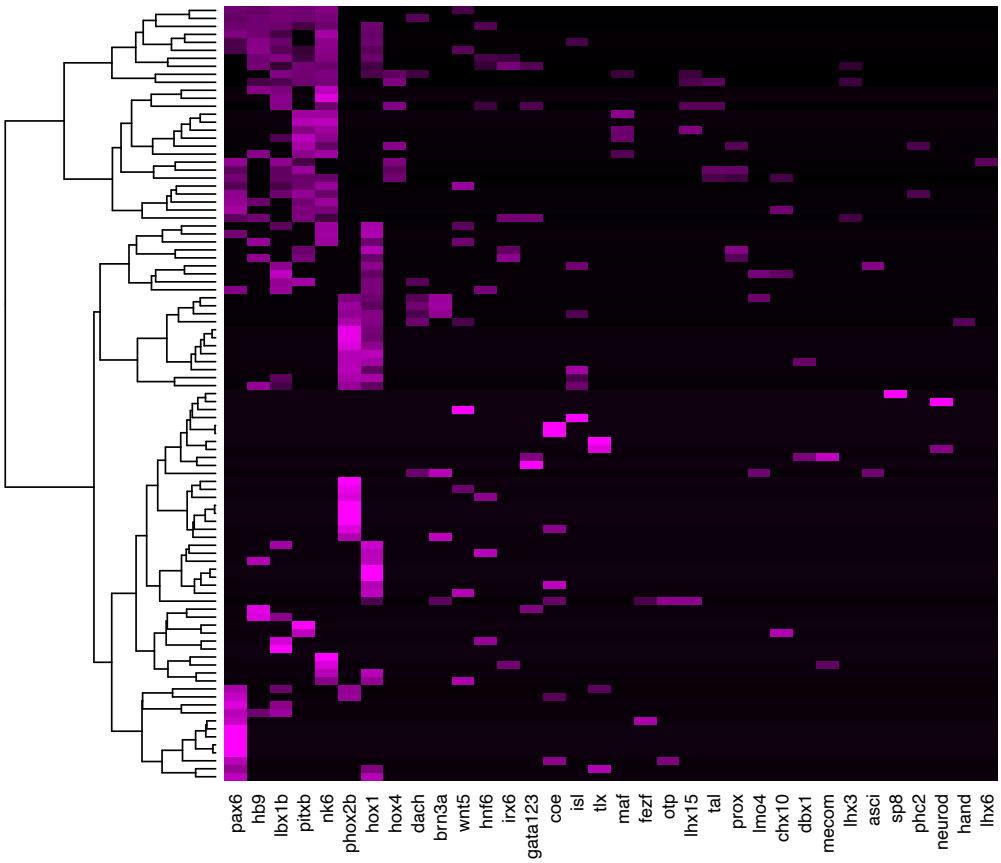
Supplementary figure 10. Clustering of the commissural neurons based on their TF-expression profile – Chebyshev

We extracted the overlap between TF expression and the segmented nuclear volume (overlap values) of the commissural neurons using the PlatyBrowser. Since overlap values under 0.1 comprise the vast majority (Figure II-2) and we consider that they indicate that a gene is not expressed in that cell, we transformed every overlap value below 0.1 to 0. The clustering was computed using the Chebyshev distance measure. The scale is a gradient between black (overlap value = 0) and magenta (overlap value = 1). Graph created with the R package 'ComplexHeatmap'.



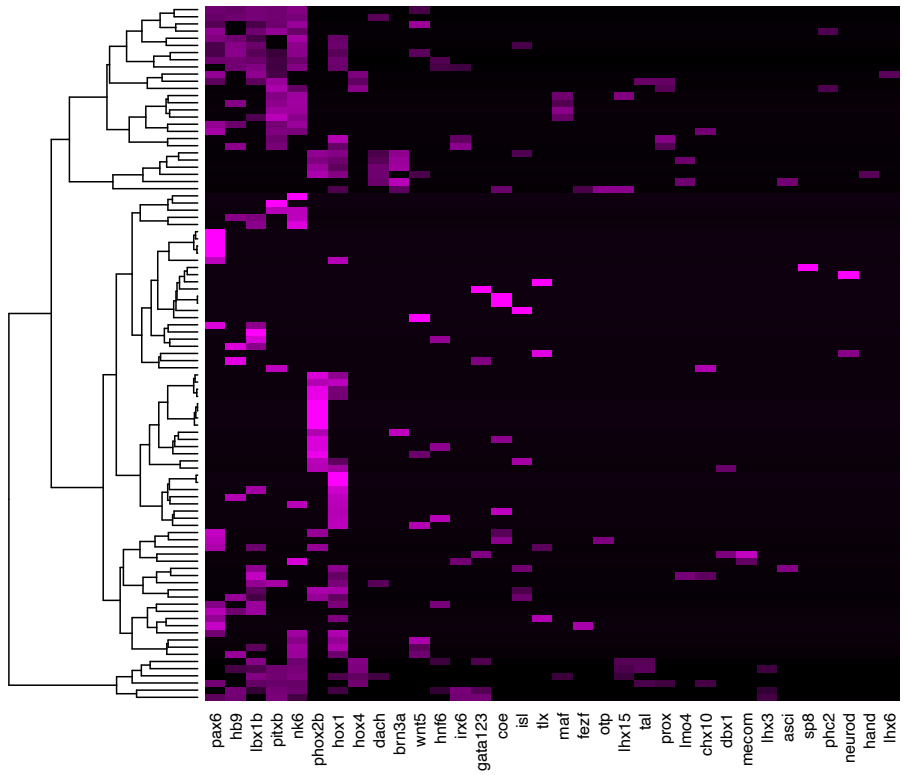
Supplementary figure 11. Clustering of the commissural neurons based on their TF-expression profile – Pearson

We extracted the overlap between TF expression and the segmented nuclear volume (overlap values) of the commissural neurons using the PlatyBrowser. Since overlap values under 0.1 comprise the vast majority (Figure II-2) and we consider that they indicate that a gene is not expressed in that cell, we transformed every overlap value below 0.1 to 0. The clustering was computed using Pearson correlation. The scale is a gradient between black (overlap value = 0) and magenta (overlap value = 1). Graph created with the R package 'ComplexHeatmap'.



