THE UNIVERSITY OF CHICAGO

ADULT NEURAL LINEAGES, TRANSCRIPTIONAL TRAJECTORIES AND GENE PROGRAMS IN THE SEA ANEMONE *NEMATOSTELLA VECTENSIS*

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

COMMITTEE ON DEVELOPMENT, REGENERATION AND STEM CELL BIOLOGY

BY
FLORA PLESSIER

CHICAGO, ILLINOIS
JUNE 2026

Copyright © 2026 by Flora Plessier

All Rights Reserved

"For every action, there's an infinity of outcomes. Countless trillions are possible, many milliards are likely, millions might be considered probable, several occur as possibilities to us as observers — and one comes true."

— *China Miéville, The Scar*

TABLE OF CONTENTS

LIST OF FIGURES vii

ACKNOWLEDGMENTS ix

ABSTRACT xi

1 INTRODUCTION 1
1.1 The evolutionary origins of neurons: an animal-specific cell type with shared eumetazoan roots 1
1.1.1 Multicellularity and the cellular division of labor 1
1.1.2 Evolution of cellular specialization 2
1.1.3 Early emergence of neural cells in animal evolution 4
1.2 Cnidarian neural cells: from common roots to innovations 8
1.2.1 A cnidarian-bilaterian gene toolkit for neural cells 8
1.2.2 Ontogeny and neural fate specification 11
1.2.3 Cnidarian neurotransmission: the old and the new 12
1.2.4 Independent diversification of the cnidarian neuropeptide repertoire . 15
1.2.5 Cnidocytes as a lineage-restricted novel mechanosensory cell 17
1.3 The sea anemone *Nematostella vectensis* as a cnidarian system for neural cell fate specification 20
1.3.1 Cnidarians in development: from induction to genetic laboratory models 20
1.3.2 The genetically tractable starlet sea anemone 22
1.3.3 Neural cell types in *Nematostella* 27
1.3.4 Neurogenesis and cnidogenesis 31
1.4 Chapters overview 34

2 METHODS FOR INVESTIGATING CELLULAR RENEWAL PROCESSES IN THE SEA ANEMONE *IN VIVO* 36
2.1 Introduction 36
2.2 Sections overview 40
2.3 Reporter-based *in vivo* imaging of neural cells 41
2.3.1 Three photoconvertible transgenic reporter lines 41
2.3.2 Single-neuron resolution imaging in live polyps 43
2.3.3 Nerve net stability highlighted by longitudinal live imaging 46
2.4 A photoconvertible reporter method for *in vivo* lineage tracing 50
2.4.1 Whole-organism photoconversion of *Nematostella* polyps 50
2.4.2 Tracking *SoxC>KikGR*$^{+}$ cells during oral regeneration 53
2.5 Spatial localization of newly-generated cells 55
2.5.1 Localization relative to anatomical cell layer 55
2.5.2 Preferential cell distribution relative to mesenteric tracts 60
2.6 Single-Cell Transcriptomic Analysis in the Regenerating Cnidarian *Nematostella vectensis* 63

2.6.1 Abstract . . . 64
2.6.2 Introduction . . . 64
2.6.3 Materials . . . 69
2.6.4 Methods . . . 71
2.6.5 Notes . . . 81
2.7 Conclusion . . . 85

3 CELLULAR AND TRANSCRIPTIONAL TRAJECTORIES OF NEURAL FATE SPECIFICATION IN SEA ANEMONE UNCOVER TWO MODES OF ADULT NEUROGENESIS . . . 86
3.1 Abstract . . . 86
3.2 Introduction . . . 87
3.3 Results . . . 89
3.3.1 Distributed peptidergic neuron generation underlies adult nerve net expansion . . . 89
3.3.2 Progeny of *FoxL2*$^+$ and *SoxC*$^+$ cells encompass neurons, cnidocytes and secretory cells . . . 95
3.3.3 Tracking of *Piwi1*$^+$ progenitor cells differentiation dynamics . . . 103
3.3.4 Direct differentiation of peptidergic neurons . . . 111
3.3.5 Two-step transcriptional maturation of cnidocytes . . . 119
3.3.6 Neural cells arising from different lineages deploy distinct TF-effector gene modules . . . 127
3.4 Discussion . . . 134
3.5 Methods . . . 138
3.5.1 Transgenic line generation . . . 138
3.5.2 Whole animal photoconversion and nerve net imaging and quantification 139
3.5.3 Single cell dissociation and KikGR$^+$ cell sorting . . . 140
3.5.4 Single cell RNA-Seq analyses . . . 144
3.5.5 Reanalysis of RNA-Seq data . . . 149
3.6 Code availability . . . 149
3.7 Data availability . . . 149
3.8 Acknowledgments . . . 149
3.9 Author Contributions . . . 150

4 DISCUSSION . . . 152
4.1 Chapter overview . . . 152
4.2 Distributed adult neurogenesis . . . 152
4.2.1 Is there a stable neural net tissue architecture? . . . 153
4.2.2 Lifelong growth and tissue renewal dynamics . . . 156
4.2.3 Distributed nerve net scaling . . . 158
4.2.4 Cellular lineage and molecular identity of progeny neurons . . . 161
4.2.5 Conserved cell-type but variable cell renewal: the case of cnidocytes . 164
4.3 Anthozoan stem cell system(s) for lifelong neurogenesis . . . 168
4.3.1 Identification of a *Piwi1*$^+$ progenitor pool . . . 168

4.3.2 Lineage restriction within the *Piwi1*$^+$ pool 169
4.3.3 Niche-organized or distributed stem cell system(s) 172
4.3.4 Are there other self-renewing cell pools? 174
4.3.5 A conserved germline program associated to multipotent stem cells . 177
4.4 Two neural gene programs at the base of Cnidaria 180
4.4.1 Ontogeny and neural fate specification in *Nematostella* 180
4.4.2 The divergent late cnidocyte neural program in *Nematostella* 183
4.4.3 Transcriptional cell-type similarities across cnidarians 184
4.4.4 Evolution of two neural developmental programs at the base of Cnidaria 187
4.5 Summary 191

A CNIDARIAN CELL-TYPE COMPARISONS 195
A.1 Cell-type comparisons across distantly related cnidarians 195
A.2 Eight pan-cnidarian cell class transcriptional programs 198
A.3 Cnidocyte, neuronal and secretory transcriptional similarities 204
A.4 Acknowledgements 210

REFERENCES 211

LIST OF FIGURES

1.1 Animal lineages phylogeny 5
1.2 Simplified Cnidaria phylogeny 7
1.3 *Nematostella vectensis* developmental stages 23
1.4 Polyp nervous system organization 26

2.1 Single-neuron resolution imaging in live *Elav>KikGR* polyps 44
2.2 Single-neuron resolution imaging of a live *FoxL2>KikGR* polyp 45
2.3 Persistent nerve net organization highlighted by *in situ* imaging 49
2.4 Efficient whole-organism photoconversion in a *SoxC>KikGR* adult 51
2.5 Pharyngeal region pre and post-photoconversion 52
2.6 Tracking *SoxC>KikGR*$^+$ cells during oral regeneration 54
2.7 *Elav>KikGR*$^+$ cell segmentation and green/orange/red cell annotation 55
2.8 Cell segmentation and landmark registration for spatial analysis 58
2.9 *Elav>KikGR*$^+$ cells spatial distribution by reporter status 60
2.10 *FoxL2>KikGR*$^+$ cells spatial distribution by reporter status 61
2.11 Spatial distribution along the primary body axis by reporter status 62
2.12 Overview of single-cell RNA Sequencing in the regenerating sea anemone *Nematostella vectensis* 68
2.13 FACS gating strategy for live cell-sized singlets with live/dead double-staining . 75
2.14 FACS gating strategy setup for live cell-sized singlets with absolute size comparison 76
2.15 Library quality metrics of a sequenced single-cell RNA-seq experiment. 79
2.16 UMI distribution in cells passing QC and clustering example from a *N. vectensis* regeneration experiment 80

3.1 Discriminating adult-born neurons by photoconversion-based temporal recording of reporter expression 90
3.2 Adult neurogenesis across the primary body axis underlies nerve net expansion . 92
3.3 Adult neurogenesis generates all *Elav*$^+$ peptidergic neuron subtypes 94
3.4 Two new photoconvertible drivers uncover neural non-expressing progeny cells . 96
3.5 Quantification of *SoxC>KikGR*$^+$ and *FoxL2>KikGR*$^+$ cells color distribution at 2, 4 and 7 days post-photoconversion in oral and aboral regions 98
3.6 *FoxL2>KikGR*$^+$ cells encompass both differentiated (neurons, cnidocytes and secretory cells) and morphologically undifferentiated cells at 7 days post-photoconversion 99
3.7 Molecular identity of *KikGR*$^+$ cells from all three drivers 100
3.8 Molecular characterization of KikGR$^+$ cells highlights peptidergic neurons, cnidocytes, secretory cells and undifferentiated cells 102
3.9 Molecular characterization of progenitor cells 103
3.10 A progenitor cell pool that self-renews and differentiates 105
3.11 Undifferentiated *Elav>KikGR*$^+$ green cells in the body wall at 7 days post-photoconversion alongside green, orange and red neurons 107
3.12 Morphologically undifferentiated *FoxL2>KikGR*$^+$ orange cells in the sup-epidermal neurite tracts at 4 days post-photoconversion 108

3.13 Orange undifferentiated *FoxL2>KikGR*$^{+}$ cells at 2 days post-photoconversion . 109
3.14 Progenitor cell reclustering reveals undifferentiated and putative differentiating states . . . 110
3.15 Generation of *Elav*$^{+}$ peptidergic neurons subtypes from *FoxL2*$^{+}$/*SoxC*$^{+}$ progenitors 112
3.16 Peptidergic neuron subtypes are end states arising from *SoxC*$^{+}$ cells . . . 114
3.17 Transient *SoxC>KikGR* expression in differentiating peptidergic neurons, secretory cells, and cnidocytes. . . . 115
3.18 Reanalysis of *Prdm14d* driver RNA-Seq data supports a peptidergic-restricted progenitor lineage . . . 117
3.19 Integrated model of adult neurogenesis . . . 119
3.20 Cnidocytes undergo a biphasic maturation . . . 120
3.21 Developing cnidocytes transiently express *SoxC>KikGR* . . . 121
3.22 Green *SoxC>KikGR*$^{+}$ developing and mature cnidocytes at 7 days post-PC . . 122
3.23 Biphasic transcriptional maturation of cnidocytes with a primary capsule program followed by a secondary neural program delayed after capsule formation . . . 123
3.24 Reanalysis of a *PouIV* KO RNA-Seq dataset supports a biphasic transcriptional maturation of cnidocytes . . . 126
3.25 The two neural lineages deploy distinct neural effector genes . . . 127
3.26 Peptidergic neurons and cnidocytes deploy distinct TF-effector gene modules . . 128
3.27 Peptidergic neurons and cnidocytes use distinct TFs with bHLH and bZIP factors for broad neural identity and a subtype-specific homeodomain code . . . 129
3.28 Cell class-specific and subtype-specific modules encompass known regulators and effector genes . . . 130
3.29 Neural class-specific and cell-type specific WGCNA gene modules . . . 131
3.30 A neural homeodomain code . . . 133
3.31 KikGR+ cell flow cytometer gating strategy . . . 142
3.32 Reporter status attribution ('green', 'orange' or 'red' cell) . . . 143
3.33 Cell identity distribution by reporter and scRNA-Seq library . . . 146
3.34 Total mRNA detected (total UMIs) by cell cluster . . . 151

A.1 Neural programs comparisons across distantly-related cnidarians . . . 196
A.2 Cnidarian scRNA-Seq atlases from distantly related species . . . 197
A.3 Cnidarian atlases by broad cell classes . . . 199
A.4 Eight cnidarian shared cell classes . . . 200
A.5 *Nematostella* and *Clytia hemisphaerica* broad cell class similarities . . . 201
A.6 *Nematostella* and *Hydra* broad cell class similarities . . . 202
A.7 *Nematostella* and *Xenia sp.* broad cell class similarities . . . 203
A.8 *Nematostella* and *Clytia hemisphaerica* cell-type transcriptome similarities . . . 204
A.9 *Nematostella* and *Hydra* cell-type transcriptome similarities . . . 205
A.10 *Nematostella* and *Xenia sp.* cell-type transcriptome similarities . . . 206
A.11 *Hydra* and *Xenia* cell-type transcriptome similarities . . . 208
A.12 *Hydra* and *Clytia* cell-type transcriptome similarities . . . 209

ACKNOWLEDGMENTS

I thank my advisor, Heather Marlow, for the opportunity to pursue this doctoral research. As her first PhD student to defend, it has been an honor to contribute to the inception and development of her lab and grow with it. I am deeply grateful for her availability and openness, and the trust she placed in me to start a multi-faceted high-risk project and bring it to fruition, as it was invaluable to this doctoral journey. I will always remember fondly all the shared moments of joy and enthusiasm, like with the first glowing transgenic animals, and I am thankful for her support during this process, encouragements to engage with the academic community, and the professional environment she fostered.

I thank the members of my thesis committee, Ilaria Rebay, Paschalis Kratsios and Nipam Patel for their constructive feedback on my manuscript, their support, timely advice and willingness to serve on my committee. I thank Ilaria Rebay for support and scientific feedback. I thank Urs Schmidt-Ott for support, advice and manuscript feedback. I thank Paschalis Kratsios for support, manuscript feedback and the inception of the DevNeuro seminar series, and fostering an atmosphere of scientific inquiry, rigor and knowledge transmission, which provided opportunities to present my work to and engage with the Neurodevelopment community. I thank François Spitz. I thank David Leclerc and Mike Olson for flow discussions. I thank Stephanie Laine Nazaire, Giovanni Tolentino Ramos and Sue Levison.

I thank current and former lab members, including Helen, Jelena, Jenks, Leroy, Ja'Marcus, Madi and Harry. I thank Ashish, Sergey, Elizabeth, Helen and Jonathan. I am deeply grateful to Zach for his support and advice throughout. I thank the DevNeuro and DRSB trainees and faculty for discussions, especially Sam for HCR advice, and Honorine, Julio, Deeptha, Annika, Manasa, Sean and Nick for discussions. From Paris, I thank Baptiste Saudemont for MARS-Seq training, Sandrine Schmutz and Sophie Novault for flow training; Marie-Pierre Mailhé, Corinne Chureau and Yann Loe-Mie for discussions. Roel's curiosity and cheerful spirits will forever be in our minds.

This PhD journey would not have been possible without the wonderful human beings that have been rooting for me at all steps of this journey, here and abroad. In Chicago, three amazingly rigorous, intellectually curious and passionate scientists have supported and helped me throughout, and I am incredibly grateful to them: Sébastien, Grace and Caroline. I am deeply thankful to Sébastien for the brilliant advice and insights throughout, always right on point, the far-retching science discussions, and his curiosity, rigor and amazing support, along with delicately worded statements, a true ray of rational enlightenment. I am grateful to Caroline, whose incisive wit, passion, grit and kindness have been a lighthouse in this journey, and for holding us together and keeping my head warm. Throughout the long dark winters and bitter days, our meetups pushed us forward, and I am grateful to them. I thank Grace for the fantastic science discussions and her enthusiasm and encouragements, and all the delicious and fun adventures. From far-away, I am deeply thankful to Thomas, whose support and patience have been infinite, despite the distance, and to my wonderful family, my dad Robert, my mother Laurence and my sister Colette, (and I am grateful for Colette's help with sea anemone silhouettes); Francine for her warm welcome and being a beacon of humanity and rationality; my grandmothers Françoise and Andrée, their grit and perseverance have been a source of inspiration; and my late grandfathers Francis and Carlo. My amazing cousins Ariane, Aurore, and Luc, and my aunt Élisabeth and uncle Benoît for the care packages, fantastic discussions and hosting me, my aunt Sabine and uncle Jeff for all the fun and delicious moments and advice, and the scarf that kept me warm throughout, my aunts Marie and Florence and uncles Yann and Bernard; and my cousins Perrine, Rémi, Eywan, Titouan, Maëlle, and Gabriel for their encouragements throughout. I thank all my amazing friends in France for their support and the wonderful shared moments, and especially Hadrien, Louise, Jean-Louis, Lise, Florence and Emma, and my far-away friends Parisa and Katya. I am grateful for all of you.

ABSTRACT

Animal tissues are built from specialized cells generated and renewed through cellular lineage decisions. Neural cells are a key animal-specific cell type to coordinate organismal function, with extensive diversity across animals. Some animal lineages, including cnidarians (sea anemones, corals, jellyfish), undergo lifelong renewal of all cell types, including neurons. The genetically tractable sea anemone *Nematostella* (Cnidaria, Anthozoa) provides an in vivo system to examine how neural lineages are structured and how distinct neural identities are specified in the context of continuous adult neural renewal. This dissertation investigates the cellular and molecular bases of lifelong neurogenesis in the growing adult. By combining photoconvertible reporter-based pseudo-lineage tracing, single-neuron resolution imaging, and targeted single-cell transcriptomics with recorded lineage status, this work reconstructs the lineage architecture and transcriptional programs underlying adult neurogenesis. This shows that the two main neural cell classes, peptidergic neurons and cnidocytes are generated in a distributed manner alongside the primary body axis, supporting a model of additive distributed neurogenesis rather than spatially restricted neurogenic niches. Newly generated peptidergic neurons are specified proportionally to pre-existing neural subtypes and can be stably integrated into a persistent nerve net architecture. Lineage analyses support that the main neuronal cell class, peptidergic neurons, arises directly from a broadly multipotent progenitor pool. In contrast, cnidarian-specific neural cells, the cnidocytes, develop from a separate cellular lineage, with a lack of a neural-exclusive bipotent progenitor state. At the transcriptional level, peptidergic neurons are specified through a common peptidergic gene regulatory program associated to their shared cellular lineage, combined with subtype-specific modules. Cnidocytes, by contrast, undergo biphasic maturation, in which a late-onset neural program, distinct from that of peptidergic neurons, is secondarily activated. Comparisons across cnidarians support the conservation of this biphasic mode of neural specification and point towards the existence of two largely distinct neural gene programs at

the base of Cnidaria. Together, these provide groundwork to understand adult neurogenesis in an anthozoan cnidarian and suggest an early divergence of neural developmental programs associated to lineage architecture, which may have constrained the diversification of neural cell types in early animals.

CHAPTER 1

INTRODUCTION

1.1 The evolutionary origins of neurons: an animal-specific cell type with shared eumetazoan roots

Cells are the fundamental unit of organisms, with the DNA-based genetic blueprint and the molecular machinery necessary for their function. With the advent of multicellularity, specialized cells are integrated into coordinated organisms, each cell expressing a subset of the genetic content required for overall fitness. Understanding how, in each organism, differentiated cells expressing distinct gene programs are specified from a shared genomic blueprint through progressively restricted cellular lineages is a fundamental question to understand the function and evolution of specialized cells. This dissertation work investigates how neural cells, an animal-specific novel cell type, are generated, using lifelong neural fate acquisition in the anthozoan cnidarian sea anemone *Nematostella vectensis* as a system to understand the development and evolution of specialized cell types.

1.1.1 Multicellularity and the cellular division of labor

For most of Earth's history, life existed in unicellular form, with the presence of genes and a DNA-based genetic code dating back to over 4 billion years ago (bya)[1]. Eukaryotic cells, with their nucleus and membrane-bound organelles, arose over 1.6 bya, alongside rising oxygen levels. From unicellular aerobic eukaryotic organisms, obligate multicellularity evolved independently at least five times, giving rise to fungi, three lineages of multicellular algae, and to animals (Metazoa), over 600 mya[2,3]. The transition from unicellular eukaryotes to obligate multicellular organisms like animals entailed the cohesive assembly and organization of cells expressing distinct specialized genetic programs from a common DNA blueprint, coordinated across the organism. This division of labor enabled the emergence of larger,

self-organizing life forms with properties such as large-scale intercellular communication, reproductive specialization, cellular regeneration and coordinated functions. Among multicellular eukaryotic lineages, animals stand out for their active movement, large body sizes, and predatory lifestyles: as heterotrophic aerobic organisms, they rely on ingestion rather than photosynthesis for their organic carbon supply. These life history traits are enabled by numerous cellular innovations, and notably their contractile muscle cells and their neurons.

The emergence of animals from unicellular eukaryotes had several consequences: (**1**) all cells within one organism share the same DNA blueprint, even as they express distinct gene programs, (**2**) a large share of this genetic toolkit has a unicellular origin, with modest increases in gene content relative to gene programs combinations, and (**3**) in animals, a single cell (the zygote) gives rise to all differentiated cells in a precisely coordinated developmental process, and cell type specification is directly linked to cellular lineage. Animal tissues are built and maintained by lineage decisions, with the expression of distinct combinatorial gene programs from a shared genetic template, and end up 'locked in' with their terminally differentiated state. Understanding these processes is foundational to the field of developmental biology.

1.1.2 Evolution of cellular specialization

In a multicellular organism, cellular diversity does not arise from difference in genome content between cells, but from the precise regulation of gene expression from a shared genome, with distinct cell populations using partially overlapping combinations of genes to support specialized functions, coordinated across the organism. Individual cells are thus distinct functional units, but they they are linked through developmental lineage relationships, shared molecular profiles and are integrated into tissues/organ structures. These similarities have provided the basis for classifying cells into cell types, across multiple levels of granularity, extending classical cytological, histological, and molecular descriptions. Importantly, cell types

are not strict well-defined discrete units, but a more conceptual hierarchical representation of a combination of shared properties, shaped by developmental history and gene program expression. This can be useful for comparing cellular identities within an organism and for retracing relationships across evolutionary distances. The advent of single cell transcriptomic and epigenomics have enabled global characterization of gene expression programs and their regulatory elements across tissues, developmental stages, and whole organisms for an increasing number of animal lineages. Building on decades of experimental work in each system, these approaches have helped expand the description and classification of specialized cells from marker-based to global reconstructions of their gene regulatory programs, thus enabling broader comparisons to study cell fate acquisition mechanisms and the evolutionary origins of cellular diversity (reviewed across[4–7] and many others). From this perspective, the evolution of novel cell types is intertwined with the evolution of new genes and new gene regulatory programs.

As such, reconstructing the evolutionary history of cell types is inherently a multiscale problem. Cells and cell types are not independent evolutionary units: within an organism, all cells share the same genome and originate from a single cell. And the expression of individual genes is an evolutionarily labile trait, that can be gained or lost multiple times. Thus, cell types are not described by single marker genes, but more robustly by the deployment of gene regulatory programs. The conservation of functionally associated gene co-expression modules, especially if associated with gene family expansions and diversification events in the ancestor, can provide support for proposed shared ancestry, and potential cell-type homology. This appears to be the case, as a whole, for a global neural cell program across bilaterians and cnidarians[8,9], likely with additional extensive lineage-specific diversification of neuronal subtypes and effector repertoires.

Across evolutionary timescales, the specification of novel cell types relies on a combination of several processes, including (**1**) new protein-coding genes, rarely arising *de novo*,

and more often from the reshuffling of existing protein domains, gene duplication and divergence; (**2**) post-transcriptional and translational mechanisms, that can give rise to distinct proteins such as alternative splicing, alternative promoter usage, RNA editing and protein modification; thus expanding the available molecular toolkit from existing genes; and (**3**) changes in regulatory architecture, including the emergence of novel cis-regulatory elements (enhancers), shifts in transcription factor binding and availability, and epigenetic modifications, which can change expression programs, and change in 3D genome configuration that may promote selective enhancer-promoter interactions.