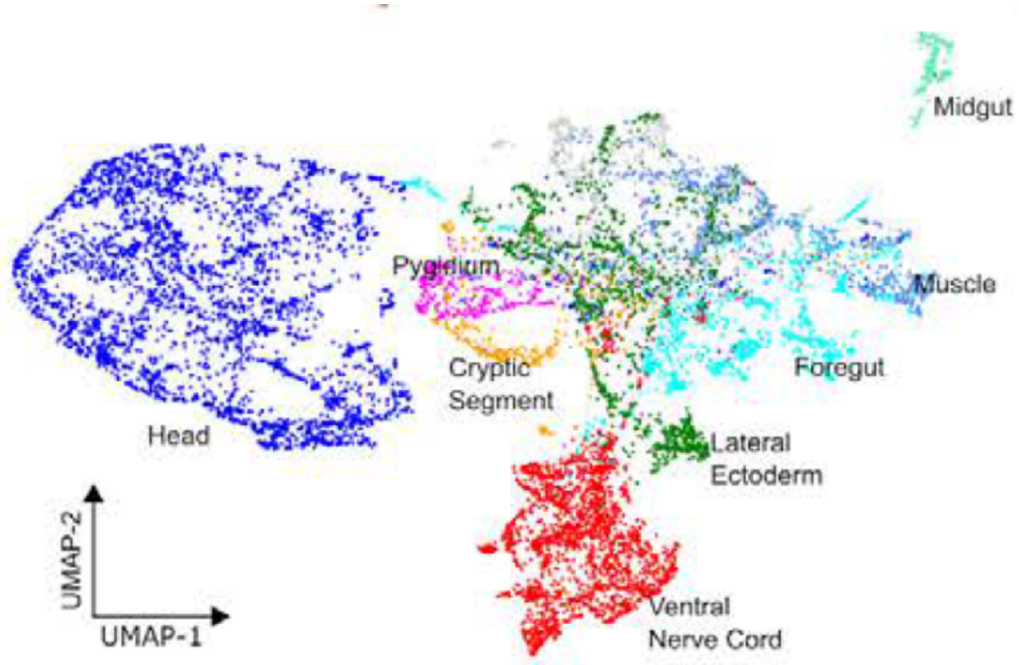
Supplementary figure 12. Clustering of the commissural neurons based on their TF-expression profile – Intersection

We extracted the overlap between TF expression and the segmented nuclear volume (overlap values) of the commissural neurons using the PlatyBrowser. Since overlap values under 0.1 comprise the vast majority (Figure II-2) and we consider that they indicate that a gene is not expressed in that cell, we transformed every overlap value below 0.1 to 0. The clustering was computed using the Intersection distance measure. The scale is a gradient between black (overlap value = 0) and magenta (overlap value = 1). Graph created with the R package 'ComplexHeatmap'.



Supplementary figure 13. Clustering of the commissural neurons based on their TF-expression profile – Canberra

We extracted the overlap between TF expression and the segmented nuclear volume (overlap values) of the commissural neurons using the PlatyBrowser. Since overlap values under 0.1 comprise the vast majority (Figure II-2) and we consider that they indicate that a gene is not expressed in that cell, we transformed every overlap value below 0.1 to 0. The clustering was computed using the Canberra distance measure. The scale is a gradient between black (overlap value = 0) and magenta (overlap value = 1). Graph created with the R package 'ComplexHeatmap'.



Supplementary figure 14. UMAP of all the segmented nuclei of the SBF-SEM dataset coloured by anatomical regions/tissues (Reproduced from Ref. [150].)

UMAP of all of the segmented nuclei of the SBF-SEM dataset, computed by Kimberly Meechan and published in Ref. [150]. It was calculated using the overlap in volume between the segmented nuclei and the expression domains of every gene in the PlatyBrowser. The anatomical regions or tissues in this graph were found through automatic segmentation followed by their manual curation[150].

VI. BIBLIOGRAPHY

1. Hyman, L. H. *The invertebrates: Echinodermata. The coelomate bilateria*. **4**, (McGraw-Hill, 1940).
2. Jékely, G., Paps, J. & Nielsen, C. The phylogenetic position of ctenophores and the origin(s) of nervous systems. *Evodevo* **6**, (2015).
3. Ryan, J. F. *et al.* The genome of the ctenophore Mnemiopsis leidyi and its implications for cell type evolution. *Science (80-.).* **342**, (2013).
4. Lichtneckert, R. & Reichert, H. Origin and evolution of the first nervous system. *Evol. Nerv. Syst.* **1**, 289–315 (2007).
5. Colgren, J. & Nichols, S. A. The significance of sponges for comparative studies of developmental evolution. *Wiley Interdiscip. Rev. Dev. Biol.* **9**, 1–16 (2020).
6. Marlow, H. & Arendt, D. Evolution: Ctenophore genomes and the origin of neurons. *Curr. Biol.* **24**, R757–R761 (2014).
7. Kin, K., Nnamani, M. C., Lynch, V. J., Michaelides, E. & Wagner, G. P. Cell-type phylogenetics and the origin of endometrial stromal cells. *Cell Rep.* **10**, 1398–1409 (2015).
8. Gans, C. & Northcutt, R. G. Neural Crest and the Origin of Vertebrates: A New Head. *Science (80-.).* **220**, 268–273 (1983).
9. Vickaryous, M. K. & Hall, B. K. Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. *Biol. Rev. Camb. Philos. Soc.* **81**, 425–455 (2006).
10. Arendt, D. *et al.* The origin and evolution of cell types. *Nat. Rev. Genet.* **17**, 744–757 (2016).
11. Arendt, D. Evolution of eyes and photoreceptor cell types. *Int. J. Dev. Biol.* **47**, 563–571 (2003).
12. Arendt, D., Hausen, H. & Purschke, G. The 'division of labour' model of eye evolution. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 2809–2817 (2009).
13. Marioni, J. C. & Arendt, D. How Single-Cell Genomics Is Changing Evolutionary and Developmental Biology. *Annu. Rev. Cell Dev. Biol.* **33**, 537–553 (2017).
14. Lee, S. & Pfaff, S. L. Synchronization of Neurogenesis and Motor Neuron Specification by Direct Coupling of bHLH and Homeodomain Transcription Factors. **38**, 731–745 (2003).
15. Thaler, J. P., Lee, S. K., Jurata, L. W., Gill, G. N. & Pfaff, S. L. LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein-protein interactions. *Cell* **110**, 237–249 (2002).
16. Rosenberg, A. B. *et al.* Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science (80-.).* **360**, 176–182 (2018).
17. Hu, G., Li, J. & Wang, G. Z. Significant Evolutionary Constraints on Neuron Cells Revealed by Single-Cell Transcriptomics. *Genome Biol. Evol.* **12**, 300–308 (2020).
18. Tosches, M. A. *et al.* Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles. *Science (80-.).* **360**, 881–888 (2018).
19. Achim, K. *et al.* Whole-body single-cell sequencing reveals transcriptional domains in the annelid larval body. *Mol. Biol. Evol.* **35**, 1047–1062 (2018).
20. Sebé-Pedrós, A. *et al.* Cnidarian Cell Type Diversity and Regulation Revealed by Whole-Organism Single-Cell RNA-Seq. *Cell* **173**, 1520-1534.e20 (2018).

21. Erwin, D. H. & Davidson, E. H. The evolution of hierarchical gene regulatory networks. *Nat. Rev. Genet.* **10**, 141–148 (2009).

22. Graf, T. & Enver, T. Forcing cells to change lineages. *Nature* **462**, 587–594 (2009).

23. Arendt, D., Bertucci, P. Y., Achim, K. & Musser, J. M. Evolution of neuronal types and families. *Curr. Opin. Neurobiol.* **56**, 144–152 (2019).

24. Erwin, D. H. & Davidson, E. H. The last common bilaterian ancestor. *Development* **129**, 3021–3032 (2002).

25. Erwin, D. H. Early metazoan life: Divergence, environment and ecology. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, (2015).

26. Haeckel, E. Die Gastraea-theorie, die phylogenetische classification des thierreichs und die homologie der keimblatter. *Jenaische Z f Naturwiss* **8**, 1–55 (1874).

27. Arendt, D., Technau, U. & Wittbrodt, J. Evolution of the bilaterian larval foregut. *Nature* **409**, 81–85 (2001).

28. Arendt, D. & Nübler-Jung, K. Dorsal or ventral: similarities in fate maps and gastrulation patterns in annelids, arthropods and chrodates. *Mech. Dev.* **61**, 7–21 (1997).

29. Steinmetz, P. R. H., Zelada-Gonzáles, F., Burgtorf, C., Wittbrodt, J. & Arendt, D. Polychaete trunk neuroectoderm converges and extend by mediolateral cell intercalation. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2727–2732 (2007).

30. Hertwig, O. & Hertwig, R. Studien zur Blättertheorie. Heft 1: Die Actinien. (1879).

31. Arendt, D., Tosches, M. A. & Marlow, H. From nerve net to nerve ring, nerve cord and brain-evolution of the nervous system. *Nat. Rev. Neurosci.* **17**, 61–72 (2016).

32. Sedgwick, A. Memoirs: on the origin metameric segmentation and some other morphological question. *J. Cell Sci.* **2**, 43–82 (1884).

33. Balfour, F. M. *A treatise on comparative embryology*. **1**, (Macmillan and Company, 1880).

34. Koizumi, O. Nerve ring of the hypostome in hydra: is it an origin of the central nervous system of bilaterian animals? *Brain. Behav. Evol.* **69**, 151–159 (2007).

35. Galliot, B. & Quiquand, M. A two-step process in the emergence of neurogenesis. *Eur. J. Neurosci.* **34**, 847–862 (2011).

36. Tosches, M. A. & Arendt, D. The bilaterian forebrain: An evolutionary chimaera. *Curr. Opin. Neurobiol.* **23**, 1080–1089 (2013).

37. Arendt, D., Benito-Gutierrez, E., Brunet, T. & Marlow, H. Gastric pouches and the mucociliary sole: Setting the stage for nervous system evolution. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, (2015).

38. Hyman, L. H. The invertebrates: Platyhelminthes and Rhynchocoela, the acoelomate Bilateria. *Invertebr. Platyhelminthes Rhynchocoela, acoelomate Bilateria.* **2**, (1951).

39. Martindale, M. Q. & Hejnol, A. A Developmental Perspective: Changes in the Position of the Blastopore during Bilaterian Evolution. *Dev. Cell* **17**, 162–174 (2009).

40. Presnell, J. S. *et al.* The Presence of a Functionally Tripartite Through-Gut in Ctenophora Has Implications for Metazoan Character Trait Evolution. *Curr. Biol.*

26, 2814–2820 (2016).