Regarding the origin of the molecular toolkit used to build novel animal cell types, genome sequencing of a diversity of animal species and their closest unicellular relatives (including choanoflagellates, filastereans, and ichthyosporeans) over the past two decades has provided insights into the evolution of the animal gene repertoire[10–13]. Overall, comparative genomics indicate that a large share of the "novel" animal genes have arisen from pre-existing protein domains already present in unicellular eukaryotes through extensive re-use, duplication and recombination, extending the molecular toolkit and giving rise to novel animal gene families. This protein-coding repertoire has been further expanded through extensive animal-specific gene duplication and sequence divergences, many of which are shared between bilaterians and cnidarians, including among transcription factors families[13–15].

1.1.3 Early emergence of neural cells in animal evolution

All animals (except for sponges and placozoans) have a nervous system. Neuron, the key cellular unit of nervous systems, is an animal innovation specialized for fast, ion-current-based, long-distance communication and signal integration. Like all specialized cell types, neuron function relies on molecular components encoded by gene families (like ion channels, pre- and post-synaptic scaffold proteins, gap junction proteins, etc..) (see below). Most of these "neural" gene families are present across unicellular eukaryotes, and thus predate the

origin of animals (and of neurons), with some families arising at the base of Metazoa[16–20]. However, from their pre-metazoan origins, there has been expansions and diversification of these gene families early in metazoan evolution, concomitant with the emergence and diversification of neurons across animal lineages (see for instance[20–23]).

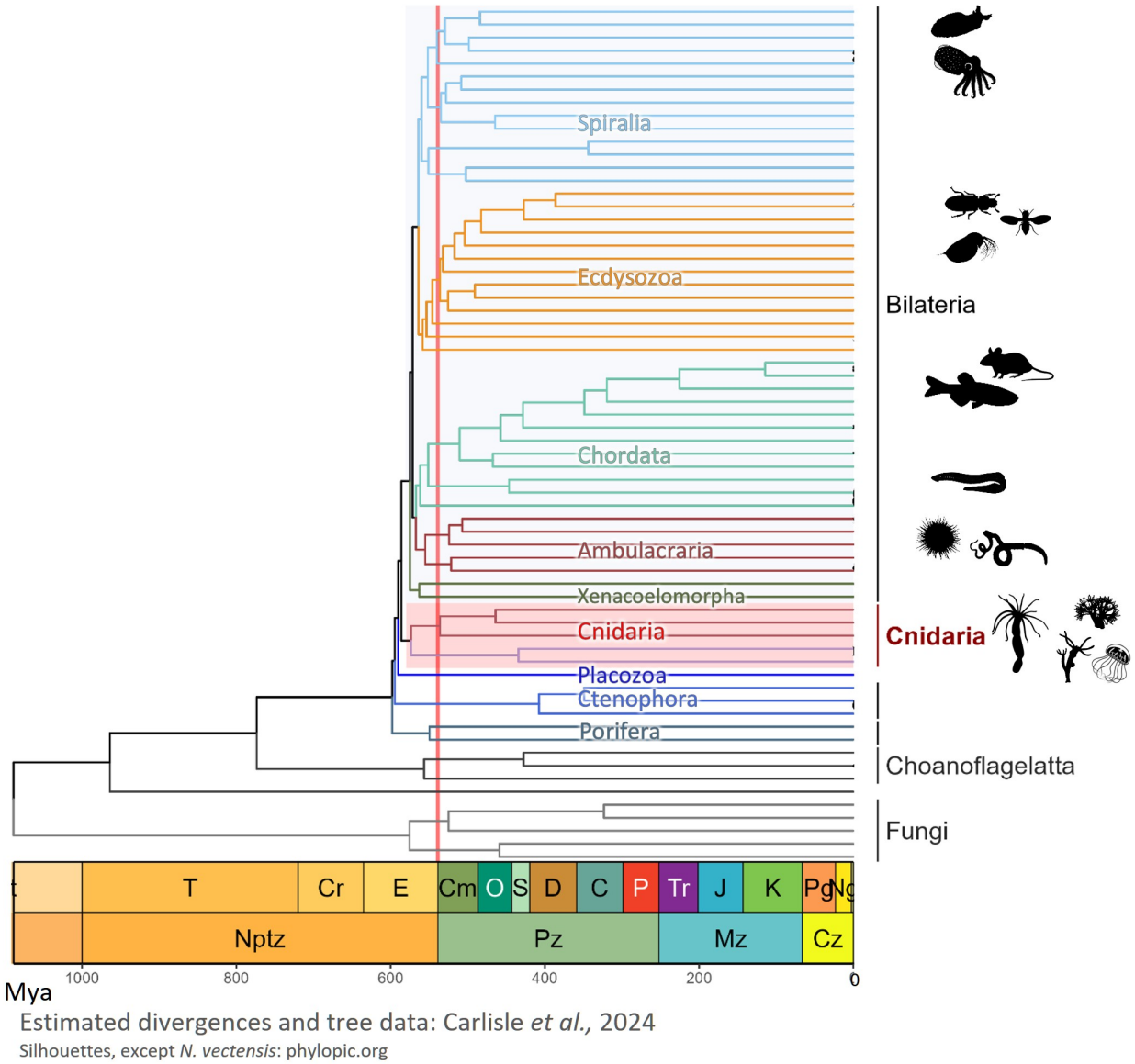

Figure 1.1: Animal lineages phylogeny

Metazoan phylogeny with estimated divergence ages across major branches (colored), with non-metazoan outgroups in black, and Cnidarians in red. Estimated divergence ages and data from[3].

Phylogenetic reconstructions (informed by paleontological data) indicate that animal lin-

eages diversified very rapidly, at least 530 Mya (likely spanning the late Ediacaran-early Cambrian periods)[3] (see Fig. 1.1). A relatively short window (at geological timescale) from the apparition of the first animals saw the emergence of the ancestors of extant poriferans (sponges), ctenophores (carnivorous comb jellies), placozoans (creeping millimeter-sized animals), cnidarians (sea anemones, corals, jellyfish), and bilaterians (all other animals, including humans), as well as the rapid diversification of all major bilaterian groups (protostomes vs. deuterostomes, and ecdysozoans vs. spiralians within the protostomes, and ambulacrarians vs. chordates within deuterostomes).

These early rapid diversification events thus saw the emergence of Eumetazoa, which encompasses both Bilateria (most animals, including humans, flies, slugs and worms) and Cnidaria (sea anemones, hydra, jellyfish, corals) (see Fig. 1.2), and is characterized by the presence of organized tissues, organismal and tissue polarity, an internal digestive cavity and a basement-membrane containing extracellular matrix, and specialized cell types, including neurons and muscle cells. Their embryonic development involves regulated patterns of cell division and migration, with the specification of spatial territories (embryonic layers) during gastrulation, that differentiate into distinct cell types.

The sequencing of non-bilaterian genomes, and of unicellular eukaryotes related to animals (like choanoflagellates, and filastereans) and the targeted investigations of orthologs of known bilaterian neural effector genes families highlighted that a large share of the gene families encoding proteins essential for neuronal function predate the emergence of neurons, or emerged at the base of Metazoa[16,25,26]. Despite the ancient and early divergences of animal lineages, phylogenetic analyses of neuronal gene families and their expression in neurons of cnidarians, ctenophores, and across bilaterians, support a broadly conserved neuronal molecular program shared between cnidarians and bilaterians, while ctenophores (which also have neural cells) exhibit both a divergent gene repertoire, and gene expression patterns in their neural cells, raising the strong possibility of an independent evolution of neurons in

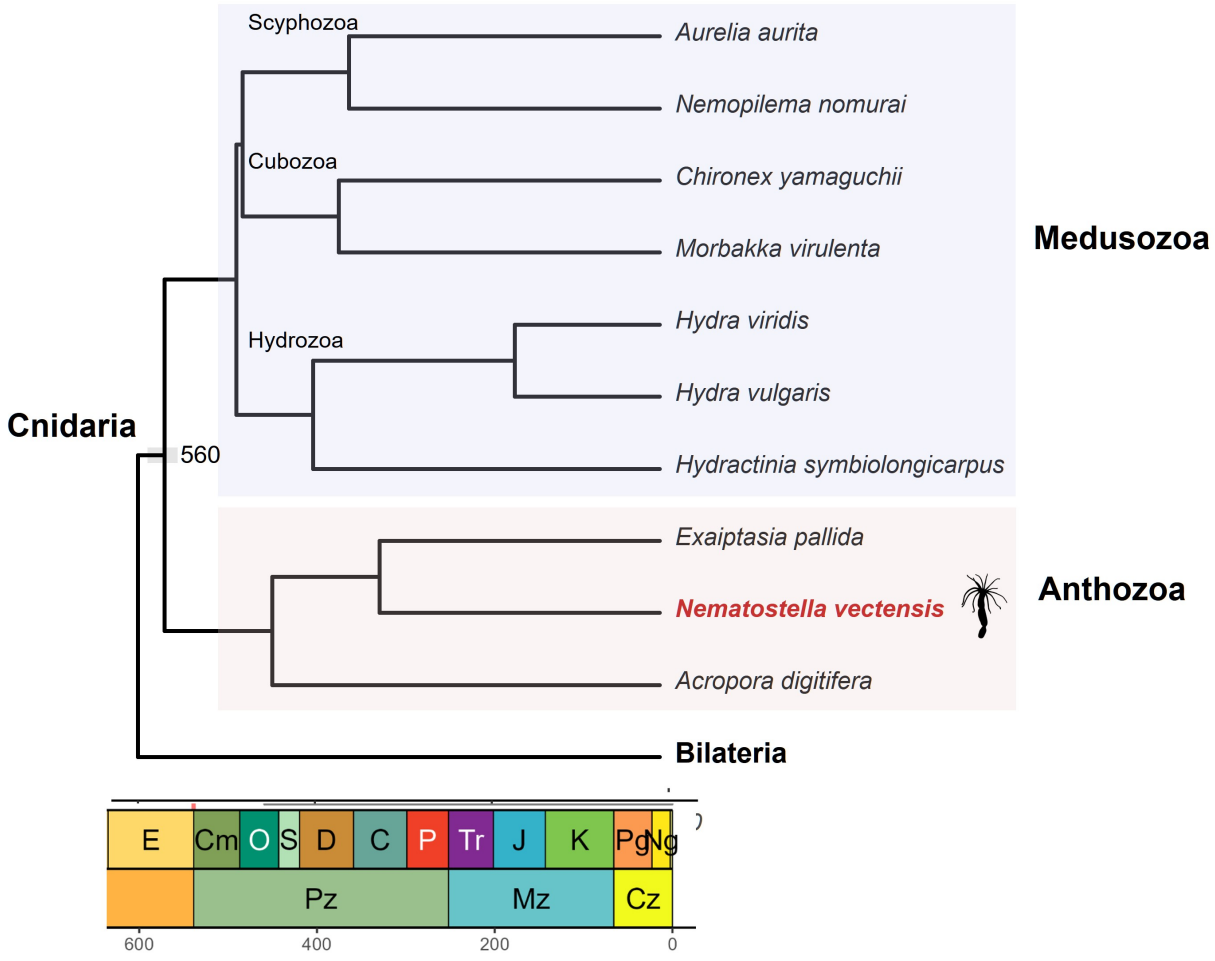

Figure 1.2: Simplified Cnidaria phylogeny

Simplified cnidarian phylogeny with Anthozoa (sea anemones, corals, sea pens) diverging from Medusozoans (which include species with jellyfish stages) (Hydrozoa, Scyphozoa, Staurozoa (not included) and Cubozoa) around 560 Mya. The sea anemone *Nematostella vectensis* is highlighted in red. Estimated divergence ages and tree data from[24].

ctenophores[8,22,27–29].

Across animals, nervous systems are not built from a single neuron type, but instead from many distinct neuronal subtypes with specialized morphologies, anatomical localizations, synaptic partners, physiology and molecular identities. A major challenge in understanding nervous system evolution is explaining how neuronal diversity is generated and maintained. In bilaterians, much of what is known about neurogenesis derives from a small number of model systems with centralized nervous systems. Thus, cnidarians, as the sister group to

Bilateria while still sharing common eumetazoan neuronal molecular program, provide an opportunity to investigate neural cell type diversification principles.

This dissertation work investigates the cellular lineages and gene programs of neural fate acquisition in the sea anemone *Nematostella vectensis* (Cnidaria, Anthozoa) during adult neurogenesis, with the goal of understanding how distinct cnidarian neural cell types are generated throughout life and how developmental programs give rise to both shared neuronal identities and lineage-specific innovations.

1.2 Cnidarian neural cells: from common roots to innovations

1.2.1 A cnidarian-bilaterian gene toolkit for neural cells

As this work focuses on the developmental lineages and gene programs of neural cells in cnidarians, using the model sea anemone *Nematostella vectensis*, we first broadly define "neural" cells properties, as characterized initially in bilaterian systems, and some of the gene families that functionally enable these properties, and their presence in cnidarians, supporting some form of evolutionary relationship between bilaterian and cnidarian neurons. The implicit assumption that this cell type (or some parts of it) can be traced back to the last common ancestor of both bilaterians and cnidarians underlies the label neural cell/neuron used to describe cells present in extant cnidarians. This is important for the interpretation of cellular renewal mechanisms used for neural cells across lineages, and their diversification.

From extensive work in bilaterians, neurons are electrically excitable cells that generate regulated ion currents (action potentials), and can transmit information to other cells. Neurons can release chemical messengers acting locally for targeted cell-cell communication over long distances at chemical synapses. On a molecular level, these properties rely on the establishment and maintenance of electrochemical gradients (a feature not restricted to animal cells or neurons), the generation of action potentials, including a rapid return to baseline membrane potential, and their propagation at a distance within specialized cellular

structures (neurites) through ion currents. Lastly, neurons can receive chemical information through their detection of neurotransmitters at post-synaptic densities, and convert electrical signal to a chemical signal through the synthesis and localized release of neurotransmitters via controlled exocytosis in the pre-synaptic density, although not all neurons do that. In bilaterians, neurons are typically classified into sensory neurons, motoneurons and interneurons, depending on their network partners and function. Neurons are specified in progressively restricted cellular lineages originating from a defined subset of the (neural) ectoderm, with this embryonic lineage restriction persisting throughout life. While many of these properties are not unique to animal cells (like fast organismal motility control by membrane depolarization and ion currents in the unicellular eukaryote *Paramecium*[30]), or to neurons (ion channels are widely expressed and regulated exocytosis is used in secretory cells for instance, and not in some neurons), their combination, and reliance on specific gene families, provides a solid common ground for defining this "cell type".

The functional properties of neurons arise from their expression of effector genes, including during their development. These neural effectors have been best identified and characterized across bilaterians, and include, but are not limited to, voltage-sensitive ion channel families[17,19,20,31,32] (including Na_v, Ca_v, K_v, DEG/ENaC/ASIC, TRP, etc.), neurotransmitter synthesis enzymes, ligand-gated ion channels (iGluR[33,34], etc.), neurotransmitter GPCRs (like metabotropic glutamate receptors, bioamine receptors, neuropeptide receptors), neurite guidance systems, neuropeptide processing machinery (prohormone convertase, etc.), synaptic machinery proteins (including pre and post-synaptic scaffold proteins), gap junction proteins (connexins in chordates, innexins otherwise). While the protein domains and gene families used to build a neuron are, in large part, older than neurons, these gene families have expanded and molecularly diversified across animal with neural cells, with early protein divergence events shared between the last common ancestor of cnidarians and bilaterians for many of these families, along with further expansions at the base of Bilateria, and within

each animal lineage.

As one historical example of such conserved neural gene family, ELAV proteins have been used as neuronal markers across bilaterian nervous systems[35]. The Embryonic Lethal Abnormal Vision-Like (ElavL) family of RNA-binding proteins was identified in *Drosophila* in the late 1980s as part of a genetic screen for neuron-specific factors required for nervous system development. Independently, related proteins were discovered in humans as targets of autoimmune responses in anti-Hu–associated paraneoplastic encephalomyelitis and sensory neuropathy. The human Hu proteins represent the vertebrate orthologs of the *Drosophila ElavL* family. As neuronal-specific markers conserved across bilaterians, ELAVL proteins constituted early candidates for neural system investigations in non-bilaterian lineages[36,37], and provided the rationale for the generation of an *Elav>mOrange* genetic driver in the sea anemone *Nematostella vectensis*[38], which is indeed broadly neuronal, as seen in earlier expression studies[36].

Cnidarians neural cells display canonical functional properties, including action potentials, chemical synapses, electrical synapses, and long-distance communication. These properties rely on macromolecular intracellular structures (like the pre and post-synaptic densities, neurites, and gap junctions), and express orthologous genes. The conservation of effector gene families and their expression and function in neurons support a common bilaterian-cnidarian neural gene program, with additional lineage-specific diversification.

In contrast, the other non-bilaterian animal lineage with neural cells, ctenophora (whose phylogenetic position is debated), shows a syncytial nerve net organization[28], which, along with other molecular features (also reviewed in[22,29] and many others), argue for a more independent origins of neurons in ctenophores; and thus support a common eumetazoan neural gene program shared between cnidarians and bilaterians.

1.2.2 Ontogeny and neural fate specification

Neural cells are specified through developmental trajectories in which pluripotent or multipotent progenitors progressively activate their gene programs. These programs rely on cis-regulatory elements (enhancers and promoters) that are bound by cell-type-specific transcription factors to coordinate gene expression.

Major transcription factor families, including Homeodomain, bZIP (basic leucine zipper), Forkhead, SOX (**S**RY-related High Mobility Group-**box** genes) and basic helix–loop–helix (bHLH) genes, diversified early in animal evolution[39–43], with specific transcription factors part of a deeply conserved toolkit for neural fate specification. In bilaterians, neurons typically arise from neuroectodermal progenitors[44] through the activity of SoxB/SoxC[45] genes and proneural bHLH factors[40] of the achaete–scute and atonal/neurogenin classes. While some of these gene families further diversified in bilaterians, like the atonal and neurogenin families[40], and the SoxB1 and SoxB2 subgroups[43,46], homologous genes from the Sox B and C subfamilies, and the bHLH achaete-scute, and atonal/neurogenin regulate neurogenesis across distantly related cnidarians[47–49], both in the anthozoan *Nematostella vectensis*[50–53], the solitary hydrozoan *Hydra* (reviewed in[48]), and the closely related colonial hydrozoans *Hydractinia echinata*[54] and *Hydractinia symbiolongicarpus*[55], thus spanning the earliest split within Cnidaria. These support the presence of an ancestral neurogenic specification program in the last cnidarian-bilaterian ancestor (over 530 Mya).

Neural fate is classically linked to embryonic ectodermal patterning and neurectoderm specification in bilaterians (broadly including neural-crest derived enteric neurons in vertebrates). In the zebrafish *Danio rerio*, where the adult forebrain shows extensive lifelong neurogenesis, adult neural stem cells also derive from neurectodermal cells specified during early embryonic stages (reviewed in[56]). However, in some protostomes such as the nematode worm *Caenorhabditis elegans*, neurons arise through invariant lineage-based programs which also give rise to non-neurons rather than from an embryonic spatially determined

neurectodermal territory. And other animal groups, including cnidarians, planarians, and some acoels, display indeterminate growth, with the lifelong generation of neuronal subtypes from broadly multipotent (like in *Hydra*[57]) to collectively pluripotent stem cells with neural-committed subsets (in planarians[58,59]), thus showing that neural differentiation can be, at least partially, uncoupled from germ layer restriction.

These lifelong neurogenic systems thus provide the opportunity to specifically investigate neural fate specification mechanisms, decoupled from the embryonic patterning and morphogenesis events at play during early life stages. As the sister group to Bilateria, with a broadly shared neural gene families toolkit and neurogenic developmental regulators, cnidarians provides informative systems to investigate neural fate specification and the diversification of neural cell types.

1.2.3 Cnidarian neurotransmission: the old and the new

The neuronal genetic toolkit is widely shared between cnidarians and bilaterians. Both possess chemical synapses, which are likely built from orthologous proteins ([17,25,60] and many others), propagate long-distance action potentials through neurites and use conserved regulators for neuronal fate, supporting deep homology. However, cnidarian neurons show independent expansion of their neurotransmitter repertoire, are organized in nerve net networks, with local condensation in several groups, but no central nervous system, and have a distinctive neural cell innovation: the cnidocyte, identified in every cnidarian species thus far.

Cnidarian neurotransmission is proposed to rely heavily on a diverse neuropeptide repertoire, with some evidence for glutamate, glycine, and GABA as well, and large areas of uncertainty regarding bilaterian small molecule neurotransmitters like acetylcholine and monoamines (serotonin, dopamine, adrenaline, noradrenaline, etc..), despite 40+ years of literature across cnidarian systems (reviewed in[61,62]). While these molecules can be de-

tected, for each putative neurotransmitter, the experimental evidence demonstrating that it is both (*i*) synthesized in the animal, (*ii*) by neural cells, (*iii*) released via controlled exocytosis, and (*iv*) capable of activating receptors on target cells to generate electrical signaling events, has been partial and sometimes ambiguous for most canonical neurotransmitter molecules[23,61–63].

As one such example, acetylcholine (Ach) is a classical excitatory neurotransmitter in bilaterians (also broadly present in non-neuronal cells including immune cells[64]), and it can be detected biochemically across cnidarians, plants, fungi and unicellular organisms[65]. There is conflicting evidence for its usage as a cnidarian neurotransmitter (see also[61,62]). On the one hand, there is no cnidarian ortholog of the bilaterian synthesis enzyme (choline acetyltransferase - ChAT), only of closely related carnitine palmitoyltransferases[23,66]. Its specific vesicular transporter (vAchT) appears to be a bilaterian innovation[67]. There is a cnidarian-specific gene family orthologous to the nicotinic acetylcholine receptors (nAchR)[68,69], with specific expression patterns, including in muscle cells in the sea anemone *Nematostella vectensis*[68,70]. Whether these receptors bind acetylcholine is currently unknown, especially as they form a cnidarian-specific clade, and there are instances of lineage-specific ligand-specificity divergences within the cys-loop ligand-gated ion channels superfamily they belong to. Notably, chemoreceptors in the octopus *Octopus bimaculoides* are also related to nAChRs, but their ligand is not acetylcholine but organic compounds[71,72]. On the other hand, behavioral and electrophysiological studies showed that exposure to high concentrations (10mM) of acetylcholine can induce tentacular contraction in the sea anemone, abolished by the nAchR antagonist mecamylamine[68], with experiments in the hydrozoan hydra showing a modulation of electrical activity associated to contraction (reviewed in[61,62]). Furthermore, Ach administration can induce cnidocyte hyperpolarization and K^+ currents, blocked by the nAchR antagonists tubocurarine or mecamylamine in a sea anemone, with cholinergic signaling thus proposed to be part of the chemo-sensory neuromodulation of cnidocyte discharge[73].

Some acetylcholinesterase enzymatic activity can be detected in the tentacle bulb but not in the nervous system in a hydrozoan jellyfish[74]. And while there is biochemical acetylcholinesterase activity detected throughout the tentacles in a sea anemone[73], expression of two cholinesterase genes showed a broadly non-neural distribution[75]. Overall, it remains unclear if acetylcholine is a neurotransmitter in cnidarians.