41. Holstein, T. W. The evolution of the wnt pathway. *Cold Spring Harb. Perspect. Biol.* **4**, 1–17 (2012).

42. Meinhardt, H. Different strategies for midline formation in bilaterians. *Nat. Rev. Neurosci.* **5**, 502–510 (2004).

43. Nielsen, C., Brunet, T. & Arendt, D. Evolution of the bilaterian mouth and anus. *Nat. Ecol. Evol.* **2**, 1358–1376 (2018).

44. Arendt, D., Denes, A. S., Jékely, G. & Tessmar-Raible, K. The evolution of nervous system centralization. *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 1523–1528 (2008).

45. Graham, V., Khudyakov, J., Ellis, P. & Pevny, L. SOX2 functions to maintain neural progenitor identity. *Neuron* **39**, 749–765 (2003).

46. Denes, A. S. *et al.* Molecular Architecture of Annelid Nerve Cord Supports Common Origin of Nervous System Centralization in Bilateria. *Cell* **129**, 277–288 (2007).

47. Skeath, J. B. At the nexus between pattern formation and cell-type specification: The generation of individual neuroblast fates in the drosophila embryonic central nervous system. *BioEssays* **21**, 922–931 (1999).

48. Arendt, D. Dispatches Animal Evolution : Convergent Nerve Cords ? *Curr. Biol.* **28**, R225–R227 (2018).

49. Martín-durán, J. M. *et al.* Convergent evolution of bilaterian nerve cords. *Nat. Publ. Gr.* **553**, 45–50 (2018).

50. Wilson, L. & T, M. M. The mechanisms of dorsoventral patterning in the vertebrate neural tube. **282**, 1–13 (2005).

51. Seredick, S. D., Van Ryswyk, L., Hutchinson, S. A. & Eisen, J. S. Zebrafish Mnx proteins specify one motoneuron subtype and suppress acquisition of interneuron characteristics. *Neural Dev.* **7**, 1 (2012).

52. Ferrier, D. E. K., Brooke, N. M., Panopoulou, G. & Holland, P. W. H. The Mnx homeobox gene class defined by HB9, MNR2 and amphioxus AmphiMnx. *Dev. Genes Evol.* **211**, 103–107 (2001).

53. Hudson, C., Ba, M., Rouvière, C. & Yasuo, H. Divergent mechanisms specify chordate motoneurons: Evidence from ascidians. *Development* **138**, 1643–1652 (2011).

54. Odden, J. P., Holbrook, S. & Doe, C. Q. Drosophila HB9 is expressed in a subset of motoneurons and interneurons, where it regulates gene expression and axon pathfinding. *J. Neurosci.* **22**, 9143–9149 (2002).

55. Von Stetina, S. E., Treinin, M. & Miller, D. M. The Motor Circuit. *Int. Rev. Neurobiol.* **69**, 125–167 (2005).

56. Martinez-Vergara, H. *et al.* Whole-organism cellular gene-expression atlas reveals conserved cell types in the ventral nerve cord of Platynereis dumerilii. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 5878–5885 (2017).

57. Von Haffner, K. *Uber den Bau und den Zusammenhang der wichtigsten Organe des Kopfendes von Hyalinoecia tubicola Malmgren*. (1959).

58. Hamaker, J. I. The nervous system of Nereis virens Sars: a study in comparative neurology. (1898).

59. Gerberding, M. & Scholtz, G. Neurons and glia in the midline of the higher

crustacean Orchestia cavimana are generated via an invariant cell lineage that comprises a median neuroblast and glial progenitors. *Dev. Biol.* **235**, 397–409 (2001).

60. Mulloney, B. & Hall, W. M. Local Commissural Interneurons Integrate Information from Intersegmental Coordinating Interneurons. *J. Comp. Neurol.* **466**, 366–376 (2003).

61. Seabrook, W. D. An electrophysiological study of the giant fiber system of the locust Schistocerca gregaria. *Can. J. Zool.* **49**, 555–560 (1971).

62. Hustert, R. & Mashaly, A. M. Spontaneous behavioral rhythms in the isolated CNS of insects–Presenting new model systems. *J. Physiol.* **107**, 147–151 (2013).

63. Lacalli, T. C. & Kelly, S. J. Ventral neurons in the anterior nerve cord of amphioxus larvae. I. An inventory of cell types and synaptic patterns. *J. Morphol.* **257**, 190–211 (2003).

64. Buchanan, J. T. & Grillner, S. Newly identified 'glutamate interneurons' and their role in locomotion in the lamprey spinal cord. *Science (80-.).* **236**, 312–314 (1987).

65. Leung, B. & Shimeld, S. M. Evolution of vertebrate spinal cord patterning. *Dev. Dyn.* **248**, 1028–1043 (2019).

66. Goulding, M. Circuits controlling vertebrate locomotion: Moving in a new direction. *Nat. Rev. Neurosci.* **10**, 507–518 (2009).

67. Alaynick, W. A., Jessell, T. M. & Pfaff, S. L. SnapShot: Spinal cord development. *Cell* **146**, 178.e1-178.e1 (2011).

68. Pierani, A. *et al.* Control of interneuron fate in the developing spinal cord by the progenitor homeodomain protein Dbx1. *Neuron* **29**, 367–384 (2001).

69. Lanuza, G. M., Gosgnach, S., Pierani, A., Jessell, T. M. & Goulding, M. Genetic identification of spinal interneurons that coordinate left-right locomotor activity necessary for walking movements. *Neuron* **42**, 375–386 (2004).

70. Zhang, Y. *et al.* Article V3 Spinal Neurons Establish a Robust and Balanced Locomotor Rhythm during Walking. 84–96 (2008). doi:10.1016/j.neuron.2008.09.027

71. Ericson, J. *et al.* Pax6 controls progenitor cell identity and neuronal fate in response to graded Shh signaling. *Cell* **90**, 169–180 (1997).

72. Helms, A. W. & Johnson, J. E. Progenitors of dorsal commissural interneurons are defined by MATH1 expression. *Development* **125**, 919–928 (1998).

73. Rickert, C., Kunz, T., Harris, K. L., Whitington, P. M. & Technau, G. M. Morphological characterization of the entire interneuron population reveals principles of neuromere organization in the Ventral nerve cord of Drosophila. *J. Neurosci.* **31**, 15870–15883 (2011).

74. Rogulja-Ortmann, A., Lüer, K., Seibert, J., Rickert, C. & Technau, G. M. Programmed cell death in the embryonic central nervous system of Drosophila melanogaster. *Development* **134**, 105–116 (2007).

75. Schmidt, H. *et al.* The embryonic central nervous system lineages of Drosophila melanogaster: II. Neuroblast lineages derived from the dorsal part of the nueroectoderm. *Dev. Biol.* **189**, 186–204 (1997).

76. Hummel, T., Schimmelpfeng, K. & Klämbt, C. Commissure formation in the embryonic CNS of Drosophila : I Identification of the required gene functions.

Dev. Biol. **209**, 381–398 (1999).

77. Moran-Rivard, L. *et al.* Evx1 is a postmitotic determinant of V0 interneuron identity in the spinal cord. *Neuron* **29**, 385–399 (2001).

78. Juárez-Morales, J. L., Martinez-De Luna, R. I., Zuber, M. E., Roberts, A. & Lewis, K. E. Zebrafish transgenic constructs label specific neurons in Xenopus laevis spinal cord and identify frog V0v spinal neurons. *Dev. Neurobiol.* **77**, 1007–1020 (2017).

79. Juárez-Morales, J. L. *et al.* Evx1 and Evx2 specify excitatory neurotransmitter fates and suppress inhibitory fates through a Pax2-independent mechanism. *Neural Dev.* **11**, 1–20 (2016).

80. Suster, M. L. *et al.* A novel conserved evx1 enhancer links spinal interneuron morphology and cis-regulation from fish to mammals. *Dev. Biol.* **325**, 422–433 (2009).

81. Heckscher, E. S. *et al.* Even-Skipped+ Interneurons Are Core Components of a Sensorimotor Circuit that Maintains Left-Right Symmetric Muscle Contraction Amplitude. *Neuron* **88**, 314–329 (2015).

82. Duman-Scheel, M. & Patel, N. H. Analysis of molecular marker expression reveals neuronal homology in distantly related arthropods. *Development* **126**, 2327–2334 (1999).

83. Harzsch, S. Ontogeny of the ventral nerve cord in malacostracan crustaceans: A common plan for neuronal development in Crustacea, Hexapoda and other Arthropoda? *Arthropod Struct. Dev.* **32**, 17–37 (2003).

84. Lacin, H. & Truman, J. W. Lineage mapping identifies molecular and architectural similarities between the larval and adult Drosophila central nervous system. *Elife* **5**, 1–28 (2016).

85. Fujioka, M. *et al.* Even-skipped, acting as a repressor, regulates axonal projections in Drosophila. *Development* **130**, 5385–5400 (2003).

86. Inamata, Y. & Shirasaki, R. Dbx1 triggers crucial molecular programs required for midline crossing by midbrain commissural axons. *Dev.* **141**, 1260–1271 (2014).

87. Friocourt, F. & Chédotal, A. The Robo3 receptor, a key player in the development, evolution, and function of commissural systems. *Dev. Neurobiol.* **77**, 876–890 (2017).

88. Tessier-Lavigne, M. & Goodman, C. S. The Molecular Biology of Axon Guidance. *Science (80-.).* **274**, 1123–1133 (1996).

89. Augsburger, A., Schuchardt, A., Hoskins, S., Dodd, J. & Butler, S. Bmps as mediators of roof plate repulsion of commissural neurons. *Neuron* **24**, 127–141 (1999).

90. Butler, S. J. & Dodd, J. A role for BMP heterodimers in roof plate-mediated repulsion of commissural axons. *Neuron* **38**, 389–401 (2003).

91. Kennedy, T. E., Serafini, T., de la Torre, J. & Tessier-Lavigne, M. Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. *Cell* **78**, 425–435 (1994).

92. Keino-masu, K. *et al.* Deleted in Colorectal Cancer Encodes a Netrin Receptor. *Cell* **87**, 175–185 (1996).

93. Rabe Bernhardt, N. *et al.* DCC mediated axon guidance of spinal interneurons is essential for normal locomotor central pattern generator function. *Dev. Biol.*

366, 279–289 (2012).

94. Ly, A. *et al.* DSCAM Is a Netrin Receptor that Collaborates with DCC in Mediating Turning Responses to Netrin-1. *Cell* **133**, 1241–1254 (2008).

95. Liu, G. *et al.* DSCAM functions as a netrin receptor in commissural axon pathfinding. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 2951–2956 (2009).

96. Serafini, T. *et al.* Netrin-1 Is Required for Commissural Axon Guidance in the Developing Vertebrate Nervous System. *Cell* **87**, 1001–1014 (1996).

97. Fazeli, A. *et al.* Phenotype of mice lacking functional Deleted in colorectal cancer (Dcc) gene. *Nature* **386**, 796–804 (1997).

98. Rabe, N., Gezelius, H., Vallstedt, A., Memic, F. & Kullander, K. Netrin-1-dependent spinal interneuron subtypes are required for the formation of left-right alternating locomotor circuitry. *J. Neurosci.* **29**, 15642–15649 (2009).