Similarly conflicting evidence exist for most monoamines transmitters including serotonin, dopamine, noradrenaline, etc.. from the biochemical detection of the molecules[76,77], including in neural cells, and functional assays, to the genomic inventory of biosynthesis enzymes (highlighting a lack of tyrosine and tryptophan hydroxylase), transporters and cognate receptors. Melatonin (and its key synthesis enzyme, N-Acetylserotonin O-methyltransferase) appears to be present across an anthozoan and a hydrozoan and associated with day/night cycle and a proposed "sleep-like" state[78–83], but whether it may act as a neurotransmitter is undetermined. Overall, usage of monoamine neurotransmitters (i.e. dopamine, adrenaline, noradrenaline, serotonin, histamine, tyramine and octopamine) has been proposed to be a largely bilaterian innovation[67].

Even for small molecule often listed as cnidarian neurotransmitters[23], like glycine and glutamate (with orthologs of bilaterian vesicular glutamate transporters and of glutamate ionotropic receptors (iGluRs[33,34], along with cnidarian-specific diversification)), the evidence can be sparse and ambiguous as to whether these molecules fit the "neurotransmitter" criteria listed above. Similarly for proton, (H^+), purinergic and nitric oxide (NO) signaling[23], and GABA[36,84], with its key synthesis enzyme, glutamic acid decarboxylase (GAD) mostly expressed in non-neural tissues in sea anemone[75]. In contrast to these, there are strong behavioral and molecular evidence for peptidergic-associated neurotransmission in cnidarians, using a diverse neuropeptide repertoire. Therefore, in this work, we use peptidergic signaling components as markers of neuronal populations.

1.2.4 Independent diversification of the cnidarian neuropeptide repertoire

Neuropeptides are key signaling peptides in animal nervous systems that mediate intercellular long-range communication[85]. These short bioactive peptides (usually 6-20 residues) can act as neurotransmitters via synaptic release to other neurons or target cells, via volumetric signaling (extra-synaptic short-range diffusion), and through circulatory systems as hormones[86]. As short bioactive peptides, neuropeptides are generated from longer propeptides. Each precursor can give rise to distinct peptides depending on their cleavage by endopeptidases like prohormone convertases (PC) and exopeptidases like carboxy and aminopeptidases, and post-translational modifications, important for stability and function, like C-terminal amidation from a peptidylglycine α-amidating monooxygenase (PAM, sometimes split into two enzymes), and N-terminal glutamine cyclization to form pyroglutamate, by glutaminyl cyclases (QC). The neuropeptide processing machinery predates neurons and metazoans, with for instance functions of PAM orthologs in cilia assembly across green algae and animals[87,88]. Additionally, short amidated peptides are also produced in neural-less placozoans[89,90] and in neurons in ctenophores[91,92], and two neuropeptide precursors could be identified across bilaterians, cnidarians, ctenophores and sponges, as well as in a unicellular relative of animals, a choanoflagellate[93], in line with the pre-metazoan origins of many neural building blocks, and thus supporting a stepwise assembly of neural programs[29,90].

Despite these pre-metazoan origins of neuropeptide neurotransmission, the almost thirty distinct neuropeptide signaling systems identified in bilaterians are, in large part, representative of an ancestral urbiletarian neuropeptide-receptor complement, not shared with cnidarians or other lineages[94–96]. And the large neuropeptide repertoire of cnidarians does not bear clear sequence similarities with that urbilaterian neuropeptide complement[95,97], although there are archetypal cnidarian neuropeptide families with RFa and GLWa C-termini, reminiscent of some of the bilaterian neuropeptide families (see below for a discussion of Wa/RFa convergence).

However, the detectability of similarities for active neuropeptides is challenging, and limited by their very short length (usually $\tilde{6}$ to 20 amino acid residues), and the divergence of their longer proprotein. Thus, as neuropeptide ligands and receptors co-evolve, sequence-based searches using their receptors have been instrumental in clarifying their evolutionary history. Investigations of the main class of neuropeptide receptors: class A rhodopsin-like G-protein coupled receptors (GPCRs) has highlighted both the clear lack of cnidarian orthologs of the different families of bilaterian-conserved neuropeptide GPCRs, as well as the extensive cnidarian-specific diversification[94–98]. Overall, these elements strongly support the independent diversification of the cnidarian peptidergic repertoire from that of Bilateria.

Due to this independent diversification, the identification of cnidarian neuropeptides has relied on a combination of high performance liquid chromatography, mass spectrometry, and sequence-based searches to characterize the complement of mature peptides and their precursor genes[98–101] (reviewed across[102–105]). *In vivo* assays in the hydrozoans polyp *Hydra* and jellyfish *Clytia hemisphaerica* have highlighted neuropeptides functions in muscle contraction, larval settlement, oocyte maturation and spawning[104,106–110], and in settlement timing in the anthozoan sea anemone *Nematostella vectensis*[111]. Currently, only a few receptor-neuropeptide pairs have been identified[109]. The full identification of their cognate receptor(s) remains technically challenging, especially given the number of putative candidate class A GPCR receptors (hundreds in each species). A large scale *in vitro* assay in *Nematostella vectensis* deorphanized 31 GPCR peptide receptors out of from 161 GPCRs tested (using a panel of 64 peptides)[98]. And a rare family of peptide-gated ion channels from the DEG/ENaC gene family has been identified in *Hydra*[112–116]. Although the DEG/ENaC gene family also includes molluscs FMRF-gated ion channels, phylogenetic, molecular and electrophysiological analyses support an independent evolution for the hydra ones (see also below).

Current data across extant cnidarians supports lineage-specific diversification of neu-

ropeptides across groups, with shared pan-cnidarian features limited to conserved C-terminal motifs like RFamide, GLWamide, and PRXamide (like PRGamide)[98,100,102–105]. These C-termini motifs are also present in some bilaterian neuropeptide families, and it is currently unclear if this similarity represents an ancestral conservation in eumetazoans[117] or convergent evolution, with the latter supported by the overall lack of conservation in their currently identified cognate receptors, and their diverging peptide sequence beyond the C-terminus. Neurons in cnidarians utilize extensively diversified neuropeptide repertoire, which as neural effector gene provide markers of neuronal subtypes specialization.

1.2.5 Cnidocytes as a lineage-restricted novel mechanosensory cell

Cnidocytes[1] are a novel cnidarian-restricted neural cell type, and a phylum-defining synapomorphy present in all extant cnidarians[118] (including highly simplified parasitic species). These mechanosensory cells harbor a single-use intracellular harpoon-releasing structure: the cnidocyst, which is used for defense, venom delivery, prey capture and/or substrate attachment. The cnidocyst capsule is assembled from taxonomically restricted genes, consistent with a single origin of cnidocytes at the base of Cnidaria, over 530 Mya. As mechanosensory (albeit specialized) cells, cnidocytes are recognized as part of the neural cell type complement in cnidarians[119,120] (and see Discussion in chapter 4).

Cnidocyte discharge is a one-time event, triggered by mechanical (like prey movement) and chemical (like prey-derived chemicals) stimuli, that can be autonomously triggered and that is regulated by neural input[121–123]. The intracellular capsule (cnidocyst) contains a tightly packed coiled thread-like long barbed tubule usually with a long filament. Upon discharge, considered to be a possible Ca^{2+}-dependent exocytic event, this thread gets ejected

1. Cnidocytes with a piercing thread are also called nematocytes. In some cnidarians like hydrozoans, all cnidocytes are nematocytes so the two terms are used interchangeably. The cnidocyte capsule is called a cnidocyst, or a nematocyst in nematocytes. Other cnidarians, like anthozoans also have non-piercing cnidocytes, like the sticky ensnaring spirocytes, which are not considered nematocytes; so in anthozoans, there are both nematocytes and non-nematocyte cnidocytes.

from the pressurized capsule at high speed and can pierce, ensnare and/or envenomate preys. The cnidocyte capsule contains a high concentration of poly-γ-glutamate, which binds large amounts of cations (2M), resulting in a high intracapsular osmotic pressure[124,125], powering its explosive capsule eversion and harpoon function. Mechanical and chemical stimulation of the cnidocyte can induce a Ca^{2+}-dependent discharge, causing the rapid eversion of the intracapsular thread structure with its barbed end, and in *hydra*, an acceleration of up to 5×10^6 g and a pressure of 7GPa at impact site on the prey reported[121,126–129]. The release of this intracellular harpoon is among the fastest cellular biological processes described, in about 3ms in the Hydrozoan *Hydra*[121], and under 15ms in the sea anemone *Nematostella vectensis* (Anthozoa)[130].

Cnidocytes are a rare example of a cellular innovation built, at least in part, from novel genes families, like the minicollagens, named for their very short triple-helix collagen domain and flanked by polyproline tracts and cystein-rich domains (CRD), and NOWA (**N**ovel **W**all **O**uter **A**ntigen) families, among others. The extremely pressure-resistant capsule is made within a post-Golgi vesicle where NOWA glycoproteins and minicollagens form disulfide bonds, assembling into a macromolecular capsule structure[126,131,132]. The tubule structure is built from one end of the capsule, with structural proteins including specific lectin (nematogalectin), chondroitin, and minicollagens, and it later invaginates into the capsule which then closes upon maturation[130,133]. The minicollagen gene family has undergone tandem duplications and expansions in cnidarian groups[24], which are differentially associated with cnidocyte subtypes[132].

Cnidocytes are present across all cnidarians, with cnidocyte subtypes historically classified based on capsule morphology, following systematic descriptions from comparative studies initiated in the 1930s[134,135]. Despite differences in capsule morphology, cnidocytes are built from a broadly conserved set of orthologous proteins, including phylum-restricted novel genes like minicollagens. There are also lineage-specific innovations, like the sticky non-

piercing spirocytes specific to anthozoans[136], which use an anthozoan-specific minicollagen gene (*NCol5*)[132]. And the associated mechanosensory cilium (the cnidocil) is much shorter in anthozoans, but elongated and contains *Nematocilin*, a central structural component in hydrozoans[137], absent from anthozoans. At the molecular level, single cell RNA-Seq atlases have currently identified just one to three molecularly distinct cnidocytes subtypes, across the sea anemone *Nematostella*[8,53] (Anthozoa, Hexacorallia), the hydrozoans *Hydra*[138,139],*Clytia hemisphaerica* (jellyfish stage), and *Hydractinia*[140,141], the scleratinian coral*Stylophora pistillata*[142] (Anthozoa, Hexacorallia), and the octocoral *Xenia sp.*[143] (Anthozoa, Octocorallia), supporting limited molecular diversification of this cell type.

Cnidocyte discharge is tightly regulated process, with the discharge involving calcium and thought to be a modified exocytic event[123]. Cnidocytes can extend thin neurite-like protrusions, whose possible connectivity and functions are unknown, and their neural circuit integration is poorly characterized, with some neural-cnidocyte synapses also described in some species from earlier electron microscopy studies[144]. As cnidocyte capsule discharge is a one-time event, these cells are renewed throughout life across cnidarians. Their lifelong renewal can involve a spatially restricted cellular conveyor-belt systems as in the hydrozoan jellyfish tentacle bulb (*Clytia hemisphaerica*)[145], or migrating nest of nematoblasts (developing cnidocytes), attached by cytoplasmic bridges and renewed from the same multipotent stem cell population as secretory cells and other neurons in the hydrozoan Hydra[49], or other non-interstitial less characterized renewal mechanisms in scyphozoan jellyfish[146], and sea anemones (see below). Cnidocytes represent a clear phylum-specific cellular innovation, treated as both a novel cell type, mechanosensory neural cells, or broadly part of the neural cell type complement specific to cnidarians depending on the era and the scientific field of study.

1.3 The sea anemone *Nematostella vectensis* as a cnidarian system for neural cell fate specification

1.3.1 Cnidarians in development: from induction to genetic laboratory models

Cnidarians have long been studied for their continuous cellular renewal and regenerative capacities. The historical prominence of the hydrozoan Hydra contributed to their representation as simple diploblastic animals with diffuse nerve nets, and provided early insight into the developmental mechanism underlying lifelong cell renewal. Over the past two decades, however, expanded genomic resources across the phylum and experimental work in multiple laboratory species have revised this view.

From the 18th century onward, with early descriptions and regeneration studies by the Swiss experimentalist Abraham Trembley, the freshwater polyp Hydra (Cnidaria, Hydrozoa) became the first cnidarian laboratory model. Its simple two-layered body plan, continuous tissue renewal, and asexual reproduction provided abundant material for transplantation experiments, which, in the early 20th century provided insights into fundamental developmental mechanisms such as induction[147]. Experiments at Marine Biological Laboratory on *Hydra* by American embryologist Ethel Browne Harvey (in addition to her graduate work at Columbia University under Edmund Wilson and Thomas Morgan) showed from tissue-transplantation experiments that part of hydra (the hypostome tissue) can induce a secondary body axis or a new head structure formed by the host tissue. This early identification of induction, published in 1909[148], preceded and likely informed grafting experiments in the 1920s in amphibians by Hilde Mangold and Hans Spemann (who had been sent her paper) in Germany, identifying a vertebrate organizer region that can instruct cell fate and patterning in adjacent cells (Spemann was later awarded the Nobel prize in 1935 "for his discovery of the organizer effect in embryonic development"). This pioneering work in a cnidarian experimentally highlighted a fundamental developmental principle. However, despite abundant adult material,

Hydra's embryos are not readily accessible, their numbers limited, and their 2 to 24 weeks embryonic dormancy phase is a limiting factor for genetic approaches. The development of genetically tractable models like the fruit fly *Drosophila melanogaster* (with the same Thomas Morgan eventually winning a Nobel in 1933 "for his discoveries concerning the role played by the chromosome in heredity"), laid the groundwork for developmental genetics research in selected laboratory model organisms, opening up detailed mechanistic studies of developmental processes[2].

Cnidarians encompass over 12,600 aquatic species (as listed in the world register of marine species), with diverse body plans and life cycles, including hydra, sea anemones, sea pens, jellyfish, and corals, with bi and tri-phasic life cycles, free living, colonial, symbiotic and parasitic species. Early developmental biology research in *Hydra* (Cnidaria, Hydrozoa) has shaped broader assumptions about cnidarian anatomy and nervous system organization: it is radially symmetrical, diploblastic, with a bi-layered body plan, contractility mediated by epitheliomuscular cells, a largely diffuse nerve net with neurons also embedded within epithelial cells. Embryos undergo prolonged dormancy, gonad formation can arise from the epidermis, and continuous displacement of each epithelial body layer drives position-dependent transcriptional changes in neurons and secretory cells[149]. While these traits are often presented in textbook as general of Cnidaria, they are perhaps more representative of the genus *Hydra*. Studies in the sea anemone *Nematostella vectensis* highlight non-epithelial muscle cells using orthologs of bilaterian contractility genes[150], and the expression of mesodermal and endodermal-associated molecular markers support a more complex, tripartite tissue organization[151], revising the simple diploblastic representation[152,153].

Neurons in cnidarians are organized in nerve nets without centralization, but their organization is far from homogeneous. Multiple lineages exhibit stereotypical condensations of neurite tracts, as in the hydrozoan *Clytia hemisphaerica*[108,110], other hydrozoans[154], cubo-

2. Several cnidarian species, including Hydra eventually became amenable to genetic perturbations approaches in the late 2000s, a century after Ethel Browne Harvey's experiments.

zoan and scyphozoan jellyfish, and anthozoan sea anemones[47]. Sensory camera-type "eyes" structures have evolved independently multiple times[155], as in cubozoans and scyphozoans, underscoring neural diversification within the phylum. Cnidarians are known for their extensive lifelong cellular renewal and regenerative abilities, but their neural cell types complement, and the cellular and molecular mechanisms of their specification and turnover in adult homeostatic renewal remain poorly understood outside of *Hydra*. In the sea anemone *Nematostella vectensis* (Anthozoa), single cell RNA-Seq atlases have revealed a diversity of neural molecular states[8,53], which, combined with its genomic resources, and experimental availability, position it as a key system for studying the developmental mechanisms of adult neural specification and diversification (see below).

1.3.2 The genetically tractable starlet sea anemone

The sea anemone *Nematostella vectensis* (also *Nematostella* hereafter), is a small burrowing sea anemone native to the North American East coast (and along the southern coast of the United Kingdom). From early description of spawning in the lab from wild-caught individuals over thirty years ago[156], it has emerged as a leading model anthozoan cnidarian about two decades ago, with the closing of its life cycle[157]. It became the first cnidarian species for which large amounts of embryos could be available in the lab year-round, opening up investigations into embryonic development and early body plan patterning and providing a comparison point with well-studied bilaterian animals. This critical milestone spurred an era of methods and tools development (further reviewed in chapter 2), and over the following decade established *Nematostella vectensis* as a leading genetically tractable cnidarian lab model species (reviewed in[158,159]).

Its genome was sequenced in 2007[160]. It was the first non-bilaterian genome assembled at the time, and was followed by the placozoan *Trichoplax adherens* (Placozoa)[161], the sponge *Amphimedon queenslandica* (Porifera, Demospongiae)[162], and *Hydra*'s in following years[66].

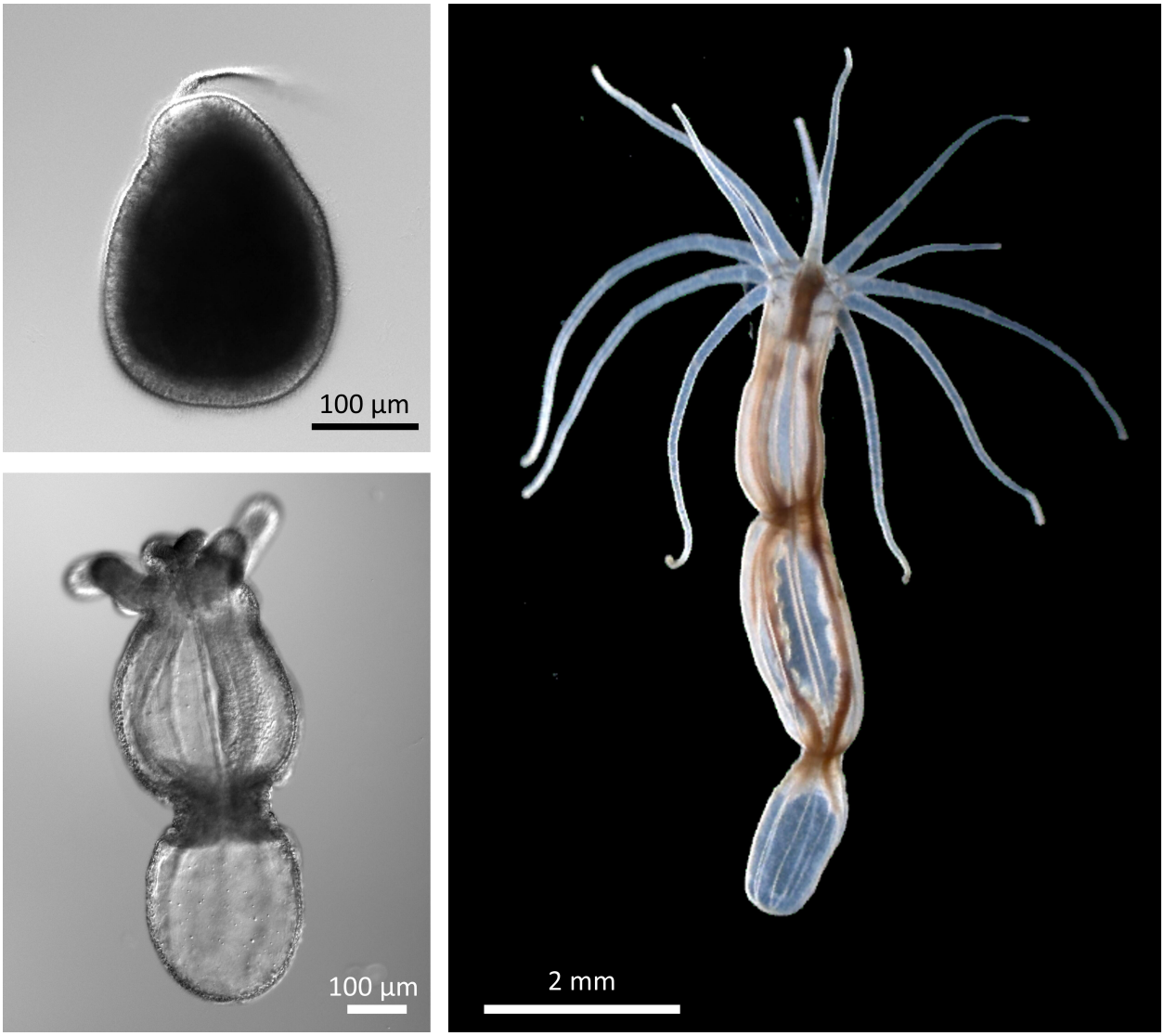

Figure 1.3: ***Nematostella vectensis*** **developmental stages**

Life stages from 4-day-old planula larva (top left, note the ciliated apical tuft at the top, and future oral opening at the bottom), primary polyp (bottom left, ca 11 days) and sexually immature polyp (right, ca 6-week-old, same stage as used in chapter 3).

Nematostella vectensis's genome (269 Mb in the latest assembly[163]) revealed a previously underappreciated level of gene families conservation with bilaterians, and especially with vertebrates, with many genes (1200+ reported) identified in *Nematostella* and in at least two of the three vertebrates genomes screened, but not in nematodes or the fruit fly, and in over 80% of comparable regions analyzed, exon-intron boundaries were conserved between

the sea anemone and human genomes[160]. Along with its conservation of signaling pathways receptors and effects, this highlighted a larger level of ancestral conservation of the eumetazoan developmental toolkit.

Nematostella (Fig. 1.3) reproduces both sexually, with separate males and females, and asexually through transverse fission, both in the laboratory and under natural conditions. Its life cycle is approximately 2-3 months in the laboratory, with lifelong growth and asexual reproduction, and its lifespan is currently undetermined. Embryonic development begins with equal radial cleavage of the fertilized egg, forming a hollow blastula. Gastrulation occurs via invagination (and ingression) at the future oral pole, and initiates the tissue structures that will become its internal mesenteries. The blastopore becomes the oral opening. The ciliated planula larva is free-swimming, with a conserved apical organ on its aboral end. Upon settlement (7 to 11 days in the lab), the planula undergoes a minimally indirect metamorphosis[164] (larval settlement), with limited tissue remodeling and elongation along the primary oral-aboral axis, including specification of four tentacle primordia and pharynx, as well as the development of the internal body cavity and maturation of the first two primary mesenteries (primary polyp stage). From 1-2mm-long primary polyps, animals progressively develop their eight mesenteries and add tentacles[165], with lifelong nutrient-regulated growth and asexual reproduction or shrinkage[166]. The overall body plan is conserved from juvenile to sexually mature polyps, with the maturation of the gonadal region within their mesenteries marking sexual maturity, at about 2-3 months in the laboratory.