99. Dominici, C. *et al.* Floor-plate-derived netrin-1 is dispensable for commissural axon guidance. *Nature* **545**, 350–354 (2017).

100. Varadarajan, S. G. *et al.* Netrin1 Produced by Neural Progenitors, Not Floor Plate Cells, Is Required for Axon Guidance in the Spinal Cord. *Neuron* **94**, 790-799.e3 (2017).

101. Charron, F., Stein, E., Jeong, J., McMahon, A. P. & Tessier-Lavigne, M. The morphogen sonic hedgehog is an axonal chemoattractant that collaborates with Netrin-1 in midline axon guidance. *Cell* **113**, 11–23 (2003).

102. Okada, A. *et al.* Boc is a receptor for sonic hedgehog in the guidance of commissural axons. *Nature* **444**, 369–373 (2006).

103. Ruiz de Almodovar, C. *et al.* VEGF Mediates Commissural Axon Chemoattraction through Its Receptor Flk1. *Neuron* **70**, 966–978 (2011).

104. Neuhaus-Follini, A. & Bashaw, G. J. Crossing the embryonic midline: Molecular mechanisms regulating axon responsiveness at an intermediate target. *Wiley Interdiscip. Rev. Dev. Biol.* **4**, 377–389 (2015).

105. Yam, P. T. *et al.* 14-3-3 Proteins Regulate a Cell-Intrinsic Switch from Sonic Hedgehog-Mediated Commissural Axon Attraction to Repulsion after Midline Crossing. *Neuron* **76**, 735–749 (2012).

106. Zou, Y., Stoeckli, E., Chen, H. & Tessier-Lavigne, M. Squeezing axons out of the gray matter: A role for Slit and Semaphorin proteins from midline and ventral spinal cord. *Cell* **102**, 363–375 (2000).

107. Chen, Z., Gore, B. B., Long, H., Ma, L. & Tessier-Lavigne, M. Alternative Splicing of the Robo3 Axon Guidance Receptor Governs the Midline Switch from Attraction to Repulsion. *Neuron* **58**, 325–332 (2008).

108. Sabatier, C. *et al.* The divergent robo family protein Rig-1/Robo3 is a negative regulator of slit responsiveness required for midline crossing by commissural axons. *Cell* **117**, 157–169 (2004).

109. Tuyen Nguyen Ba-Charvet, K. *et al.* Slit2-Mediated Chemorepulsion and Collapse of Developing Forebrain Axons In the spinal cord and peripheral nervous system, axons can be repelled by a number of molecules. Most of these are members of the Semaphorin family of chemorepel. *Neuron* **22**, 463–473 (1999).

110. Nawabi, H. *et al.* A midline switch of receptor processing regulates commissural axon guidance in vertebrates. *Genes Dev.* **24**, 396–410 (2010).

111. Charoy, C. *et al.* Gdnf Activates Midline Repulsion by Semaphorin3B via NCAM during Commissural Axon Guidance. *Neuron* **75**, 1051–1066 (2012).

112. Dickson, B. J. & Zou, Y. Navigating intermediate targets: the nervous system midline. *Cold Spring Harbor perspectives in biology* **2**, 1–17 (2010).

113. Evans, T. A. & Bashaw, G. J. Axon guidance at the midline: of mice and flies. *Current Opinion in Neurobiology* **20**, 79–85 (2010).

114. Ducuing, H., Gardette, T., Pignata, A., Tauszig-Delamasure, S. & Castellani, V. Commissural axon navigation in the spinal cord: A repertoire of repulsive forces is in command. *Semin. Cell Dev. Biol.* **85**, 3–12 (2019).

115. Chédotal, A. Roles of axon guidance molecules in neuronal wiring in the developing spinal cord. *Nat. Rev. Neurosci.* **20**, 380–396 (2019).

116. Fradkin, L. G. *et al.* The Drosophila Wnt5 protein mediates selective axon fasciculation in the embryonic central nervous system. *Dev. Biol.* **272**, 362–375 (2004).

117. Howard, L. F. J., Brown, H. E., Wadsworth, B. C. & Evans, T. A. Midline axon guidance in the Drosophila embryonic central nervous system. *Semin. Cell Dev. Biol.* **85**, 13–25 (2019).

118. Klämbt, C., Jacobs, J. R. & Goodman, C. S. The midline of the drosophila central nervous system: A model for the genetic analysis of cell fate, cell migration, and growth cone guidance. *Cell* **64**, 801–815 (1991).

119. Chan, S. S. Y. *et al.* UNC-40, a C. elegans homolog of DCC (Deleted in Colorectal Cancer), is required in motile cells responding to UNC-6 netrin cues. *Cell* **87**, 187–195 (1996).

120. Finger, J. H. *et al.* The netrin 1 receptors Unc5h3 and Dcc are necessary at multiple choice points for the guidance of corticospinal tract axons. *J. Neurosci.* **22**, 10346–10356 (2002).

121. Merz, D. C., Zheng, H., Killeen, M. T., Krizus, A. & Culotti, J. G. Multiple signaling mechanisms of the UNC-6/netrin receptors UNC-5 and UNC-40/DCC in vivo. *Genetics* **158**, 1071–1080 (2001).

122. Simanton, W. *et al.* Conservation of arthropod midline netrin accumulation revealed with a cross-reactive antibody provides evidence for midline cell homology. *Evol. Dev.* **11**, 260–268 (2009).

123. Cebrià, F. & Newmark, P. A. Planarian homologs of netrin and netrin receptor are required for proper regeneration of the central nervous system and the maintenance of nervous system architecture. *Development* **132**, 3691–3703 (2005).

124. Marlétaz, F., Peijnenburg, K. T. C. A., Goto, T., Satoh, N. & Rokhsar, D. S. A New Spiralian Phylogeny Places the Enigmatic Arrow Worms among Gnathiferans. *Curr. Biol.* **29**, 312-318.e3 (2019).

125. Harris, R., Sabatelli, L. M. & Seeger, M. A. Guidance cues at the Drosophila CNS midline: Identification and characterization of two Drosophila Netrin/UNC-6 homologs. *Neuron* **17**, 217–228 (1996).

126. Brankatschk, M. & Dickson, B. J. Netrins guide Drosophila commissural axons at short range. *Nat. Neurosci.* **9**, 188–194 (2006).

127. Mitchell, K. J. *et al.* Genetic analysis of Netrin genes in Drosophila: Netrins guide CNS commissural axons and peripheral motor axons. *Neuron* **17**, 203–215 (1996).

128. Dorsten, J. N., Kolodziej, P. A. & VanBerkum, M. F. A. Frazzled regulation of myosin II activity in the Drosophila embryonic CNS. *Dev. Biol.* **308**, 120–132 (2007).

129. Furrer, M. P., Kim, S., Wolf, B. & Chiba, A. Robo and Frazzled/DCC mediate dendritic guidance at the CNS midline. *Nat. Neurosci.* **6**, 223–230 (2003).

130. Hiramoto, M. & Hiromi, Y. ROBO directs axon crossing of segmental boundaries by suppressing responsiveness to relocalized Netrin. *Nat. Neurosci.* **9**, 58–66 (2006).

131. Hattori, D., Millard, S. S., Wojtowicz, W. M. & Zipursky, S. L. Dscam-Mediated Cell Recognition Regulates Neural Circuit Formation. *Annu. Rev. Cell Dev. Biol.* **24**, 597–620 (2008).

132. Andrews, G. L. *et al.* Dscam guides embryonic axons by Netrin-dependent and -independent functions. *Development* **135**, 3839–3848 (2008).

133. Zinzen, R. P., Cande, J., Ronshaugen, M., Papatsenko, D. & Levine, M. Evolution of the Ventral Midline in Insect Embryos. *Dev. Cell* **11**, 895–902 (2006).

134. Morita, S., Shiga, Y., Tokishita, S. & Ohta, T. Analysis of spatiotemporal expression and function of the single-minded homolog in the branchiopod crustacean Daphnia magna. *Gene* **555**, 335–345 (2015).

135. Cebrià, F., Guo, T., Jopek, J. & Newmark, P. A. Regeneration and maintenance of the planarian midline is regulated by a slit orthologue. *Dev. Biol.* **307**, 394–406 (2007).

136. Kidd, T. *et al.* Roundabout controls axon crossing of the CNS midline and defines a novel subfamily of evolutionarily conserved guidance receptors. *Cell* **92**, 205–215 (1998).

137. Battye, R., Stevens, A. & Jacobs, J. R. Axon repulsion from the midline of the Drosophila CNS requires slit function. *Development* **126**, 2475–2481 (1999).

138. Xu, Y. Q., Li, X., Zhong, Y. & Zheng, Y. F. Evolution and diversity of axon guidance Robo receptor family genes. *J. Syst. Evol.* **00**, (2020).

139. Seeger, M., Tear, G., Ferres-Marco, D. & Goodman, C. S. Mutations affecting growth cone guidance in drosophila: Genes necessary for guidance toward or away from the midline. *Neuron* **10**, 409–426 (1993).

140. Georgiou, M. & Tear, G. Commissureless is required both in commissural neurones and midline cells for axon guidance across the midline. *Development* **129**, 2947–2956 (2002).

141. Keleman, K. *et al.* Comm Sorts Robo to Control Axon Guidance at the Drosophila Midline. *Cell* **110**, 415–427 (2002).

142. Keleman, K., Ribeiro, C. & Dickson, B. J. Comm function in commissural axon guidance: Cell-autonomous sorting of Robo in vivo. *Nat. Neurosci.* **8**, 156–163 (2005).

143. Yang, L., Garbe, D. S. & Bashaw, G. J. A frazzled/DCC-dependent transcriptional switch regulates midline axon guidance. *Science (80-.).* **324**, 944–947 (2009).

144. Raible, F. *et al.* Evolution: Vertebrate-type intron-rich genes in the marine annelid Platynereis dumerilii. *Science (80-.).* **310**, 1325–1326 (2005).

145. Hui, J. H. L. *et al.* Features of the ancestral bilaterian inferred from Platynereis dumerilii ParaHox genes. **13**, 1–13 (2009).

146. Christodoulou, F. *et al.* Ancient animal microRNAs and the evolution of tissue identity. *Nature* **463**, 1084–1088 (2010).

147. Arendt, D., Tessmar-Raible, K., Snyman, H., Dorresteijn, A. W. & Wittbrodf, J. Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. *Science (80-.).* **306**, 869–871 (2004).

148. Tessmar-raible, K. *et al.* Conserved Sensory-Neurosecretory Cell Types in Annelid and Fish Forebrain : Insights into Hypothalamus Evolution. 1389–1400 (2007). doi:10.1016/j.cell.2007.04.041

149. Dray et al. Hedgehog signaling regulates segment formation in the annelid Platynereis. **329**, 339–343 (2010).

150. Martinez-Vergara, H. M. *et al.* Whole-body integration of gene expression and single-cell morphology. *bioRxiv* 2020.02.26.961037 (2020). doi:10.1101/2020.02.26.961037

151. Deerinck, T. J., Bushong, E. A., Thor, A. & Ellisman, M. H. NCMIR METHODS FOR 3D EM : A NEW PROTOCOL FOR PREPARATION OF BIOLOGICAL SPECIMENS FOR SERIAL BLOCK FACE SCANNING ELECTRON MICROSCOPY. 6–8 (2010).