Adults are solitary polyps, about 1-3 cm in length. Their body is a tube-like structure, with an oral opening with a pharynx region and tentacles for prey capture and ingestion at one end, and closed by the physa region at the other, with a small physal pore opening at the extremity. The body is comprised of an epidermal monolayer separated by an extracellular matrix layer (mesoglea), with a gastrodermal layer that continues into attached internal tissue structures called mesenteries. The oral opening is surrounded by 12 to 20

tentacles with both piercing cnidocytes and sticky non-piercing cnidocytes (spirocytes) for food capture and venom delivery, and is continuous with the central gastrovascular cavity, which functions as both a digestive chamber and a hydrostatic skeleton in the absence of mineralized hard parts. The primary body axis is oral-aboral, with the oral side originating from the embryonic blastoporal region following larval settlement. Its internal tissue structures, the mesenteries, are specified during embryonic and larval development, with Hox genes involved in patterning of the secondary body axis, the directive axis[167,168]. Mesenteries run longitudinally throughout its body column, tapering off on the aboral end, and are functionally regionalized, with circular (smooth-like) and longitudinal (parietal and mesenteric) muscles[150,169], bilateral nerve condensations with neural cell bodies, a gonadal and trophic region, and a cnidoglandular tract with secretory cells[151] and cnidocytes at its distal edge, in close contact with captured prey for digestion.

Like most sea anemones of the order Actiniaria, *Nematostella vectensis* (phylum: Cnidaria, subphylum: Anthozoa, class: Hexacorallia, order: Actiniaria, family: Edwardsiidae, genus: Nematostella), has a fixed number of mesenteries (8 in this species), established early during embryonic development, and thus a corresponding fixed number of neurite plexuses, independent of body size. These are paired longitudinal neurite tracts localized on each side of the parietal muscles, with uncharacterized cellular composition, connectivity and function. There are also sparse neural cells in the mesenteric muscle region in the mesenteries. While external morphology does not reveal asymmetry, the internal anatomy is bilaterally symmetrical (visible in the siphonoglyph and mesenteric muscles placement) rather than radially symmetrical, with symmetry breaking during embryonic and larval stages regulated in part by BMP signaling[170–172].

Its cell types include around eight broad classes, with additional uncharacterized cells[8,53]. These include epithelial epidermis cells, gastrodermis cells (internal cells with metabolic support and storage functions), mucous secretory cells and digestive secretory cells[151], muscle

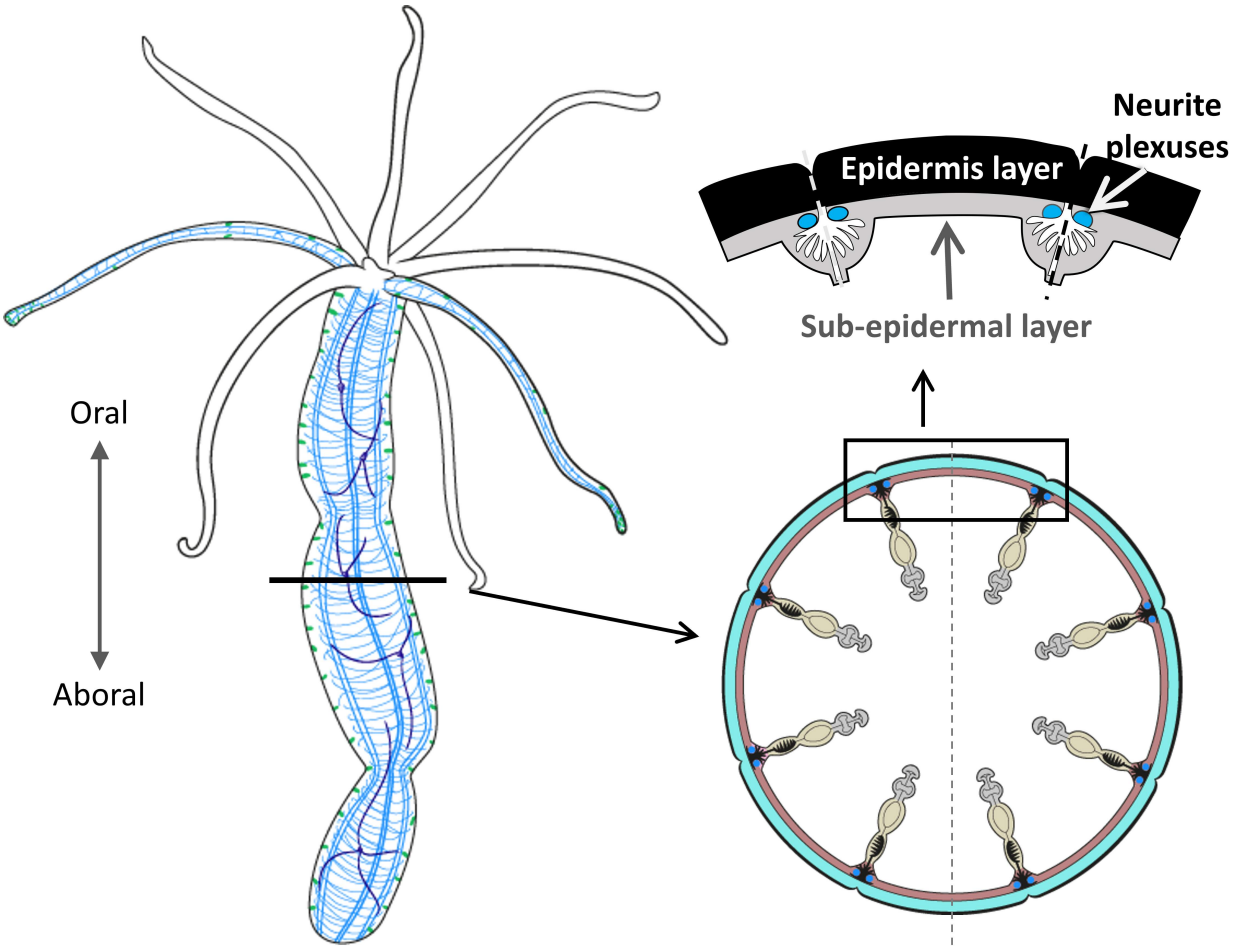

Figure 1.4: Polyp nervous system organization

Schematic of nervous system organization. (***left***) Distributed nerve net with epidermis-embedded neurons (dark blue), and nerve net tracts (blue), and cnidocytes (green). (***right, top***) Body wall anatomy, with the position of the neurite plexuses (in blue), and the extracellular matrix and parietal muscle (in white). (***right, bottom***) Internal anatomy, with the epidermis monolayer (cyan), sub-epidermal gastrodermis layer (brown), and mesenteries with parietal muscle and retractor muscles (in black) and trophic and gonadal region in yellow and ciliary and cnidoglandular tract in grey. Line schematics by Colette Plessier, from F.P. pictures and drawings.

cells[70,173], with longitudinal parietal and mesenteric muscles along the oral-aboral axis, circular muscles in the body wall and tentacle-specific retractor and circular muscles. There are also putative progenitors cells, and a diversity of neural cells, including cnidocytes and many neuronal states identified from single cell atlases[8,53,174].

1.3.3 Neural cell types in *Nematostella*

The nervous system is organized as a nerve net, with localized condensations of neurites and associated soma along the primary body axis. The organization of neural circuits and the functional roles of neural cell populations remains mostly uncharacterized. Based on morphology, three neural cell types are described: ganglion neurons (the largest class), cnidocytes, and putative sensory cells. For clarity, while cnidocytes are also considered neural cells, we use the term "neurons" hereafter to refer specifically to non-cnidocyte neural cells.

Characterization of neural cell types in *Nematostella*[47,119,175,176] has relied on whole-mount in situ hybridization of candidate genes (orthologs of bilaterian neural markers as well as cnidarian-specific genes), primarily in the early embryo and larval stages, transgenic reporter lines labeling defined cell populations[38,101,177,178], and single-cell RNA-sequencing atlases[8,53] (see methods overview in chapter 2). These approaches have identified both broadly expressed and cell-type-restricted molecular markers, with preliminary morphological descriptions for a small number of neuronal subtypes.

Broadly neuronal markers include an ELAV family member[36], which has been used to generate a broadly neuronal reporter line[38]. Prohormone convertases, which are endopeptidases of the proprotein convertase subtilisin/kexin (PCSK) family involved in neuropeptide maturation, are also expressed broadly in neurons. *Nematostella* pro-hormone convertases encompass orthologs of bilaterians subfamilies, as well as unclassified expansions with uncharacterized expressions[179].

Putative sensory cells embedded in the body wall epidermis express *FoxQ2d*, and display 2-4 short basiepithelial processes[180]. Their function and whether they are neurons remain unresolved, and they are labeled by an anti-bilaterian FMRFamide antibody, but are not labeled by the *Elav>mOrange* reporter line[180].

The transcription factor *Prdm14d* also marks a specific subpopulation of body wall neurons, localized primarily within the mesenteries[181]. *Prdm14d* is also expressed at low levels

in uncharacterized progenitors, which give rise to non-*Prdm14d*-expressing uncharacterized neuron subtypes[181].

Neuropeptides in *Nematostella*

Specific neuronal markers include neuropeptide precursor genes. These are often highly expressed, and thus can be clearly detected in in situ mRNA hybridization experiments, or antibody-staining based experiments. Indeed, in early studies, cross-reactive antibodies targeting vertebrate FMRFamide highlighted a dense network of RFamide$^+$ neurons throughout the animal from early planula to primary polyp stages[36,38]. From early studies of individual genes, to genome scanning and mass spectrometry approaches, there is now over a dozen precursors genes identified in *Nematostella*[52,92,98,100,101,105,171,177,182], in addition to recently deorphanized GPCR neuropeptide receptors from a large pharmacological screen from synthesized peptides[98]. Neuropeptides are secreted molecules, through controlled exocytosis, and their precursor gene thus contain a signal peptide (for synthesis in the endoplasmic reticulum) and usually a propeptide removed by cleavage during maturation. Each neuropeptide precursor genes can generate 1 to 30+ small peptides almost always with the same C-terminal residue[98,100,103–105]. Precursor genes are currently named for the C-terminal end of their short peptide, like a RPamide precursor (named RPa with **a** for the C-terminal amidation), and include RPa, HIRa, PRGa, QWa, WFSa, LRWa, GLWa among others, sometimes with multiple precursor genes[100].

Their precursors are often highly expressed, and some show regionalized expression, like VRHa and HIRa that are specific to the tentacular region. In the planula, an RPa and a GLWa are expressed in distinct neuron populations[183], and expression of some of these neuropeptides in published atlases noted their specific expression in neuronal subsets[92,98]. In adult polyps, one LWamide precursor is expressed in the neural tracts at the base of the mesenteries, as well as in tentacles, while an RPamide precursor is specifically detected in the

tentacles in juvenile polyps[8]. A reporter line for a LWamide precursor (*LWa-like>mCherry*) labels sparse neurons throughout the tentacles, pharyngeal region, mesenteric tracts and outside in the body wall, and partially overlaps with another neuronal subtype-specific marker named serum amyloid-like, from a serum amyloid-likeA>mCherry reporter line[177,178].

It is hypothesized that these neurally-expressed peptides act as neurotransmitters, with currently little characterization of their neurobiology in *Nematostella* beyond the recent deorphanization of select neuropeptide GPCR receptors[98], a GLWa knock-out line with a developmental delay phenotype[111], and sparse behavioral assays using exogenous peptide soaking (see[101]). There are studies of neuropeptide function in other cnidarians, like scleratinian corals, where RFamides can induce polyp contraction[184], as well as studies in hydrozoans highlighting roles in development, contraction and others. In addition to these short amidated peptides, there are several insulin-like prohormones, expressed in digestive secretory cells in the cnidoglandular region, with the exception of an insulin-like peptide with a "neural-like" expression pattern in polyps[151]. Its function is uncharacterized.

Moreover, a screen for toxin-like genes identified multiple peptide precursors specifically expressed in neurons[101], and generating detectable short peptides. These include *Aeq5-like1*, (similar to Acrorhagin 5a from the sea anemone *Actinia equina*), and related protein Aeq5-like2, as well as *Class8-like*, (similar to an Avd8e toxin), and several ShK-like peptides, shared a structure and a cysteine arrangement with the potassium channel blocker toxin ShK from the sea anemone ***S****tichodactyla* ***h****elianthus*. *ShK-like2* is expressed as two isoforms generated by alternative transcription start site usage (*ShK-like2a* and *2b*). All toxin-related genes are expressed in neurons during planula and primary polyp stages, except *ShK-like1*, which is expressed in cnidocytes. For example, in situ hybridization for *Class8-like* labels multipolar ganglion neurons with long processes in the body wall but not tentacles. It is unclear if these are used as neurotransmitters, with incubation with *ShK-like1* or *ShK-like2* peptides inducing body contraction; and morpholino-mediated knock-down of *ShK-like2* causing a

developmental delay[101], although the underlying mechanisms are unknown.

Across these neurally-expressed peptide precursors, despite clear neuronal expression, direct evidence for synaptic transmission or neuromodulatory function is currently lacking. These genes are currently used as subtype-specific neural marker genes.

Cnidocytes in *Nematostella*

Cnidocytes are broadly distributed throughout the animal, embedded in the epidermis monolayer, in the pharynx and in the mesenteries mostly in the cnidoglandular tract, and in the tentacles, with an increased concentration at the tips[185]. *Nematostella* also produces structures enriched for nematocytes, the nematosomes[186] that are released by its mesenteries into its internal cavity, and maternally loaded into the egg jelly mass along with toxins (like *NvePtx1*[187]) and thought to help protect eggs from predators. Cnidocytes encompass three subtypes described morphologically[132], including two closely related nematocyte (piercing cnidocytes) subtypes: the smaller and most abundant basitrichous haplonema (12μm in length, 3μm in diameter, $\tilde{9}$0% of cnidocytes in the body, and $\tilde{4}$0% in the tentacles), and the larger microbasic mastigophores (17-22μm in length, 3μm in diameter). The non-piercing spirocyte is found in the tentacles (around 25-30μm-long). At the molecular level, in single cell RNA-Seq atlases, two main subtypes are identified: nematocytes (with subsets hypothesized to match the morphologically identified nematocytes) and spirocytes[8,53], in addition to a cnidocyte-associated secretory population[187].

Cnidocytes express several minicollagen genes, including *N-Col1*, *N-Col3*, *N-Col5*[132], a spinalin-like gene (although its domain composition appears distinct from that of hydra's[130]), nematogalectin-like genes, used to build its capsule and harpoon structure. An anti-NvNCol-3 antibody recognizes maturing nematocytes but not the mature ones, likely due to inaccessibility of the CRD domain epitope[132]; although it still recognizes the spirocyte capsule. Spirocytes also express similar genes, with some minicollagens that are either

nematocyte-specific (like one *minicollagen1* gene) or spirocyte-specific (like *N-Col5*), and also express *FoxL2*. *N-Col3* mRNA is further cleared during cnidocyte maturation when the capsule is being built[188].

Additional markers of cnidocytes and cnidocyte-associated secretory cells include the largely uncharacterized **N**ematocyte **E**xpressed **P**roteins (NEP) genes. NEP genes were identified by mass spectrometry of the soluble protein fraction from discharged cnidocytes in *Nematostella*, with the corresponding genes named as *NEP1* to *NEP20*, largely from unrelated protein families[189]. Their expression is restricted to cnidocytes and poorly characterized cnidocyte-associated secretory cells[187].

1.3.4 Neurogenesis and cnidogenesis

Embryonic neurogenesis

Insights into neurogenic regulators in *Nematostella* comes from embryonic perturbation studies, as later developmental stages (including juvenile polyps) remain technically less accessible. Gene function has been assessed through morpholino or shRNA-mediated knockdown, mRNA overexpression, pharmacological manipulation, and a few constitutive knockouts, typically monitored via changes in neural marker expression (e.g. *Elav*, neuropeptides, or cnidocyte structural genes) (see also chapter 2).

Overall, embryonic neurogenesis in *Nematostella* relies on both ectodermal and endodermal cellular sources that remain incompletely characterized[38]. Dispersed ectodermal cells already express neurogenic regulators at the blastula stage (*SoxB(2)* (an unclassified basal SoxB gene) and an atonal/neurogenin related gene (named *Nem10*[39]/*NvAth-like* (atonal homolog-like)[51]/*NvArp3*[171])[50,51].

Early embryonic neurogenesis is considered to be independent of BMP (Bone Morphogenetic Protein, part of the Transforming Growth Factor-*beta* (TGF-*beta*) superfamily) signaling, while modulation of BMP signaling can have pro- or anti-neurogenic effect at the

planula stage[171]. Moreover, pharmacological inhibition of MAPK signaling in embryos decreased neurogenic markers[177], with abnormal embryonic development. The specific role of these signaling pathways in neural fate acquisition is poorly understood. Neural phenotypes may reflect direct effect on fate specification and/or indirect consequences of broader embryonic patterning changes and domain expansions, especially in light of recent work on germ layer specification involving MAPK, *beta*-catenin, and Notch signaling[152,153] to establish three layers reminiscent of bilaterian ectoderm, mesoderm and endoderm from earlier work[151]. Interpretation is further limited by the lack of inducible or cell-type-specific perturbations, as many studies rely on changes in neural marker abundance or expression domains, which can shift secondary to patterning alterations rather than reflecting specific roles in neural fate acquisition. Similarly, Notch signaling inhibition results in a lack of cnidocytes[190], later refined as the accumulation of immature developing cnidocytes[51,191], and an increase in cells positive for neuronal markers in the embryo and planula[51]. Notch appears to play a negative role in neurogenesis in the embryo, upstream of *SoxB(2)*, *atonal-related* and *AshA*[51].

Embryonic regulators of cnidocyte fate

Several studies identified transcription factors involved in specification of cnidocyte fate in embryos[136,188,192,193], these include *PaxA*[188], the zinc finger named *ZNF845*[193] (not an ortholog of vertebrate ZNF845), *NvSox2* (not an ortholog of mammalian SOX2, nor a *SoxB* gene but a basally-branching uncharacterized Sox gene not associated to known bilaterian SOX gene families and named *NvSox2*, here identified as *Sox_unk* to minimize confusion)[43,136], the nuclear receptor *NvNR12*, and the bZIP *cnidoJun*[194]. *ZNF845* is downstream of *Sox(B)2* and upstream of *PaxA*, and *NvSox2* is involved in fate specification of cnidocyst morphological features. Homeodomain transcription factor *POUIV* (ortholog of mammalian *Brn3*) is required for the maturation of cnidocytes, and not for the initial speci-

fication of the cnidocyte lineage, as indicated by the accumulation of immature cnidocytes in a constitutive *POUIV* knock-out in both planula and primary polyps[120]. This requirement of *POUIV* for cnidocyte maturation is cell-autonomous[195]. *POUIV* is also present in tentacular mechanosensory cells[196,197], alongside developing and mature cnidocytes. Knock-outs of Lysine-specific demethylase 1 (*Lsd1*) and Co-repressor of REST (*CoREST*), components of the CoREST complex involved in cell differentiation and homeostasis across bilaterians, showed defect in cnidocyte maturation, with accumulation of immature cnidocytes and a cell-autonomous rescue[195,198].

Moreover, single-cell RNA-Seq atlas identified a distinct maturation state for cnidocyte from in silico reconstruction and noted the expression of calcium-binding proteins and ion channels, along with a decrease in minicollagen gene expression[53]. This is further supported by earlier studies indicating that mature cnidocytes no longer contain minicollagen mRNAs even as they have mature capsules[132,188].

Embryonic regulators of neurons and cnidocyte development

SoxB(2), and an atonal-related protein have been identified as upstream regulators of both peptidergic neuron and cnidocyte development in embryonic/larval stages, but whether they act in the same progenitor pool is unknown. Furthermore, an achaete-scute related gene, *NvAshA*[52] (also known as *Nem17*[39]) is a peptidergic-specific embryonic neuronal regulator. And there is partial co-expression of *SoxB(2)* with *FoxQ2d* (a marker of putative sensory cells) at the mid-planula stage, consistent with a possible role for *SoxB(2)* in specification for these cells as well[180]. *NvInsm1* (*insulinoma1*) is a transcription factor expressed broadly across $Elav^{+}$ neurons and secretory cells, but not in cnidocytes, and is downstream of *SoxB(2)* at early life stages. Downregulation of *insm1* decreases secretory cell markers, but not peptidergic neuron markers at planula stage[199]. It is unknown if it is part of the same peptidergic and secretory lineage. Exogenous GABA or baclofen (a $GABA_{B1R}$-specific

agonist) induce a delay in metamorphosis, and decrease expression of upstream neuronal regulators (like *SoxB(2)*) as well as downstream neuronal effector *Elav*, with more disorganized neurons and shorter neurites; and also decrease expression of cnidocyte regulators (like *PaxA*) and effectors (like the minicollagen gene *Ncol3*)[200]. This developmental delay phenotype is similar to the GLWamide KO, although it is unknown if the molecular mechanisms are similar[111].

Overall, embryonic/larval neurogenesis is not restricted to post-gastrulation neurectoderm and remains incompletely understood. Upstream embryonic regulators of peptidergic neuron and cnidocyte development include *SoxB(2)* and *Nem10_Atonal-related*, with *AshA* identified as a peptidergic-specific regulators, and *ZNF845*, *PaxA*, *NvSox_unk(NvSox2)*, *NR12* and *POUIV* (POUIV also regulates tentacular hair cells) as cnidocyte-specific regulators. Specifically the cellular populations and lineage restriction that give rise to neural cells (including cnidocytes) are not characterized, with some regulators described as specific to peptidergic neurons or to cnidocytes. Neural cells identities overlap at the molecular level from atlas data between polyps and larval stages[8,53], with a few larval-specific neurons disappearing during larval settlement[183], supporting the conservation of differentiated neural cell states across life stages. And LWamide-like neurons are a small subset of the total neuronal population and their number scale with body size in the juvenile polyp[178,201].

Reporter lines for markers of the putative progenitor cells identified in scRNA-Seq atlases[8,53], including *Piwi1*, and *Vasa2* support that a pool of cells expressing these markers can give rise to somatic derivatives, including *Elav*$^{+}$ neurons[202,203] in the polyp. Their cellular lineages and restrictions are uncharacterized.

1.4 Chapters overview

The growing adult sea anemone *Nematostella* (Cnidaria, Anthozoa) provides a tractable in vivo system to examine neural lineage and fate specification in a lifelong homeostatic growth

context, independent of embryonic patterning and large-scale morphogenetic events. While transcriptomic studies have described adult neuronal states, the lineage relationships and regulatory programs that generate distinct neural populations remain to be characterized. In this dissertation work, we leverage the genetic tractability of *Nematostella* to investigate the cellular and molecular bases of neural renewal in an anthozoan cnidarian. By combining photoconvertible pseudo-lineage tracing with single neuron-resolution imaging and targeted single-cell transcriptomics, this work reconstructs progenitor dynamics, lineage relationships, and transcriptional trajectories underlying adult neurogenesis. Chapter 2 develops quantitative in vivo tools to study adult neural renewal in *Nematostella*, including whole-organism photoconversion–based cell tracing from new transgenic lines, high-resolution imaging of individual neurons within the nerve net, spatial mapping of newly generated cells relative to anatomical structures, and fluorescence-based isolation of reporter-labeled cells for molecular profiling by single-cell RNA sequencing, with oral regeneration post-amputation as an example. Chapter 3 then uses these tools to reconstruct cellular and transcriptional trajectories underlying adult neurogenesis. We find newly generated neural cells distributed across nerve net territories, and their differentiation is characterized by identifiable progenitor and end states, with proportional scaling of subtype identities. This work further identifies lineage restrictions, and gene programs linked to neural cell classes and subtypes. In the Discussion (Chapter 4), these findings are integrated to provide a broader view of neural fate specification and stem cell systems for cellular renewal and we consider their implication for the evolution and diversification of neural cell types. Together, the work provides a lineage-informed view of adult neural renewal in a cnidarian model and contributes to understanding how diverse neural cell types can be specified in an early-branching eumetazoan lineage.