152. Wanner, A. A., Genoud, C., Masudi, T., Siksou, L. & Friedrich, R. W. Dense EM-based reconstruction of the interglomerular projectome in the zebrafish olfactory bulb. *Nat. Neurosci.* **19**, 816–825 (2016).

153. Wolf, S. *et al.* The Mutex Watershed: Efficient, Parameter-Free Image Partitioning. in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **11208 LNCS**, 571–587 (2018).

154. Rueckert, D. *et al.* Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imaging* **18**, 712–721 (1999).

155. Fischer, A. H. L., Henrich, T. & Arendt, D. The normal development of Platynereis dumerilii (Nereididae, Annelida). *Front. Zool.* **7**, 31 (2010).

156. Martinez-Vergara, H. Descriptive and functional approaches for a system-level understanding of Platynereis dumerilii and the evolution of the locomotor circuits of Bilateris. (2016).

157. Kübler, I. The role of Neogenin/DCC on commissure formation in Platynereis dumerilii. (2018).

158. Lai, H. C., Seal, R. P. & Johnson, J. E. Making sense out of spinal cord somatosensory development. *Development* **143**, 3434–3448 (2016).

159. Al-Mosawie, A., Wilson, J. M. & Brownstone, R. M. Heterogeneity of V2-derived interneurons in the adult mouse spinal cord. *Eur. J. Neurosci.* **26**, 3003–3015 (2007).

160. Alvarez, F. J. *et al.* Postnal phenotype and lacalization of spinal cord V1 derived interneurons. **493**, 177–192 (2010).

161. Imondi, R. & Kaprielian, Z. Commissural axon pathfinding on the contralateral side of the floor plate: A role of B-class ephrins in specifying the dorsoventral position of longitudinally projecting commissural axons. *Development* **128**, 4859–4871 (2001).

162. Wilson, N. H. & Stoeckli, E. T. Sonic Hedgehog regulates its own receptor on postcrossing commissural axons in a glypican1-dependent manner. *Neuron* **79**, 478–491 (2013).

163. Blanchette, C. R., Perrat, P. N., Thackeray, A. & Bénard, C. Y. Glypican is a modulator of netrin-mediated axon guidance. *PLoS Biol.* **13**, 1–29 (2015).

164. Arber, S. *et al.* Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity. *Neuron* **23**, 659–674 (1999).

165. Simionato, E. *et al.* atonal- and achaete-scute-related genes in the annelid Platynereis dumerilii: Insights into the evolution of neural basic-Helix-Loop-Helix genes. *BMC Evol. Biol.* **8**, 1–13 (2008).

166. De Rosa, R., Prud'homme, B. & Balavoine, G. Caudal and even-skipped in the annelid Platynereis dumerilii and the ancestry of posterior growth. *Evol. Dev.* **7**, 574–587 (2005).

167. Marlow, H. *et al.* Larval body patterning and apical organs are conserved in animal evolution. *BMC Biol.* **12**, 1–17 (2014).

168. Backfisch, B. *et al.* Stable transgenesis in the marine annelid Platynereis dumerilii sheds new light on photoreceptor evolution. *Proc. Natl. Acad. Sci.* **110**, 193–198 (2013).

169. Tessmar-Raible, K. G. K. The evolution of sensory and neurosecretory cell types in bilaterian brains. *Fachbereich Biol.* (2004).

170. Lauri, A. *et al.* Development of the Annelid Axochord: Insights into notochord evolution. *Science (80-.).* **345**, 1365–1368 (2014).

171. McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv Prepr. arXiv1802.03426* (2018).

172. Li, Y. *et al.* Conserved gene regulatory module specifies lateral neural borders across bilaterians. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E6352–E6360 (2017).

173. Hummel, T., Schimmelpfeng, K. & Klämbt, C. Commissure formation in the embryonic CNS of Drosophila. II. Function of the different midline cells. *Development* **126**, 771–779 (1999).

174. Huerta-Cepas, J. *et al.* PhylomeDB v3.0: An expanding repository of genome-wide collections of trees, alignments and phylogeny-based orthology and paralogy predictions. *Nucleic Acids Res.* **39**, 556–560 (2011).

175. Musser, J. M. *et al.* Profiling cellular diversity in sponges informs animal cell type and nervous system evolution. *bioRxiv* 758276 (2019). doi:10.1101/758276

176. Tessmar-raible, K., Steinmetz, P. R. H. & Snyman, H. Fluorescent two-color whole mount in situ hybridization in Platynereis dumerilii (Polychaeta , Annelida), an emerging marine molecular model for evolution and development. **39**, 5–7 (2005).

177. Jékely, G. & Arendt, D. Cellular resolution expression profiling using confocal detection of NBT/BCIP precipitate by reflection microscopy. *Biotechniques* **42**, 751–755 (2007).

178. Longair, M. H., Baker, D. A. & Armstrong, J. D. Simple Neurite Tracer: open source software for reconstruction, visualization and analysis of neuronal processes. *Bioinformatics* **27**, 2453–2454 (2011).

179. Chari, R., Yeo, N. C., Chavez, A. & Church, G. M. sgRNA Scorer 2.0: a species-independent model to predict CRISPR/Cas9 activity. *ACS Synth. Biol.* **6**, 902–904 (2017).

180. Hwang, W. Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229 (2013).

181. Prud'homme, B. *et al.* Arthropod-like Expression Patterns of engrailed and wingless in the Annelid Platynereis dumerilii Suggest a Role in Segment Formation. *Curr. Biol.* **13**, 1876–1881 (2003).