CHAPTER 2

METHODS FOR INVESTIGATING CELLULAR RENEWAL PROCESSES IN THE SEA ANEMONE *IN VIVO*

2.1 Introduction

Our understanding of biological processes is limited by what we can technically observe, measure, manipulate, and infer. Cnidarian systems have been studied for several centuries, but major technical advances in the last twenty years have opened up new areas of investigations. These stepwise developments, briefly reviewed below, together with the increasingly accumulating biological knowledge, have made it possible to begin tracking cellular lineages and cell type specification trajectories, thus investigating the mechanisms underlying their life-long renewal abilities. This would help inform our understanding of animal cellular renewal processes, and provide insight into the evolutionary history of neural cells, an animal-specific cell type.

Over the last two decades, *Nematostella vectensis* emerged as a leading cnidarian laboratory model species, largely due to its availability in the lab, and the resources generated for this species and associated eras of technical advances. In the first decade (2000s), foundational work established protocols for year-long lab availability of all life stages, *in situ* hybridization for spatial mRNA detection, morpholino-based and pharmacological perturbations in embryos, and assembled a first scaffold-level genome with associated gene models[157,160,204–206]. In the second decade (ca. 2010s), development of meganuclease-based stable random genomic insertion[169] led to the generation of several tissue and cell-type-specific reporter lines[38,50,52], followed by development of TALEN and CRISPR-based gene knock-out methods[120,207] and later CRISPR-based knock-ins reporters[194]. In parallel, genomic resources for *Nematostella* were expanded, with improved gene annotations and epigenome-wide marks[208]. Collectively, these facilitated the integration of molecular

evolution analyses with functional experiments, providing insight into a diversity of biological processes, including microRNA pathways evolution, secondary body axis patterning, germ layer evolution, cnidocyte development, and embryonic neural development, among others[36,50,151,170,171,177,188,191,209–212]. In parts due to the availability of embryos and embryonic-restricted gene knockdown approaches, studies of neurogenesis in *Nematostella* focused on embryonic stages and targeted select genes of interest (orthologs of known bilaterian neurogenic genes from the *Elav*, *SoxB*, *atonal* and *achaete-scute* families). These foundational works characterized molecular and cellular aspects of embryonic neural processes in *Nematostella*[36,38,47,50–52,171,177,201].

Complementary to these targeted approaches, whole-organism single cell transcriptomic atlases provided a global view of the cell populations present in the animal across life stages[8,53]. By characterizing the transcriptome-wide programs of all cellular states, these atlases also expanded the candidate gene approach from targeting orthologs of known bilaterian genes to genome-wide candidates present in cells of interest. These atlases also revealed the previously underappreciated molecular diversity of neuronal cell states, with the characterization of several dozens of transcriptionally-defined neural states in *Nematostella*. Atlases also identified putative "progenitor" states[8,53], tentatively labeled as such for their lack of cell-type-specific marker genes and expression of cnidarian orthologs of genes associated with bilaterian germline stem cells, and other multipotent stem cells (discussed in section 4.3) . And a hypothetical shared neuroglandular lineage for neural and secretory cells has been proposed solely from the UMAP 2D embedding representation of the data[53].

However, cell atlases are static maps of the cell states present in the organism, and do not provide experimental cellular lineage evidence. Computational reconstructions of cellular lineages from such static snapshot is challenging[213], and further complicated in this case by the number of cell states described (over two dozens for neural cells, depending on granularity) and the lack of *a priori* knowledge of which cell states are intermediate versus

terminally differentiated populations, as they might all be concomitantly present across life stages. The differentiation trajectories, source population(s) and regulators of lifelong neural fate specification are experimentally undetermined.

In contrast, cellular lineage can be determined via genetic cell labeling approaches, for instance through the inheritance of reporter fluorescent protein like GFP[1] in progeny cells and/or genetic conditional labeling strategies. Thus, combining scRNA-Seq snapshots with genetic labeling can link molecular identities of neuronal cells to their cellular lineages and differentiation trajectories.

Both conditional and inducible genetic approaches are widely used in bilaterian genetic model organisms for lineage tracing (mouse, zebrafish, *Drosophila* flies, *C. elegans* nematodes), but these are often two component systems and rely on characterized genetic drivers. In *Nematostella*, the longer generation time (2.5-3+ months in the lab), the lack of established two-component genetic reporter tools, and the small number of described drivers were limiting factors for targeted genetic approaches, especially as the cells and processes underlying adult cellular renewal were unknown. Here, to study lifelong neurogenesis *in vivo*, we generated transgenic constitutive reporter lines, using an established broadly neuronal driver (*Elav* driver, from[38]), as well as from two new candidate genes (*FoxL2* and *SoxC*) identified in single cell atlas[8].

A major limitation of using constitutive fluorescent reporter to investigate cellular lineages is discriminating between expressing cells and non-expressing progeny cells that also carry the fluorescent reporter protein. This is especially challenging in animals with continuous growth and active cellular renewal like cnidarians, where, at any given stage, all

1. The most influential contribution of cnidarians to modern biology is the characterization of the green fluorescent protein GFP, identified in the jellyfish *Aequorea victoria* (Cnidaria, Hydrozoa), in the 1960s by Osamu Shimomura. The gfp gene was cloned in the 90s (Douglas Prasher), and used in living organisms as a genetic marker by Martin Chalfie. From this characterization, work by Roger Tsien and others engineered several related fluorescent proteins with distinct spectra, enabling *in vivo* tagging of multiple cells or proteins of interest. This discovery was awarded the Chemistry Nobel prize in 2008 (to Martin Chalfie, Osamu Shimomura and Roger Tsien).

differentiated and differentiating cellular states might be present, and thus, it is not possible to utilize developmental time as a proxy for differentiation stage.

However, this very limitation can be leveraged for lineage tracing by using a photoconvertible reporter protein. In a modified cellular pulse-chase experimental setup, the photoconverted reporter protein can identify cells that previously expressed the reporter but no longer express it. These cells represent the non-expressing progeny of actively expressing cells. Continuous reporter expression during this temporal "chase" period post-photoconversion enables simultaneous tracking of both active and past expression states. This provides a non-invasive way to investigate cellular lineages *in vivo*.

To investigate lifelong neurogenic processes in the sea anemone *Nematostella vectensis*, we developed a transgenic reporter approach using a photoconvertible reporter protein: KikumeGreenRed (KikGR thereafter)[214]. KIKGR is a GFP-like spectrum fluorescent protein that forms a homotetrameric complex and that can be irreversibly cleaved by UV-spectrum light, thus shifting to a distinct broad orange-red spectrum (excitation/emission maxima: 507/517nm (non-photoconverted) to 583/593nm (photoconverted))[215]. The first green-to-red photoconvertible fluorescent protein, *Kaede*, was discovered in the open brain stony coral *Trachyphyllia geoffroyi* (Cnidaria, Anthozoa, Hexacorallia, Scleractinia)[216]. This endogenous protein emits a GFP-like fluorescence that can permanently shift to a broad orange-red fluorescence spectrum upon UV or near UV illumination. Studies of the molecular bases of its photoconvertible properties identified the key residues involved and led the the rational engineering of the KikGR from a GFP-like protein isolated in the stony coral *Favia favus*, also a scleractinian coral, with similar photoactivable properties as *Kaede* but increased brightness[215]. This property is critical for *in vivo* imaging. KikGR has been used as an optical marker for cellular and subcellular labeling and tracking *in vivo* and *in vitro*, as well as for multi cell-scale precise spatial labeling combined with scRNA-Seq in the developing mouse limb[214,217].

In this dissertation work, we combined the transcriptome-wide scRNA-Seq characterization of cell states with a targeted genetic labeling approach to investigate the cellular lineages and transcriptional programs that give rise to differentiated neurons in the sea anemone *Nematostella vectensis*. Thus, we generated stable constitutive photoconvertible reporter lines expressing the KikGR fluorophore in the sea anemone *Nematostella vectensis* under the control of either a published broadly neural driver (*Elav>KikGR* line, driver from[38]), or of new regulatory elements identified from the expression pattern of their endogenous gene the single cell atlas using the associated chromatin accessibility and available epigenetic marks data (*SoxC>KikGR* and *FoxL2>KikGR* reporter lines)[8].

2.2 Sections overview

This chapter presents methodological advances for investigating adult developmental processes in the genetically tractable sea anemone *Nematostella vectensis*. These approaches are then utilized in chapter 3 to identify cellular lineages for neurogenesis in the homeostatic growth context.

First, section 2.3 describes the generation of three new photoconvertible transgenic reporter lines (*Elav>KikGR*, *FoxL2>KikGR*, *SoxC>KikGR*), and a single-cell resolution imaging method, as well as partial photoconversion experiments in *FoxL2>KikGR* animals, shedding light on the temporal stability of the nerve net organization.

Section 2.4 then describes an efficient whole-organism fluorescent reporter photoconversion method for polyps and the imaging-based single-cell resolution characterization of both photoconverted progeny cells and newly-expressing cells in live adults, with a proof-of-concept tracking *SoxC>KikGR*$^{+}$ cells during oral regeneration.

To quantify cellular dynamics from *in situ* 3D-imaging, section 2.5 describes a cell segmentation approach associated to spatial registration applied to live confocal z-stacks to quantify and localize green, orange and red cells in the body wall nerve net, highlighting

spatial distribution changes across nerve net structures during cellular renewal.

Finally, this chapter details in section 2.6 a combination of established methods to target fluorescently labeled cell populations of interest in adult *Nematostella vectensis* for single cell RNA-Sequencing-based molecular profiling, with the characterization of amputation-induced regenerated and uninjured oral regions. This section (2.6) is reproduced from a published protocol chapter[218].

Together, these methodological contributions lay the foundation for the targeted study of adult neurogenic processes in sea anemone detailed in chapter 3, and provide early insights into the tissue dynamics and cellular processes at play during homeostatic growth and injury-induced regeneration in the sea anemone *Nematostella vectensis*, further discussed in chapter 4.

2.3 Reporter-based *in vivo* imaging of neural cells

2.3.1 Three photoconvertible transgenic reporter lines

New candidate genes for reporter drivers were identified from the initial single cell atlas, due to their specific expression in the putative progenitor cell population. Two candidate genes for reporter drivers were identified from the first RNA-Seq atlas[8] as they were expressed in the putative progenitor cell population, thus targeting these cells and enabling tracking their progeny. The two selected genes were: *FoxL2* and *SoxC*, with their respective gene identifiers: LOC5520088 and LOC116611748 from the jaNemVect1.1 *Nematostella vectensis* genome assembly (from the Darwin Tree of Life (DToL) project – Wellcome Open Research[163] and associated NCBI GCF_932526225.1 gene models. Reporter plasmids were generated by cloning the genomic regions (2.2 to 2.4kb) upstream of the ATG start codon of either *Elav*(from[38])/*FoxL2* or *SoxC*, to drive expression of the KikGR fluorescent protein (from[214]); in the pNvT-MHC::mCh plasmid backbone(Addgene plasmid #67943)[169] (see also Methods in chapter 3). Stable transgenic lines were established from mosaic F0 animals

using meganuclease-mediated transgenesis with the donor plasmid above co-injected with I-SceI meganuclease enzyme (modified from[169]).

Clarifying *KikGR*$^{+}$ cells nomenclature

Nematostella genes and proteins are variably identified throughout the literature: gene names and the gene model identifiers are not standardized, and conversion based on sequence similarities often result in one-to-many or many-to-many matches between gene model sets. There are at least five publicly available set of gene models in use (v1g, NVE, NvERTx, NV2, LOC, from their starting identifying scheme), respectively from three different genome assemblies (JGI, JGI, JGI, Technau/Gibson lab, Wellcome Sanger Institute/Darwin Tree of Life, respectively); and genes are written *gene*, *Gene*, *NvGene*, *Nvgene*, with multiple names in use for the same protein sequence across different publications, making the unambiguous identification of genes from one primary literature source to the other somewhat challenging.

For standardization, genes are in priority uniquely identified as *Gene* with the jaNemVect1.1 LOC-based gene model identifiers (NCBI gene models from the Wellcome Darwin Tree of Life *Nematostella vectensis* assembly); and the full gene name is standardized as "*Gene* - LOCidentifier" or "*Genename1_genename2* - LOC identifier" with multiple historical annotations listed. Note that the protocol reproduced in section 2.6 was published prior to the availability of this genome and associated gene models and thus uses the same merged annotation as in[8] with the JGI initial genome assembly[160] for gene models.

The protein product of a gene is usually indicated as PROTEIN, but this has not been the case across the *Nematostella* literature either, with *Protein* and *NvProtein* used as well. So while the product of the KikGR gene is the KIKGR protein, its presence in either its native (KikumeGreen, KIKGREEN) or photoconverted form (KikumeRed, KIKRED) can be confusing. For clarity, we refer throughout the paper as cells that contain the KikGR fluorescent protein (in either state) as *driver>KikGR*$^{+}$ cells, and specify whether these are

red cells (with photoconverted KikGR and lacking new KikGR signal), green cells (with the new unphotoconverted KikGR signal but not the photoconverted KikGR signal), or orange cells (with both). Thus, a *SoxC>KikGR*$^{+}$*/Elav>KikGR*$^{+}$*/FoxL2>KikGR*$^{+}$ cell contains some KikGR protein (photoconverted, unphotoconverted, or both), but is not necessarily a cell actively expressing this driver.

2.3.2 Single-neuron resolution imaging in live polyps

Detection of fluorescent-reporter-expressing cells in *Nematostella* has commonly relied on antibody-based detection for signal amplification in fixed tissues, via a secondary fluorophore-conjugated antibody[38,50,51,120,181]. However, this approach was not feasible or desirable in our case, as we aimed to distinguish photoconverted KikGR from newly synthesized KikGR across cells, and no antibody specifically recognizing the photoconverted versus the unphotoconverted form of KikGR exists. As the nervous system is organized as a distributed nerve net, identifying the anatomical locations at which new neurons are added is particularly informative. Mapping the spatial localization of KikGR^{+} cells and integrating reporter expression information (i.e. presence of the red photoconverted tracer label) can therefore reveal whether new neuron addition is biased towards a specific growth zone. Thus, we established a non-invasive method for *in situ* imaging of KikGR^{+} cells in intact, live polyps, at single cell resolution within the body wall.

We developed a custom imaging chamber for cellular-resolution non-invasive imaging of nerve net neuron in intact, live adult *Nematostella vectensis*. This simple imaging chamber consists of an aluminum or hard plastic plate with a Ø2cm central opening, allowing the animal to be mounted in seawater medium between two coverslips (bottom: 24x50mm, ca 0.145mm thickness and top: square 22x22mm, ca. 0.17 mm thickness). The spacing between the two coverslips is provided by 5 layers of regular transparent tape (Scotch, ca. <0.5mm in total chamber depth). Animals are alive and immobilized in a 3.5% (w/v) solution of

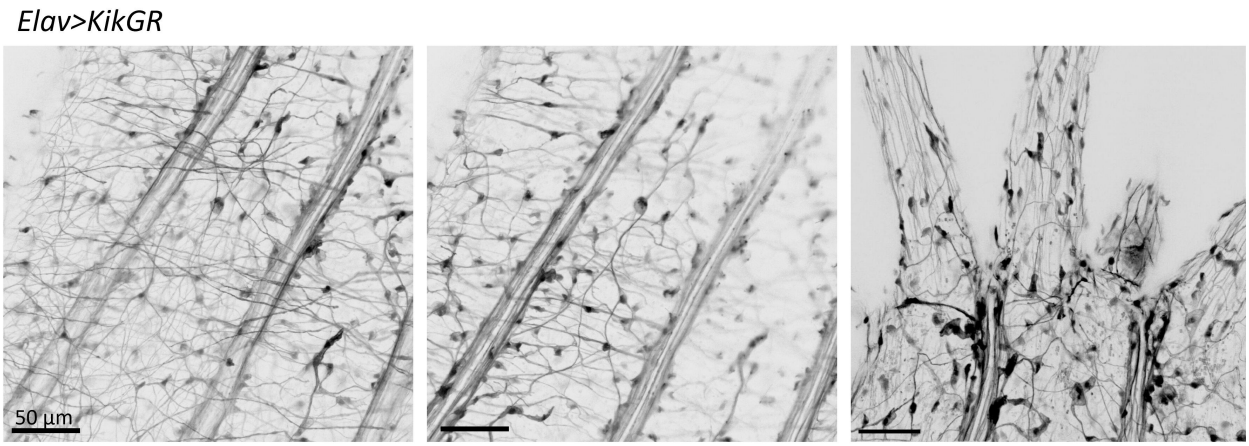

Figure 2.1: Single-neuron resolution imaging in live *Elav>KikGR* polyps

Confocal imaging (40x) of *Elav>KikGR*$^+$ intact live polyps in the body wall nerve (left and middle, two distinct optical layers) and at the base of the tentacles highlighting individual neurons. Note that the neural plexuses stop at the base of the tentacles. Scale bar is 50 μm.

magnesium chloride ($MgCl_2$) in *Nematostella* medium, according to standard protocols (see section 2.6). The combination of external compression, which flattens the body wall surface, and the internal hydrostatic pressure from the live polyp, which keeps the tissue under tension, is critical for generating high-quality images. This confining open design enables gas exchange, and allows imaging with a single-cell resolution in intact, live polyps at 40X magnification with a laser scanning confocal microscope Fig. 2.1.

Elav>KikGR expression broadly mirrors what has been previously described for *Elav>mOrange*[38], with the imaging chamber and increased brightness of the KikGR fluorophore as compared to mOrange enabling its clear detection in live animals, with a cellular resolution and without relying on antibody-based histoimmunofluorescence signal amplification.

Consistent with previous descriptions of *Elav>mOrange* expression[38], the majority of *Elav>KikGR*$^+$ cells throughout the animal display a neuronal morphology, with cell bodies detected primarily in the sub-epidermal layer level, across both the neurite plexuses, and in the inter-mesenteric region in the body wall, and throughout the tentacles.

The *Elav* regulatory element driving KikGR expression is highly expressed, but we could also achieve single-cell resolution imaging with the two new drivers we generated, which are

expressed at a lower level across diverse cell types.

The *Elav* regulatory element drives robust KikGR expression, with ongoing expression across neurons (see also chapter 3), resulting in high KikGR protein levels across cells. By comparison, KikGR expression driven by the two new reporters described here *FoxL2>KikGR* and *SoxC>KikGR* is overall much lower. This is consistent with the higher endogenous mRNA recovery of *Elav* compared to *SoxC* or *FoxL2* (as is typical for transcription factor), and with these driver being largely downregulated in terminally differentiated neurons (as shown in chapter 3).

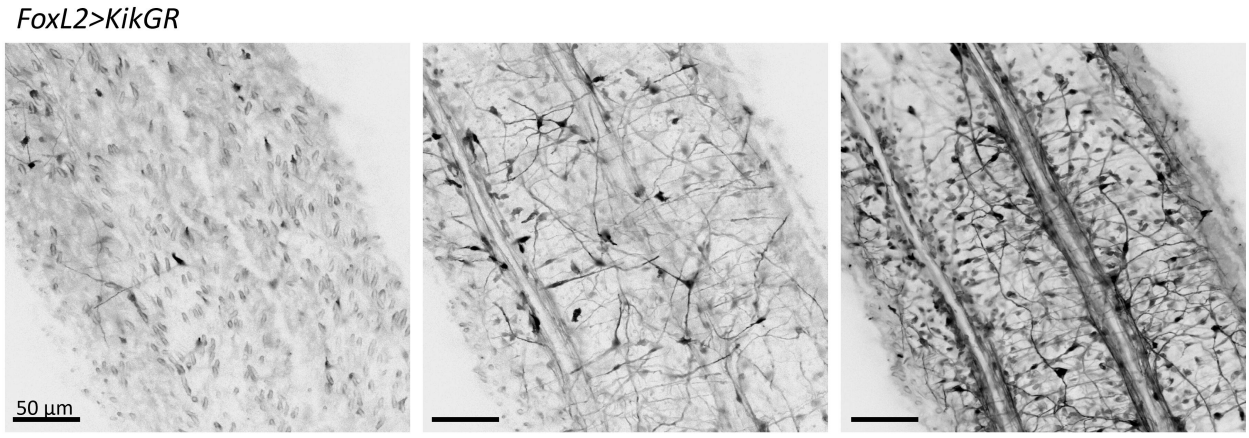

Figure 2.2: Single-neuron resolution imaging of a live *FoxL2>KikGR* polyp

Body wall confocal imaging (40x) of a *FoxL2>KikGR*$^+$ intact live polyp (same area, distinct optical layers) highlighting cnidocytes, neurons, and uncharacterized cells. Left to right from the outside of the animal towards the internal body cavity. Note the clearly visible double-tracts of the neurite plexuses and the multilayered neuron bodies localization. Scale bar is 50 μm.

Despite the lower level of KikGR signal, using different acquisition parameters also enabled single-cell in situ imaging for *FoxL2>KikGR* and *SoxC>KikGR* polyps. Indeed, *FoxL2>KikGR*$^+$ cells signal is an order of magnitude lower than *Elav>KikGR*$^+$ cells, except for tentacular spirocytes, which have high levels of KikGR signal, consistent with the high *FoxL2* endogenous expression in these cells. Even in cells with a low KikGR signal (i.e. across the body wall), the cytosolic subcellular localization of the KikGR protein (both native and photoconverted) highlights cell shape outline (including neurites), while being

excluded from intracellular structures such as the cnidocyte capsule, secretory granules, and cell nuclei, allowing morphology-based identification of cell types in intact animals from live imaging, even for cells where it is lowly expressed.

Neurons are also labeled by the *FoxL2>KikGR* driver, but its expression is much broader: encompassing a large array of cells, including neurons, sparse secretory cells, cnidocytes (developing and mature), and several distinct morphologically uncharacterized cells as well Fig. 2.2. Accordingly, in sexually immature adult polyps, there are between 1 in 4 to 1 in 3 cells in the body wall with mesenteries, as assessed by flow cytometry (see also chapter 3). This high fraction is in part due to the expression of this driver, at low levels, throughout epidermis epithelial cells. It is currently unclear whether this mirrors endogenous *FoxL2* expression (very lowly detected in single-cell RNA-Seq data, for several technical reasons), or if this might represent ectopic transgene expression. Expression in epidermis epithelial cells is uniform, and continuous, at low levels.

2.3.3 Nerve net stability highlighted by longitudinal live imaging

Polyps did not exhibit adverse effect from the mounting and imaging process and could be re-imaged multiple times. The latter property prompted us to test the tracking of tissue dynamics by re-imaging the same area over a week in intact animals, to examine changes in the body wall nerve net. This highlighted the presence of neurons with the same position and neurite branching pattern, hinting at a previously uncharacterized temporal stability in nerve net organization, at least at the week-long temporal scale assessed.

Repeated imaging over several days of the same area relies on the ability to localize that region in a polyp at each imaging session. This is challenging due to the small area imaged at the confocal at single-neuron resolution (less than 4% of body wall surface area at around 6 weeks) and compounded by the lack of visible landmarks in the cylindrical body of the sea anemone. Its body plan is not radially symmetrical, but displays morphological bilateral

symmetry, defining a secondary body axis (the directive axis). However, this bilateral symmetry is not outwardly visible in live polyps, as it is apparent in transverse tissue sections (i.e. bilaterally symmetrical siphonoglyph groove in the pharynx, and mesenteric retractor muscle placement in all eight mesenteries, also reviewed in[172]), and thus could not be used for orientation.

We therefore used a partial photoconversion strategy to identify the same nerve net region for repeated imaging over 9 days post-photoconversion to track body wall nerve net stability. After confocal-based photoconversion of a rectangular body wall region of a 6 week-old *FoxL2>KikGR*, animals were re-imaged at 2, 4, 7 and 9 days. The photoconverted KikGR protein could be detected after at least 9 days *in vivo* with the rectangular photoconverted region clearly detectable, showing no gross tissue reorganization in that timeframe. Individual neurons could be tracked from their distinctive localization, neurite morphology and branching, providing insights into the organization and stability of the nerve net over time.

Thus, identical neurons could be clearly identified throughout the 9-day period, revealing some stability of nerve net cellular composition and organization over time (Fig. 2.3). These neurons' cell bodies are embedded in both the epidermal and sub-epidermal layer, and located both in the neural tracts parallel to the parietal muscle as well as outside of it. This is further discussed in chapter 4. In addition to the identification of *FoxL2>KikGR*$^{+}$ neurons stably present in the nerve net, *in vivo* imaging at single-neuron resolution also revealed a clearly stereotypical neurite orientation branching across the nervous system. Namely, there are three classes of anatomically-delineated neurite branching patterns, each associated to the localization of the neuron soma in the body wall. Neurons in the epidermal monolayer, with a sensory cell morphology, are mostly tripolar, with variable neurites patterns. Then, neurons whose cell body are embedded in the subepidermal layer outside of the neurite tracts display a clear circumferential orientation, localized above the neurite plexuses. Going deeper

towards the mesenteries, neurons whose cell bodies are located within neurite plexuses have a mostly longitudinal oral-aboral orientation within these neural plexuses.

This global stereotypical neurite orientation in the body wall nerve net mirrors what has been described from a single neuronal subtype (*LWa>mCherry*-expressing neurons in[178]) and is thus seen across all neurons labeled by our three drivers (*FoxL2>KikGR*, *Elav>KikGR*, and *SoxC>KikGR*), which, collectively, as shown in chapter 3, encompass the currently described molecular diversity of neuronal subtypes. See chapter 4 section for further discussion.

In addition to the identical *FoxL2>KikGR*$^{+}$ neurons present in the body wall nerve net, "new" *FoxL2>KikGR*$^{+}$ neurons are detectable at 7 days post-photoconversion across anatomical layers, i.e. cells that were not present in the same location or with the same morphology at day 0. Whether these cells are neurons migrating from other areas or newly differentiated since day 0 cannot be determined with this partial photoconversion experimental setup (see chapter 3). Altogether, these experiments establish a proof of concept for single-cell detection of neurons in live intact polyps. The repeated imaging further reveals both nerve-net stability and its dynamic changes over a week (see discussion in chapter 4).

However, these underscored several limitations of the confocal-based partial photoconversion strategy: (1) confocal-based piecewise photoconversion is not efficient or fast enough to completely photoconvert an entire ca. 6 weeks old polyps given its large surface area (several mm in body length); (2) re-imaging the exact same area is not scalable across polyps; (3) while the boundary of the photoconverted region can be visualized, determining which cells are negative for the photoconverted KikGR is challenging. The latter would severely limit the biological interpretability of the photoconversion status and we thus explored other complete whole-organism photoconversion methods. Thus, to track cellular lineages, we aimed at photoconverting the entire polyp and tracking the day0-labeled cells and detecting newly-expressing cells over several days in an unperturbed context.

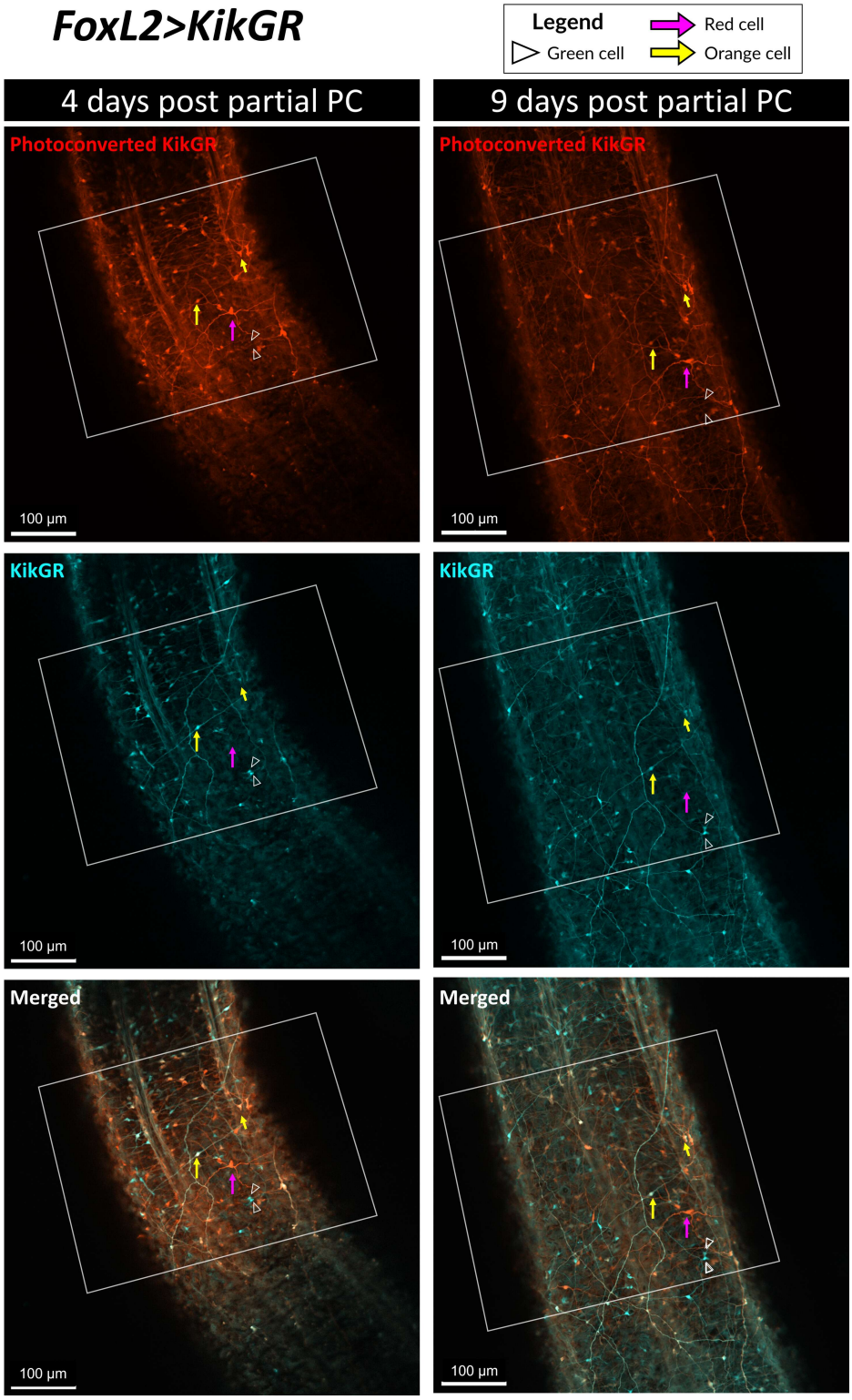

Figure 2.3: Persistent nerve net organization highlighted by *in situ* imaging

Partial photoconversion (approximate boundary in white) of ca. 6 week-old *FoxL2>KikGR* polyps with stable neurites branching at 4 and 9 days post-photoconversion with the same neuron bodies identified.

2.4 A photoconvertible reporter method for *in vivo* lineage tracing

2.4.1 Whole-organism photoconversion of *Nematostella* polyps

For full polyp photoconversion, given the size of the animals (~2.5mm in length at 6 weeks), total and efficient confocal-based whole-organism photoconversion did not seem to be technically achievable using a laser-scanning confocal-based approach. And the large dynamic range of fluorescent protein expression across drivers (see above) hindered efficient photoconversion of both lowly-expressing and highly-expressing cells in the same animal, with photo-bleaching occurring when the number of photoconversion cycles was increased to photoconvert the highly-expressing cells.

We tested alternative approaches, and used an epifluorescence microscope (Zeiss AxioImager M2) with a long (6-8s) exposure time at 20X magnification with filter set 49 (Zeiss), used to detect DAPI signal. From a full spectrum LED light source (Xcite 120LED, Excelitas), this provided a strong UV/near-UV excitation centered around 360 nm, on a large area that could be manually tiled to efficiently photoconvert an entire animal held in the custom imaging chamber and immobilized using standard $MgCl_2$-based method. This protocol achieved efficient and complete photoconversion for all *KikGR* reporter lines in a short timeframe (<5min per animal), with no detectable photobleaching. See Fig. 2.4 for an example of a *SoxC>KikGR* animal before and after photoconversion, and Fig. 3.31 below for a flow cytometer characterization example of unphotoconverted and photoconverted animals). The GFP-like pharyngeal signal is autofluorescence and does not represent KikGR signal (see Fig. 2.5). Complete photoconversion (i.e. no unconverted green KikGR fluorophore) was more challenging in tentacular *FoxL2>KikGR*$^+$ spirocytes (a non-piercing cnidocyte subtype), possibly due to their much higher fluorophore expression level than any other cell type. Animals exhibited full recovery after the procedure, with no morphological defect observed >100 photoconverted animals.

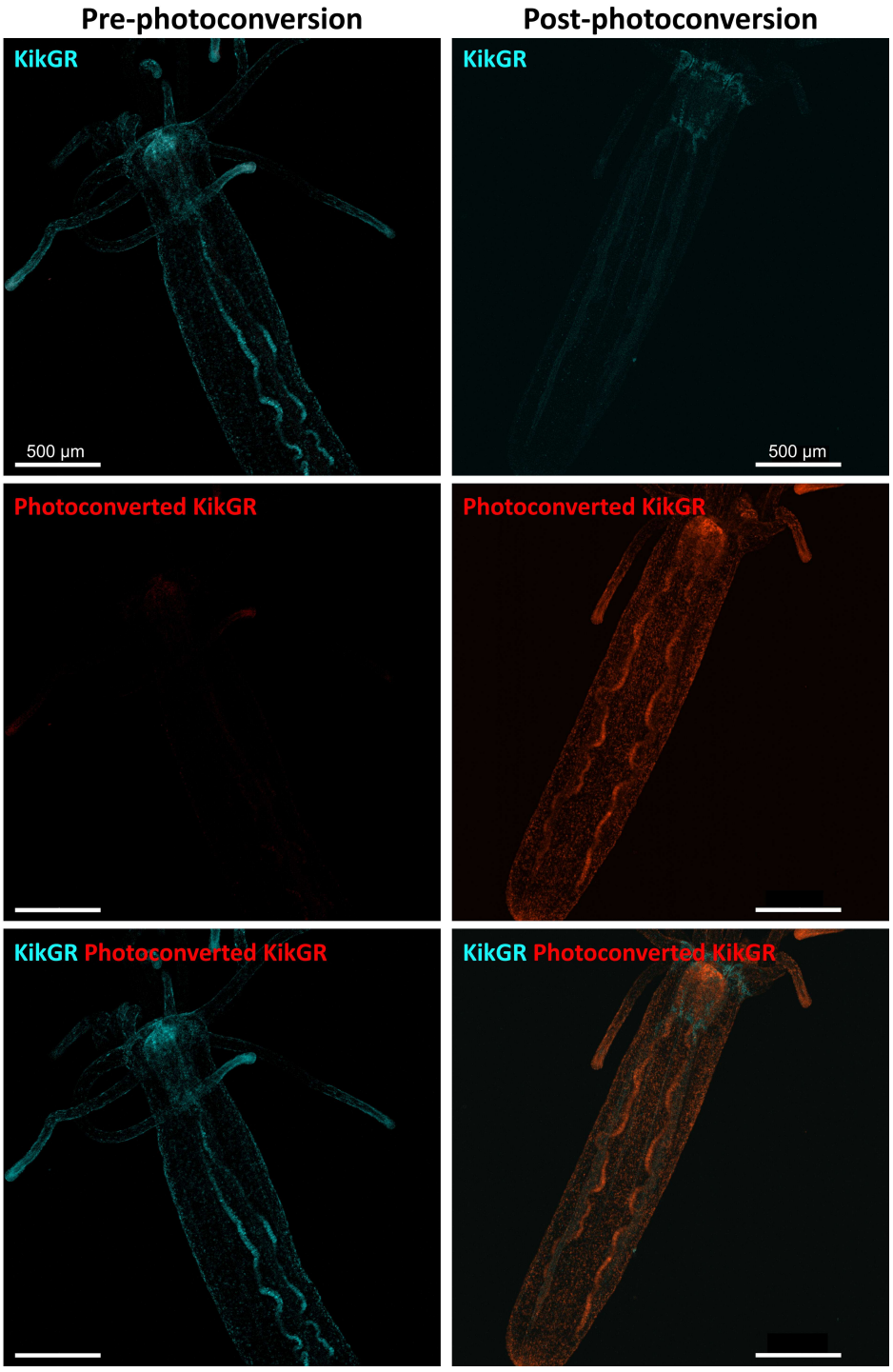

Figure 2.4: Efficient whole-organism photoconversion in a *SoxC>KikGR* adult

Whole animal confocal imaging (5X) of a 2.5mm long sexually immature *SoxC>KikGR* adult (ca. 6 weeks old) pre- and post-KikGR epifluorescence-mediated photoconversion. Note the presence of pharyngeal GFP-like cellular autofluorescence on the oral end at the base of the pharynx and non-cellular autofluorescence at the base of each tentacle.

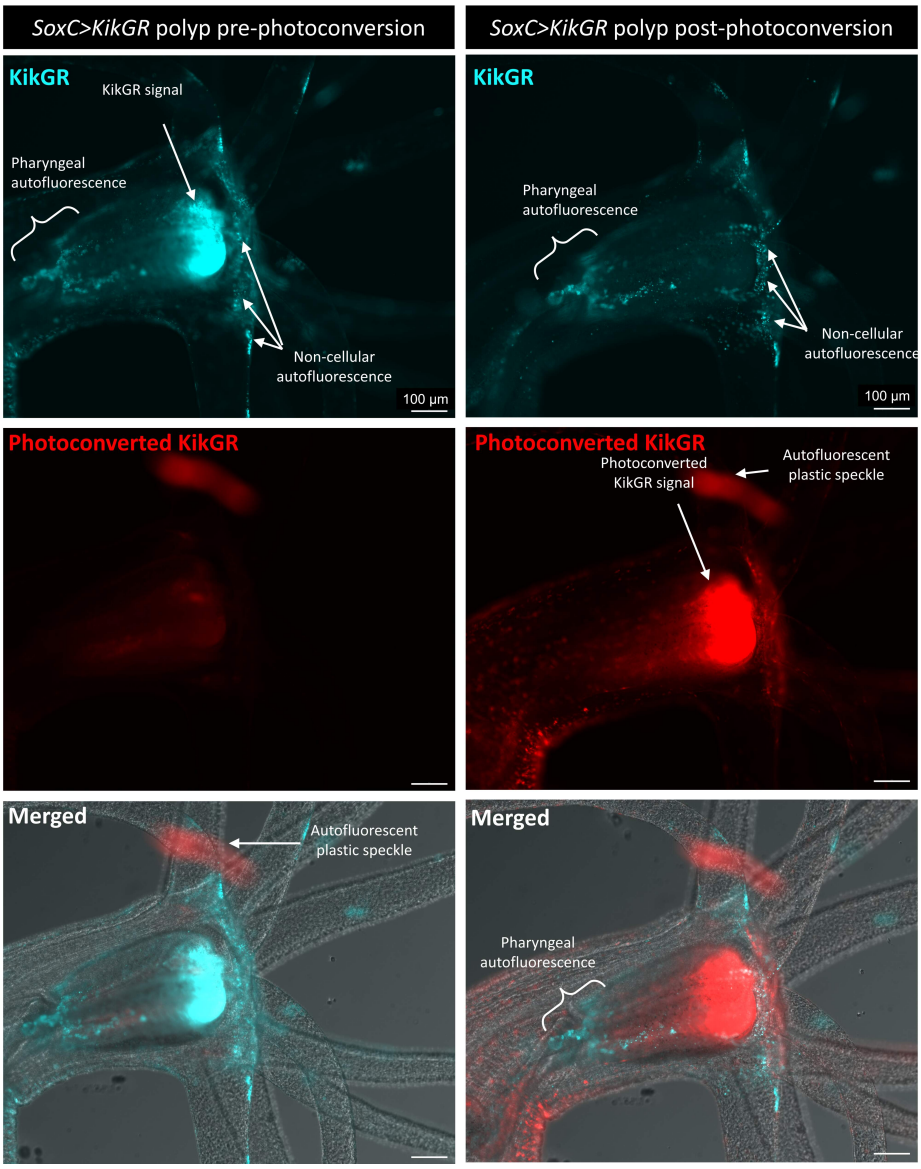

Figure 2.5: Pharyngeal region pre and post-photoconversion

Pharyngeal region pre- and post-KikGR photoconversion of a *SoxC>KikGR* adult (ca. 6 weeks old). Internal GFP-like autofluorescence at the base of the pharynx, and the noncellular epidermal GFP-like autofluorescent signal at the base of each tentacle. 20X epifluorescence image at the depth of the middle of the pharynx, so individual cells in the body wall and tentacle areas are out of focus here.

2.4.2 Tracking *SoxC>KikGR*$^{+}$ cells during oral regeneration

As a proof of concept for tracking labeled cells and newly-expressing cells after whole-animal photoconversion, we photoconverted *SoxC>KikGR*$^{+}$ polyps, amputated their oral end and assessed the presence of previously labeled cells (with the photoconverted KikGR signal). At 4 days post-amputation and photoconversion, cells with the photoconverted KikGR and lacking new KikGR protein (red cells hereafter), cells with only the newly-expressed KikGR signal (green cells) and cells with both (orange cells) are detectable throughout the tentacles (Fig. 2.6).

Morphologically mature red cnidocytes (with their intracellular capsule and neurites) are present throughout the tentacles including at their tip, across all animals. These red late cnidocytes come from cells that contained the photoconverted KikGR protein at day 0 and are not expressing cells in the 4-6 days afterwards. Either those day 0 *SoxC>KikGR*$^{+}$ red cells were already late cnidocytes, which could indicate that mature cnidocytes from other regions are re-used in the formation of new tentacle tissue and/or that there is widespread cell migration; or transiently expressing *SoxC>KikGR*$^{+}$ cells that were not yet cnidocytes give rise to these. This is further discussed in chapter 4. Anecdotally, we observed very few green developing cnidocytes in the aboral region of oral-amputated polyps at 4 days post-amputation, while these were abundant in the non-amputated condition.

Altogether, the whole-photoconversion is efficient, and can be used to track photoconverted KikGR^{+} cells over at least 7 days *in vivo*, discriminating between neurons, developing cnidocytes, mature cnidocytes, secretory cells, and amoeboid-like cells across the body column and tentacles. This identified red, orange and green cells, but it was qualitative. The molecular identity of *KikGR*$^{+}$ cells is described in chapter 3, so we focus here on their relative localization within the body wall nerve net. As the nerve net shows stereotypical organization (section 2.3.3), identifying the relative position of the green/orange/red KikGR^{+} cells is important to understand where new cells are added relative to the existing ones, and

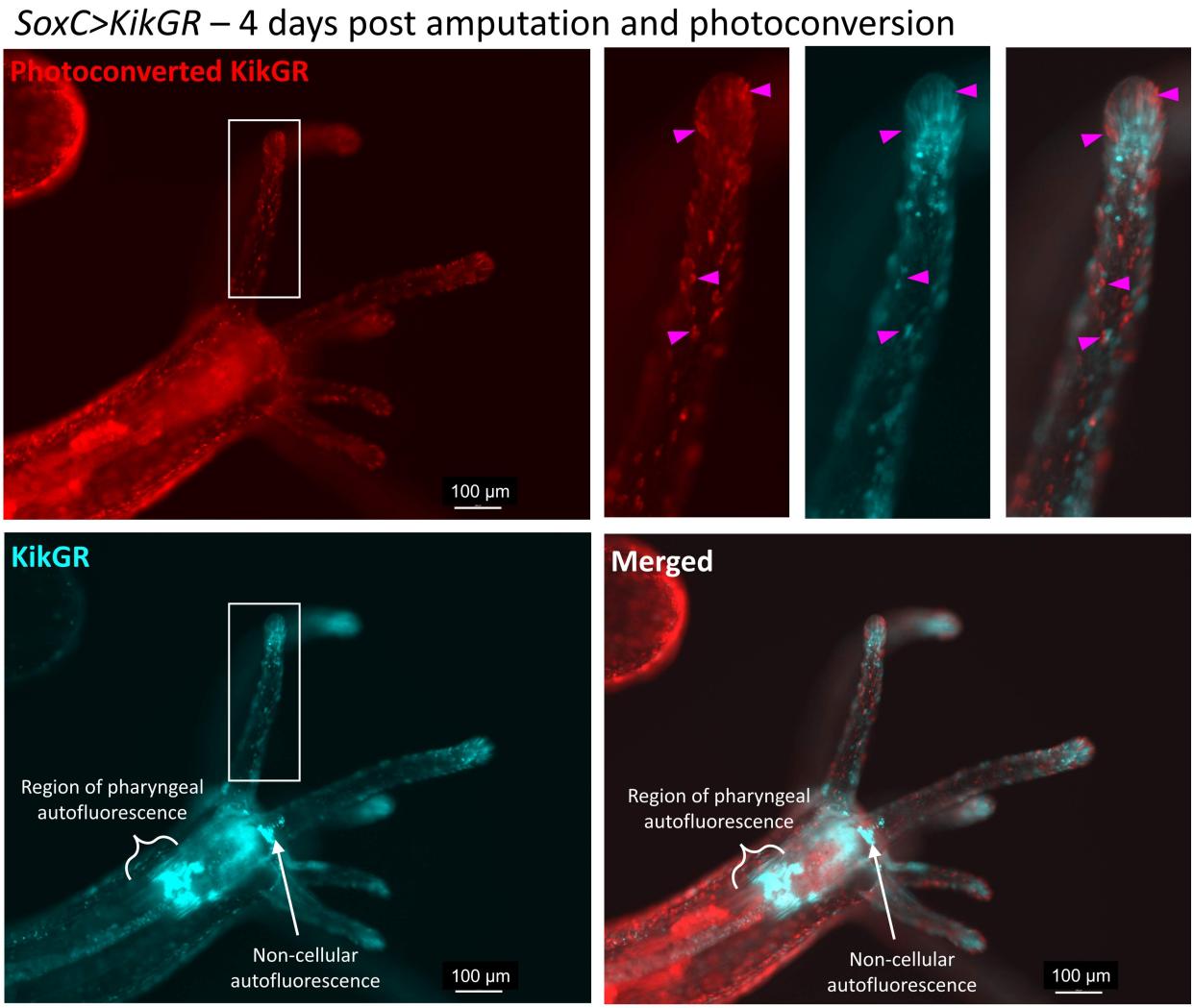

Figure 2.6: Tracking *SoxC>KikGR*$^{+}$ cells during oral regeneration

Oral side at 4 days post-oral amputation and photoconversion (PC) of a ca. 6-week-old *SoxC>KikGR* polyp (20X epifluorescent microscope). Individual red cells and green cells are clearly detectable, with red late cnidocytes detected throughout the tentacles in all animals with their intracellular darker capsule (magenta arrowheads). The bright GFP-like (in cyan) signal at the base each tentacle in the epidermal layer is not KikGR signal but non-cellular GFP-fluorescent particles (white arrow) and the GFP-like internal signal at the base of the pharynx is autofluorescence (as in Fig. 2.4 and Fig. 2.5). Oral part of the pharynx has cellular KikGR signal as well.

specifically whether neurons are only added to a specific subset of the nerve net or broadly distributed spatially.

2.5 Spatial localization of newly-generated cells

2.5.1 Localization relative to anatomical cell layer

To quantify the presence of green/orange/red $KikGR^{+}$ cells in the body wall, and analyze where these cells are localized relative to anatomical landmarks, we used confocal imaging, sampling the body wall area across epidermal monolayer and subepidermal body wall areas, combined with user-trained cell segmentation machine learning methods and morphological landmark registration and analysis. This highlighted the presence of green/orange/red KikGR+ cells across all layers and anatomical locations, and on the oral and aboral regions of the body wall, further supporting a model of active distributed nerve net scaling (see chapter 3).

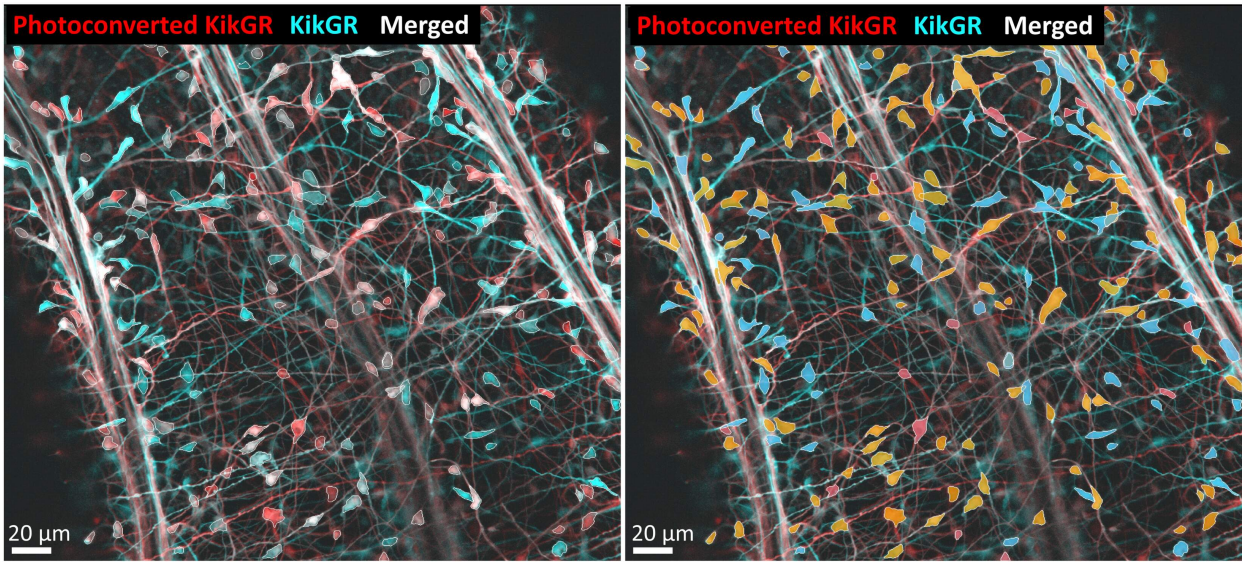

Figure 2.7: *Elav>KikGR*$^{+}$ cell segmentation and green/orange/red cell annotation
Example of a cell segmentation output using the trained Cellpose model with individual neuron bodies outlines on a single imaging slice of a *Elav>KikGR*$^{+}$ polyp at 4 days post-photoconversion (**left**). Segmented cell bodies are annotated in blue (green cells), orange or red (**right**). Scale bar is 20 μm. Note that the green cells are localized throughout the neurite plexuses and inter-mesenteric area.

To quantify KikGR^{+} cells in the body wall nerve net, single-neuron resolution confocal z-stacks were acquired with the same parameters for each driver at 2, 4 and 7 days post-

photoconversion in polyps. Cell bodies, excluding neurites, were manually annotated for Cellpose 2.0[219,220] (a user-trainable machine learning segmentation tool) model training. The resulting model was used for automated cell segmentation for all images, separately for the photoconverted KikGR signal and the newly-expressed KikGR signal channels. Detected cell bodies from each channel were merged across the z-stack volume (see single slice example in Fig. 2.7)and used to quantify the number and relative distribution of green, orange and red cells in each image as shown in chapter 3.

The distribution of detected green/orange/red *KikGR*$^{+}$ cells is different between drivers, reflecting their distinct expression dynamics. In brief, *SoxC>KikGR+* cells are in majority either green or red, with less orange cells detected, consistent with a transient expression of the driver associated with early differentiation. In contrast, *FoxL2>KikGR*$^{+}$ cells are a majority of orange or red cells, consistent with a continuous expression in a source population and a progressive downregulation in mature cells, except for cnidocytes, where it is secondarily upregulated at the late cnidocyte stage. Thus, we focus here on the distribution of green/orange/red *Elav>KikGR*$^{+}$ and *FoxL2>KikGR*$^{+}$ cells. Red *KikGR*$^{+}$ cells can be detected throughout at least 7 days post-photoconversion *in vivo*, with the green cell fraction increasing with increasing time post-photoconversion, as predicted. Global cell quantifications across drivers, regions and days post-photoconversion are presented in chapter 3.

We will focus here on the localization of these cells relative to anatomical landmarks. As neurons are localized in specific regions in the body wall, we further analyzed the position of the detected *KikGR*$^{+}$ cells relative to both cell layer compartment (epidermal monolayer vs sub-epidermal) and to its distance to the nearest mesentery, where neurite plexuses are located. By tracking newly-generated neurons using a whole-animal photoconversion-based approach, we can combine the cell color information with the spatial location of each cell to investigating where these cells are localized, i.e. restricted to a specific nerve net area or broadly distributed. This would provide insight into the neurogenic process by indicating

whether newly-generated neurons are preferentially localized in the tracts, in-between these, or in the epidermal layer.

For spatial analysis, morphological landmarks were registered for all images, delineating rectangular regions between two double neurite tracts (i.e. one spatial analysis region is between two dashed lines in Fig. 2.8A, with one to two such regions by imaging stack). For registration, anonymized greyscale .tiff z-stacks were manually annotated in ImageJ at the center of each mesenteric tract between the two neurite bundles at the depth where the above neurites perpendicularly cross the nerve tracts across all images (indicated by the pink X in Fig. 2.8). This morphological landmarks were used to define a standard rectangular region of interest between two mesenteries, encompassing the same anatomical region, and to define the epidermis vs sub-epidermal boundary (the dark vs. light gray layers in Fig. 2.8A). From these registered coordinates, a piecewise affine transformation (scikit-image library[221]) for each annotated region was used to normalize each layer into a standardized rectangle area of interest. As the z-stack layers are not fully parallel to the internal tissue layers, depth was interpolated using the registered coordinates with a reference surface computed using multilevel B-splines with the mba.point function (with m = 2, n=1, h = 3) (MBA package in R). The depth distribution of cell bodies centroids within the standardized area was bimodal, with a sharp peak for the epidermal monolayer, used to classify cell centroids as part of the epidermis or sub-epidermal compartment.

The spatial distribution of all KikGR^{+} cell bodies (over 20,000 segmented cells bodies in total) across the registered areas fully recapitulates the two-layers anatomy Fig. 2.8B of the body wall. In the epidermis monolayer, cell density is lowest where the mesenteries originate (black dotted line, distance to closest mesentery defined as 0) and becomes uniform with increasing distance. This mirrors the epidermis monolayer cell density, which is uniform, except for a dip at the location of the mesentery. In the sub-epidermal layer, cell density is minimal at distance = 0 (where parietal muscles are inserted), then there is a sharp peak at

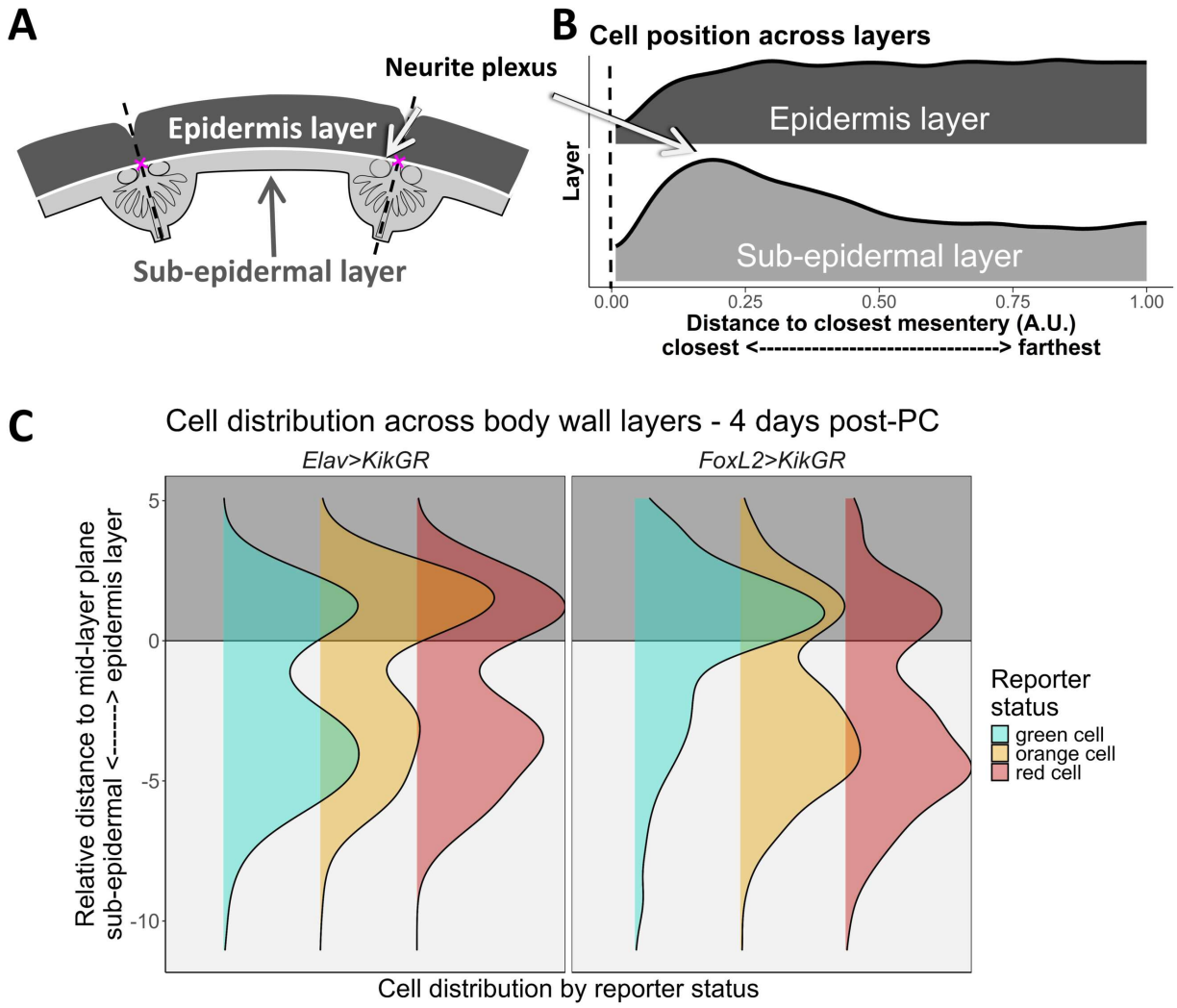

Figure 2.8: Cell segmentation and landmark registration for spatial analysis

(**A**) The body wall has an epidermal monolayer and a sub-epidermal compartment where the neurite tracts are located. The pink X locates the registered anatomical landmarks. (**B**) Observed distribution of cell bodies centroids (21,085 segmented cells with spatial information) across the epidermis and sub-epidermal compartment with their distance to the closest mesenteric tract (the black dotted line shown in **A**). (**C**) Bimodal distribution of cell bodies highlighting the two cellular compartments, with green, orange and red *Elav>KikGR*$^+$ and *FoxL2>KikGR*$^+$ cells present (4 days post-photoconversion, oral body wall).

the dense neurite plexus region (distance of about 0.2 A.U.), and a decreasing density farther from the mesentery tract origin, mirroring, tissue organization. This profile is similar across drivers, highlighting the distributed presence of KikGR$^+$ cell bodies across anatomical layers,

matching overall cell density, across both the epidermal and sub-epidermal compartment in the body wall, for all drivers and days post-photoconversion, as seen in individual images.

Each detected KikGR$^+$ cell within the standardized volume is annotated with its anatomical cell layer (epidermal or sub-epidermal), and its distance to the nearest mesenteric tract (the dotted black line in Fig. 2.8. As neurite organization differs systematically across nerve-net subsets according to their anatomical location, this spatial annotation, combined to the reporter status information (green/orange/red cell) provides insight into where new neurons are incorporated. Accordingly, green, orange and red *Elav>KikGR*$^+$ cells are detected across both cell layers (shown at 4 days-post-photoconversion in the oral region in Fig. 2.8C, with similar results at all days post-photoconversion and regions assayed), with higher number of cells detected in the sub-epidermal layer, consistent with the increased density of neuronal cell bodies there. Green *Elav>KikGR*$^+$ encompass newly-generated neurons, with a mostly neuronal morphology, and have cell bodies within the tracts, in the inter-mesenteric sub-epidermal layer, and on the epidermal layer. This supports a model where new neurons are incorporated throughout the nerve net.

In the epidermal layer, as the KIKGR protein is cytosolic the epithelial epidermis cells express the *FoxL2>KikGR* driver at a low level, slowly accumulating the KIKGR protein, and are uniformly *KikGR*$^+$ with individual epithelial cells not segmented and so these cells are not counted or present in the spatial analysis. Interestingly, the *FoxL2>KikGR*$^+$ cells in the epidermal layer are mostly green or red, consistent with the large fraction of mature cnidocytes detected (see chapter 3); while the cells localized in the sub-epidermal layer are orange or red, consistent with a majority of neurons. Overall, green, orange and red *FoxL2>KikGR*$^+$ cells are detected across both cellular compartments, with a preferential localization of green cells in the epidermal layer, consistent with the presence of mature (green) cnidocytes in the epidermal monolayer (see also Fig. 3.6).

2.5.2 Preferential cell distribution relative to mesenteric tracts

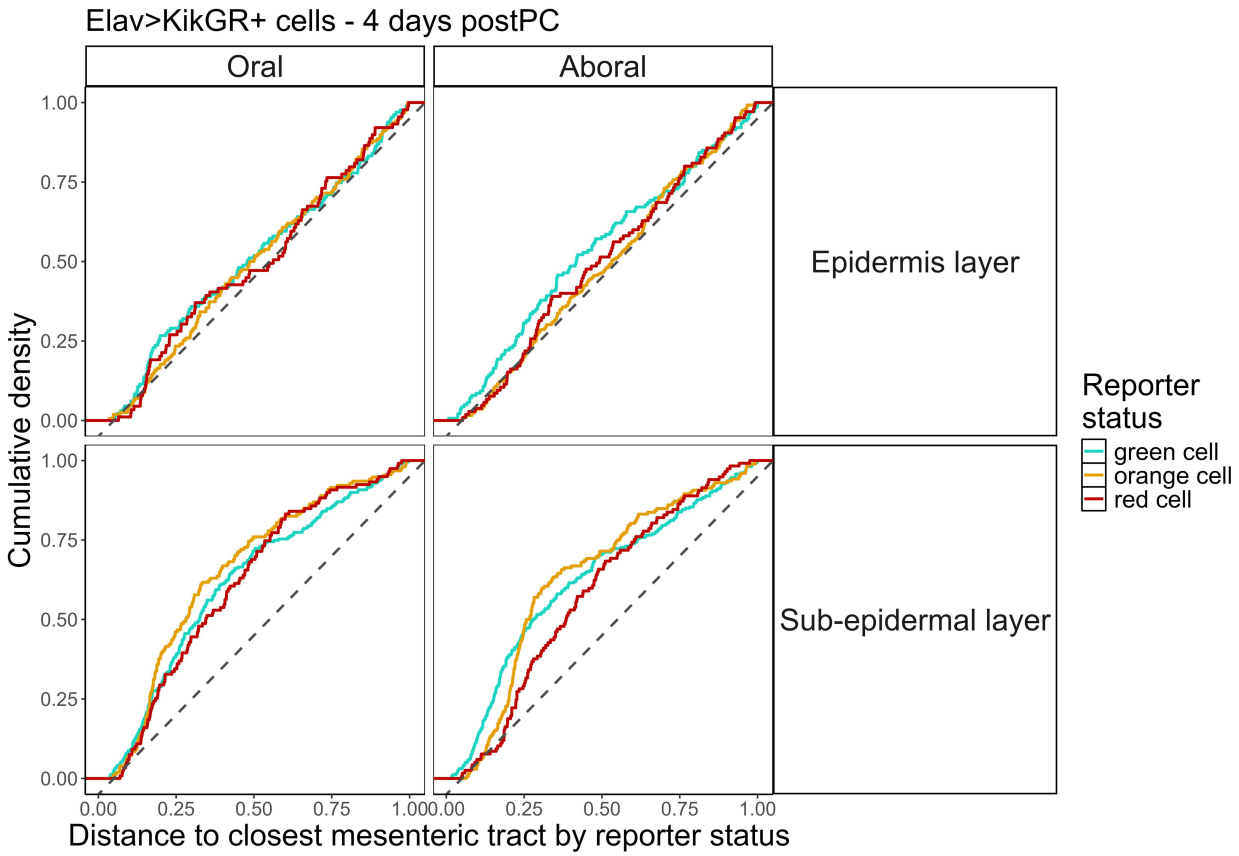

Figure 2.9: *Elav>KikGR*$^+$ cells spatial distribution by reporter status

Cumulative density distribution of the position of *Elav>KikGR*$^+$ cells relative to the closest mesenteric tract across the epidermal and sub-epidermal layers at 4 days-post-photoconversion, in oral and aboral regions. Grey dotted line indicates a uniform distribution after the first dip (slope=1, intercept is -0.05). Green and orange *Elav>KikGR*$^+$ cells show a slight shift towards the mesenteries sub-epidermal layer but not in the epidermal compartment. Position 0 corresponds to the arrowhead in Fig. 2.8A/B.

We then characterized the spatial distribution of green, orange and red KikGR$^+$ cells relative to the mesentery insertion (black dashed line in Fig. 2.8), as this could reveal whether newly generated neurons are preferentially integrated into specific anatomical locations. In particular, we assessed whether green KikGR$^+$ cells were enriched near the neurite plexuses, which would indicate a biased incorporation of new neurons into a restricted subset of the nerve net rather than uniform addition across the tissue.

Distribution of *Elav>KikGR*$^+$ cells is mostly uniform in the upper layer in the oral and aboral regions as shown at 4 days post-photoconversion in Fig. 2.9. This suggests that there is no specific localized integration of the newly-generated neurons but that they are distributed throughout. This contrasts with the subepidermal layer, where the overall cell distribution is shifted toward the proximal region where the neurite plexus is located, consistent with a higher neuronal density in this area. Red *Elav>KikGR*$^+$ cells additionally showed a trend toward displacement away from the mesenteric tract.

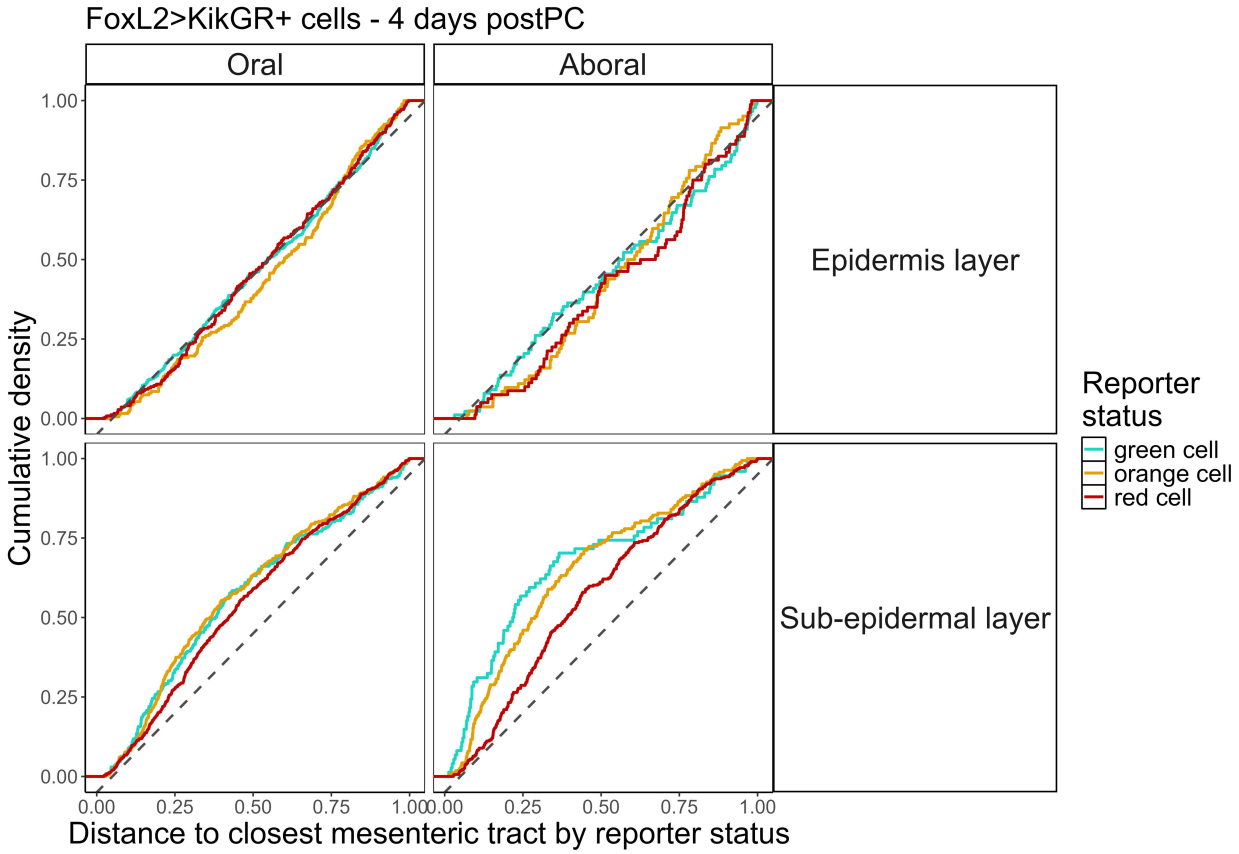

Figure 2.10: *FoxL2>KikGR*$^+$ cells spatial distribution by reporter status

Cumulative density distribution of the position of *FoxL2>KikGR*$^+$ cells relative to the closest mesenteric tract across the epidermal and sub-epidermal layers at 4 days-post-photoconversion, in oral and aboral regions. Grey dotted line indicates a uniform distribution (slope=1, intercept is -0.05). Green and orange *FoxL2>KikGR*$^+$ cells show a slight shift towards the mesenteries sub-epidermal layer but not in the epidermal compartment. Position 0 corresponds to the arrowhead in Fig. 2.8A/B.

This cell distribution is similar for *FoxL2>KikGR*$^+$ cells, with a mostly uniform distri-

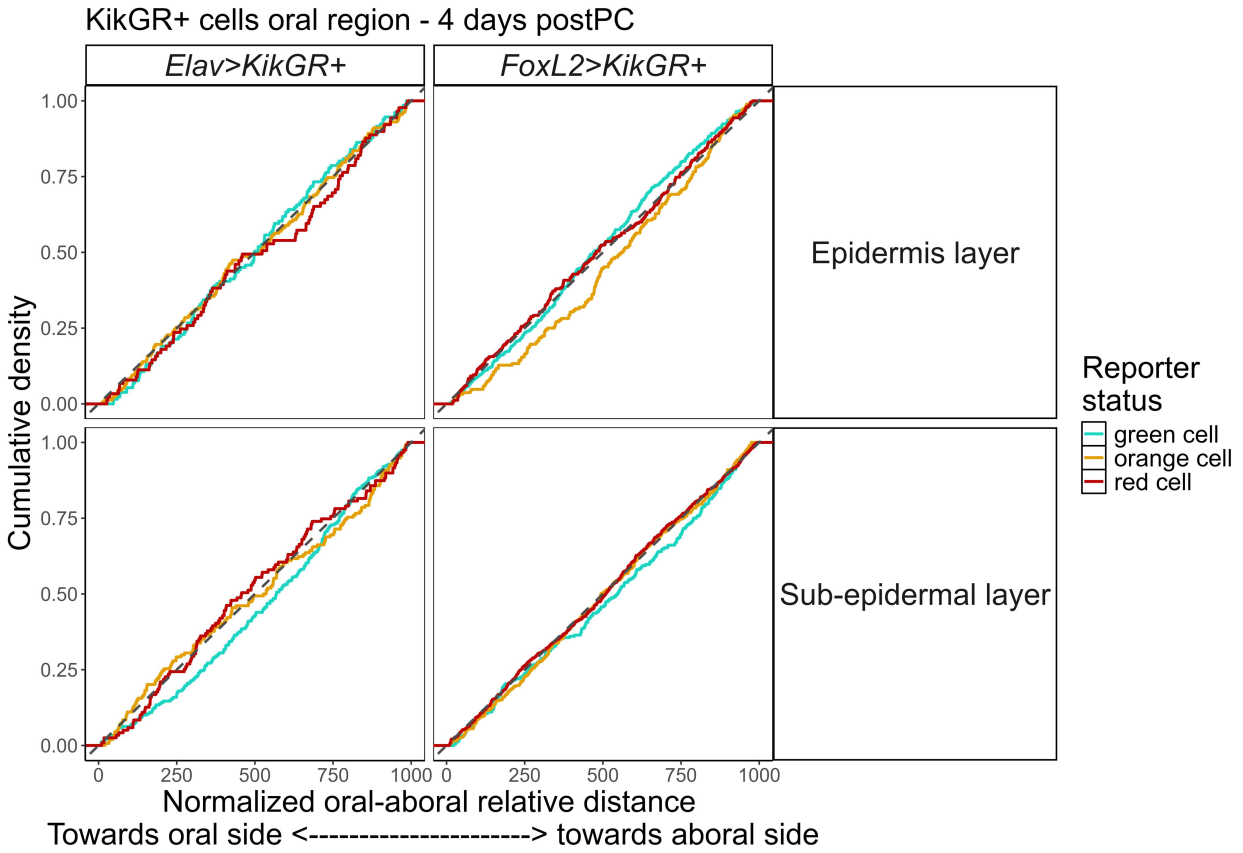

Figure 2.11: Spatial distribution along the primary body axis by reporter status

Cumulative density distribution of the position of *Elav>KikGR*$^+$ and *FoxL2>KikGR*$^+$ cells relative to the primary body axis localization (i.e. 0 is closest to the oral side) across the epidermal and sub-epidermal layers at 4 days-post-photoconversion, in the oral region. Similar distribution is observed in the aboral region. Grey dotted line indicates a uniform distribution.

bution in the epidermal layer, and an increased cell density closer to the mesenteric tract in the sub-epidermal compartment. There may be a slight shift in the cell distribution with green and orange cells located slightly more towards the mesenteric tracts.

We also assessed whether this distribution was shifted in the oral-aboral axis, as a local shift might be indicative of a globally oriented generative process in the animal. There was no obvious oral-aboral shift in the distribution of green, orange and red cells across the *FoxL2>KikGR* and *Elav>KikGR* drivers, shown for oral regions at 4 days post-photoconversion in Fig. 2.11, with the same distribution for aboral regions. This is consistent

with a globally distributed cell renewal process, and indicates it is not preferentially oriented.

Overall, green, orange and red cells are present across anatomical compartments, distributed according to the overall cell density, with a possible slight shift in the distribution of green and orange cells towards the mesenteric tracts in the sub-epidermal layer (Figs. 2.9 and 2.10). We further discuss possible scenarios of neuron differentiation supported by these experimental data in light of chapter 3 in chapter 4.

2.6 Single-Cell Transcriptomic Analysis in the Regenerating Cnidarian *Nematostella vectensis*

This protocol (section 2.6) is reproduced from[218], part of a thematic Methods in Molecular Biology collection on Whole Body Regeneration, edited by Brigitte Galliot and Simon Blanchoud:

Plessier, F., Schmutz, S., Novault, S., & Marlow, H. (2022). *Single-Cell Transcriptomic Analysis in the Regenerating Cnidarian Nematostella vectensis* **Methods in Molecular Biology**, 2450, 565–581. doi:10.1007/978-1-0716-2172-1_30.

Authors contribution: Co-authors Sandrine Schmutz (S.S.) and Sophie Novault (S.N.) (research engineers at the Flow Cytometry Platform of Institut Pasteur, Paris, France) provided flow cytometry training, and the initial gating strategy, developed the colorimetric bead-based single cell sorting setup validation assay, described it in note 17, provided note 16 and approved the manuscript. H.M conceptualized and supervised the project. H.M and F.P. developed the cell dissociation protocol and designed experiments. F.P. performed experiments and generated figures. F.P. and H.M. wrote the manuscript.

The example data presented in the protocol is sc RNA-Seq data from control and regenerating *SoxC>KikGR*$^{+}$ polyps, mapped to the[160] scaffold-level *Nematostella vectensis* genome, and a combined gene bin annotation (as used in[8]), which predate the chromosomal-level assembly and gene models used in chapter 3.

2.6.1 Abstract

Cnidarians have historically served as excellent laboratory models for regenerative development given their capacity to regrow large portions of the adult organism. This capacity is notably absent or poorly developed in the powerful genetic laboratory models *Drosophila*, *C. elegans*, and mouse. Increasingly, development of genetic and genomic resources and the application of next-generation sequencing-based techniques in cnidarian systems has further expanded the potential of cnidarian regenerative models. Here, we present a workflow for the characterization of the regenerative response in the sea anemone *Nematostella vectensis* utilizing fluorescence-activated cell sorting and a plate-based single-cell RNA-sequencing pipeline. This approach can characterize the transcriptional response during regeneration in distinct populations of cells, thus providing a quantitative view of a whole organism process at cellular resolution.

2.6.2 Introduction

The ability to regenerate is widespread among marine invertebrates. Understanding the molecular and developmental basis of regeneration in a diversity of taxa has the potential to provide insight into shared principles of the regenerative response. Cnidarians (Hydra, jellyfish, sea anemones, corals) belong to one of the earliest-branching metazoan phyla and are well known for their extensive regenerative abilities[149,222,223].

The sea anemone *Nematostella vectensis* is an experimentally tractable Cnidarian species for studies into mechanisms of regeneration. Significant advances in genetic and genomic resources, ease of culture, and conservation of gene content with Bilaterian taxa contribute to the utility of this sea anemone as a developmental and regenerative model system[119,159,224]. An experimentally bisected *Nematostella vectensis* polyp can regenerate two whole animals in about a week[222]. How the sea anemone regenerates lost tissue, maintains axial coordinates, and generates the appropriate number and distribution of constituent cell types during

regeneration or asexual reproduction is not well understood.

Studies focusing on the cellular dynamics following oral pole regeneration in juvenile polyps have highlighted the involvement of cellular proliferation and apoptosis[222,225]. Whole-animal transcriptomic time-courses have been employed to uncover genes modulated during the regeneration process[226–228]. However, the behaviors of individual cell populations during regeneration have not been molecularly or functionally characterized in *Nematostella vectensis* thus far[229].

In order to better characterize the process of regeneration, additional tools that allow for the characterization of behaviors of single cells are needed. As regeneration is dynamic, gaining an understanding of the molecular events happening in individual cells throughout this process is critical.

Over the past few years, newly developed single-cell RNA-Sequencing (scRNA-Seq) techniques have enabled investigators to sample the transcriptome of single cells from complex tissues and whole organisms, to assess their developmental trajectories and reconstruct gene regulatory networks governing cell identity and function[230]. Different single-cell sequencing technologies have been developed ranging from plate-based, droplet or microfluidic platforms, each with distinct advantages and limitations[231]. Recently, single-cell RNA-Seq of whole *Nematostella vectensis* polyps and planula larvae via a plate-based approach, MARS-seq, has enabled the generation of a molecular atlas of the cell types present in the sea anemone[8]. In combination with targeted genetic labeling and trajectory inference methods, scRNA-Seq data from regenerating animals could enable the reconstruction of key cellular trajectories, as well as the relative contributions of different cell populations to the newly regenerated tissues.

From cell suspension, most scRNA-seq protocols rely on a similar workflow, where single cells are physically isolated and lysed, in wells or droplets, captured mRNA molecules are reverse transcribed and amplified, and multiplexed dsDNA libraries are generated and se-

quenced[231]. Complete, high-viability dissociation of starting tissue is a critical pre-requisite of scRNA-seq. Single cells can then be isolated through Fluorescence-activated cell sort-ing (FACS), which enables a single cell to be separated from a cell suspension with a high degree of purity. Following single-cell isolation, mature mRNAs are often captured through their polyadenylated tails using barcoded oligos bearing polyT stretches, then reverse transcribed into cDNAs before being amplified via PCR or *in vitro* transcription[232]. During capture or in subsequent steps, cell- and/or pool-specific barcodes are added, to enable attribution of the sequenced reads to a specific cell and sample of origin. To limit the impact of amplification artifacts, unique molecular identifiers (UMIs), short stretches of random DNA bases, are attached to the captured mRNA molecules[232,233]. Duplicate reads from the same mRNA molecule can be merged by UMI. For data processing and analysis, several quality control and clustering analysis packages have been developed[234–236].

In this protocol, we focus on a single-cell RNA-sequencing experiment in a regenerating *Nematostella vectensis* polyp including the dissociation of samples of interest into live single-cell suspensions, the flow cytometer parameters and gating strategy to sort live single cells into separate wells of a 384-well plate, as well as an overview of quality control steps after library sequencing, read demultiplexing and mapping, QCs, and general notes and considerations specific to single-cell RNA-Seq *Nematostella vectensis* data. Briefly, polyps along a regeneration time-course following oral pole amputation are dissociated into single-cell suspensions at timepoints of interest, then live single cells are sorted on a flow cytometer into single wells of a 384-well plate (Fig. 2.12). Following cell lysis and mRNA capture with barcoded oligonucleotides in individual wells, multiplexed cDNA libraries are generated using the MARS-Seq single-cell RNA-Seq method[232,237]. The scRNA-seq library preparation method has been adapted from an established workflow and an extensive protocol detailing the experimental steps, and an associated computational pipeline has recently been published[237].

After library sequencing, reads are filtered, demultiplexed, and mapped using the MARS-Seq computational pipeline onto the *Nematostella* genome to generate count tables, listing the number of recovered molecules from each gene in each sequenced cell[160,237] (Fig. 2.12D). The count tables can then be filtered and clustered using various clustering methods for biological analyses[238] (Fig. 2.12d). Here, we generate an example clustering using the MetaCell clustering package, a published pipe-line for the analysis of scRNA-seq data[234].

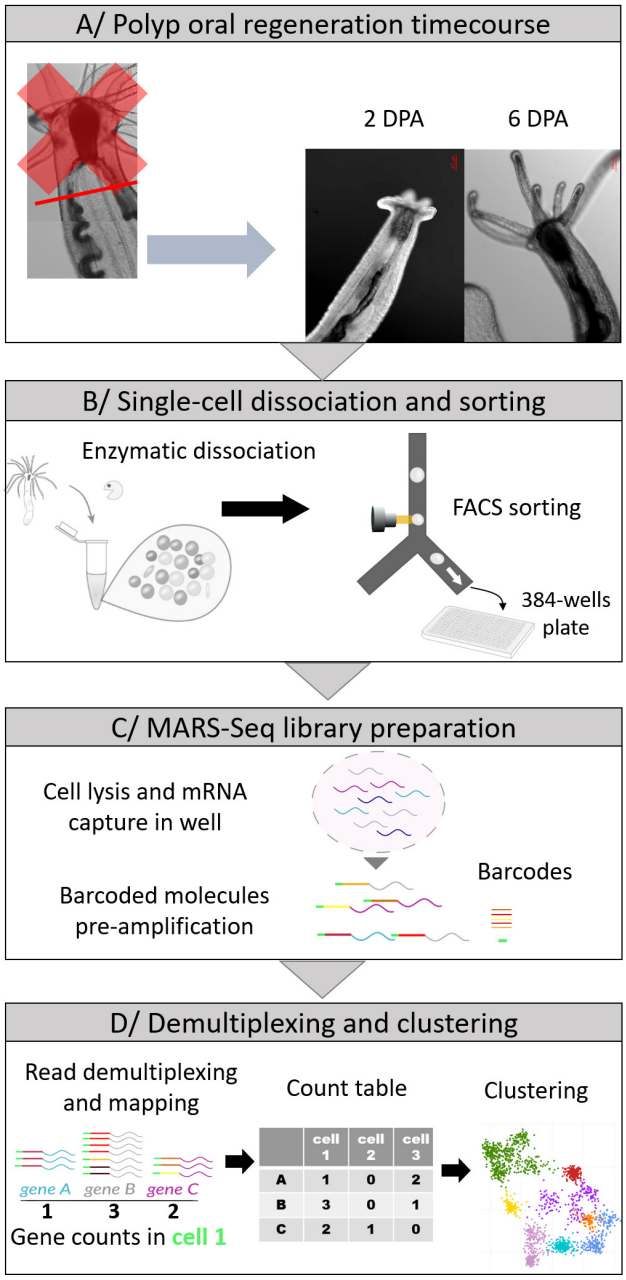

Figure 2.12: Overview of single-cell RNA Sequencing in the regenerating sea anemone *Nematostella vectensis*

(**A**) DIC images of the oral region of uncut control, 2 and 6 day post-amputation (dpa) juvenile *Nematostella vectensis* polyps (5 weeks) at room temperature. Amputation site is indicated with a red line. (**B**) Enzymatic dissociation of sample into a single-cell suspension and single-cell FACS sorting into 384-well plates. (**C**) Simplified overview of the labeling strategy for single-cell-derived mRNA which includes barcoding for each cell in each well using capture oligonucleotides with a polyT-tract, a random unique molecular identifier (UMI) and a cell barcode. (**D**) Multiplexed single-cell RNA-seq short-read libraries are sequenced, reads are filtered, mapped, and demultiplexed to attribute recovered molecules, through their unique UMIs tags, to their cell of origin. Cells are then clustered using unsupervised clustering based on gene expression for downstream analyses.

2.6.3 Materials

All solutions should be prepared using ultrapure water and stored at room temperature (RT) unless otherwise stated.

1. *Nematostella* medium: 12.8 g/L artificial seawater salt in distilled water.
2. 24 h starved 5- to 8-week-old juvenile animals.
3. Six-well plates.
4. Sterile scalpel blades.
5. 3-mL plastic Pasteur pipets.
6. 7% (w/v) magnesium chloride solution in ultrapure water.
7. Low-binding 1.5-mL tubes (e.g., Biozym #710176).
8. Calcium-Magnesium-Free Artificial Seawater (CMFSW) for *Nematostella* dissociation medium: 495 mM NaCl, 9.7 mM KCl, 27.6 mM $NaHCO_3$, 50 mM Tris–HCl pH 8. Filter sterilize.
9. Live cell marker: 1 μg/μL calcein AM resuspended in dimethyl sulfoxide (DMSO). Protect from light and store at -20°C (see Note 1).
10. Dead cell marker: 1 mg/mL propidium iodide ready-to-use solution. Protect from light and store at 4°C (see Note 1).
11. Collagenase solution: 5 mg/mL Liberase^{TM} resuspended in nuclease-free water. Store aliquots at -20°C.
12. Gelatin solution: 1 mg/mL gelatin. Heat to dissolve, filter.

13. Gelatin-coated P200 tips: pipette up and down gelatin solution, let the tips dry overnight, store at RT.

14. Gelatin-coated P1000 tips.

15. Trimmed P200 gelatin-coated tips: remove 1–2 mm of dry coated tip length using a scalpel blade.

16. Dissociation medium: 32.4 mL CMFSW, 17.6 mL ultrapure H_2O, filter at 0.4 μm. Check salinity against sample salinity.

17. Staining medium: 2 μL/mL live cell marker, 3 μL/mL dead cell marker in dissociation medium. Prepare fresh. Protect from light.

18. Fluorescence-Activated Cell Sorter (e.g., BD FACS ARIAIII), with laser/filters combinations appropriate for the dyes used (see Note 2).

19. Flow cytometer sizing beads: 1.34 μm, 3 μm, and 6 μm diameter.

20. Frozen cell capture plates: 384-wells plates filled with 2 μL polyT barcoded capture oligos in lysis solution stored at -80°C (see Note 3).

21. *Nematostella vectensis* genome and gene annotation (see Note 4).

22. MARS-Seq library preparation pipeline (see Note 5).

23. DNA electrophoresis equipment (e.g., TapeStation, Agilent).

24. DNA concentration fluorometer (e.g., Qubit, Thermo Fisher Scientific).

25. MARS-Seq computational pipeline (see Note 6).

26. scRNA-seq data clustering pipeline (see Note 7).

2.6.4 Methods

Cell Dissociation of Regenerated and Control Samples

1. Select 5- to 8-week-old siblings of similar size (see Note 8).
2. Transfer one polyp into the lid of a 5-cm plastic dish filled with
3. 1.5 mL Nematostella medium.
4. Repeat step 2 for the remaining siblings, in separate lids.
5. Wait for 20 min for the polyp to relax, until the tentacles are fully extended.
6. Add drop-by-drop, on the side of the lid, 1.5 mL of 7% $MgCl_2$.
7. Wait for 30s for full immobilization (see Note 9).
8. Amputate the oral end by cutting perpendicularly to the oral-aboral axis below the pharynx in one straight cut. See red line in Fig. 2.12a for the amputation plane used here.
9. Gently transfer the aboral end to a well of a six-well plate filled with 3 mL Nematostella medium using a plastic Pasteur pipette with a wider opening (see Note 10).
10. Keep the plate at 22°C.
11. Repeat steps 5–8 three more times to transfer the replicates to separate wells.
12. Transfer the two control polyps to the remaining wells.
13. Keep the plate in an incubator at 22°C.
14. Do not feed control or regenerating animals.
15. Wait until the desired timepoint is reached (see Note 11).

16. Transfer each sample in a separate 1.5-mL low-binding tube using a plastic pipette.

17. Spin down each sample by brief centrifugation (<1 s) on a tabletop centrifuge.

18. Remove the superficial liquid layer, while keeping each sample immersed.

19. Add 1 mL of dissociation medium to wash.

20. Repeat steps 16–18 for another wash step.

21. Adjust total volume to 300 μL with dissociation medium (see Note 12).

22. Add 3 μL of thawed liberase$_{\mathrm{TM}}$ solution.

23. Gently pipette up and down slowly at RT (>1.5 s per pipetting cycle) using a trimmed P200 gelatin-coated tip (see Note 13). Use different tips for each sample to avoid cross-contamination.

24. Transition to an uncut gelatin-coated P200 tip with a narrower opening when the diameter of the remaining particles is small enough to pass through the new opening.

25. Add 30 μL of 0.5 M EDTA to stop the dissociation.

26. Add 420 μL of dissociation medium.

27. Gently pipette up and down slowly with gelatin-coated P1000 tips to homogenize.

28. Fill the P1000 tip with the cell suspension.

29. Place the tip directly atop the filtering cap of the 5-mL FACS tube, at an angle.

30. Gently eject the cell through the filter, aiming at the side of the tube (see Note 14).

31. Add 750 μL of staining medium.

32. Slowly pipette up and down with a gelatin-coated P1000 tip to mix.

33. Protect from light and place tubes in a tube holder atop ice.

34. Bring tubes to the flow cytometer facility (see Note 15).

Live Single-Cell Sorting Using Flow Cytometry for Single-Cell RNA-Sequencing

1. Initialize flow cytometer according to manufacturer's instructions, including laser alignment testing and time delay setup for sorting (see Note 16).

2. Adjust plate holder arm coordinates so that the sort stream deposits a single droplet within the center of each well of a cell capture plate (see Note 17).

3. Display the forward scatter-area (FSC-A) signal and side scatter-area (SSC-A) signal in the biexponential scale for all samples and adjust the photomultipliers (PMT) voltages and thresholds to visualize the dynamic range of particles (Fig. 2.13A/B) (see Note 18).

4. Select singlet particles from all the events by excluding multiplets based on their forward scatter-width vs. forward scatter-height signal (gated population is in blue in Fig. 2.13C).

5. Set the photomultiplier voltages for the channels associated to the live and dead cell dyes fluorescent channels using positive (stained) and negative (fluorescence-minus one—FMO) populations (see Note 19).

6. Select live singlet particles from the singlet population (in blue in Fig. 2.13c) by gating the populations according to the live cell marker and dead cell marker (gated population in purple in Fig. 2.13D).

7. Select the live single cell-sized particles (3–20 μm) to sort from the live singlet population (in purple in Fig. 2.13d) based on their SSC-A vs. FSC-A signal (gated population

in cyan in Fig. 2.13e). The setup of the cell-sized gate is presented in Fig. 2.14 (see Note 20).

8. Thaw the 384-well capture plates.

9. Sort one single cell of interest (precision mode: single cell with yield mask: 0, purity mask: 32 and phase mark: 16) per well in each 384-well capture plate, leaving four control wells empty.

10. Spin down the plate at 1000 rcf for 1 min and seal the plate.

11. Freeze on dry ice until the rest of the plates have been sorted.

12. Store at -80°C until processing (see Note 21).

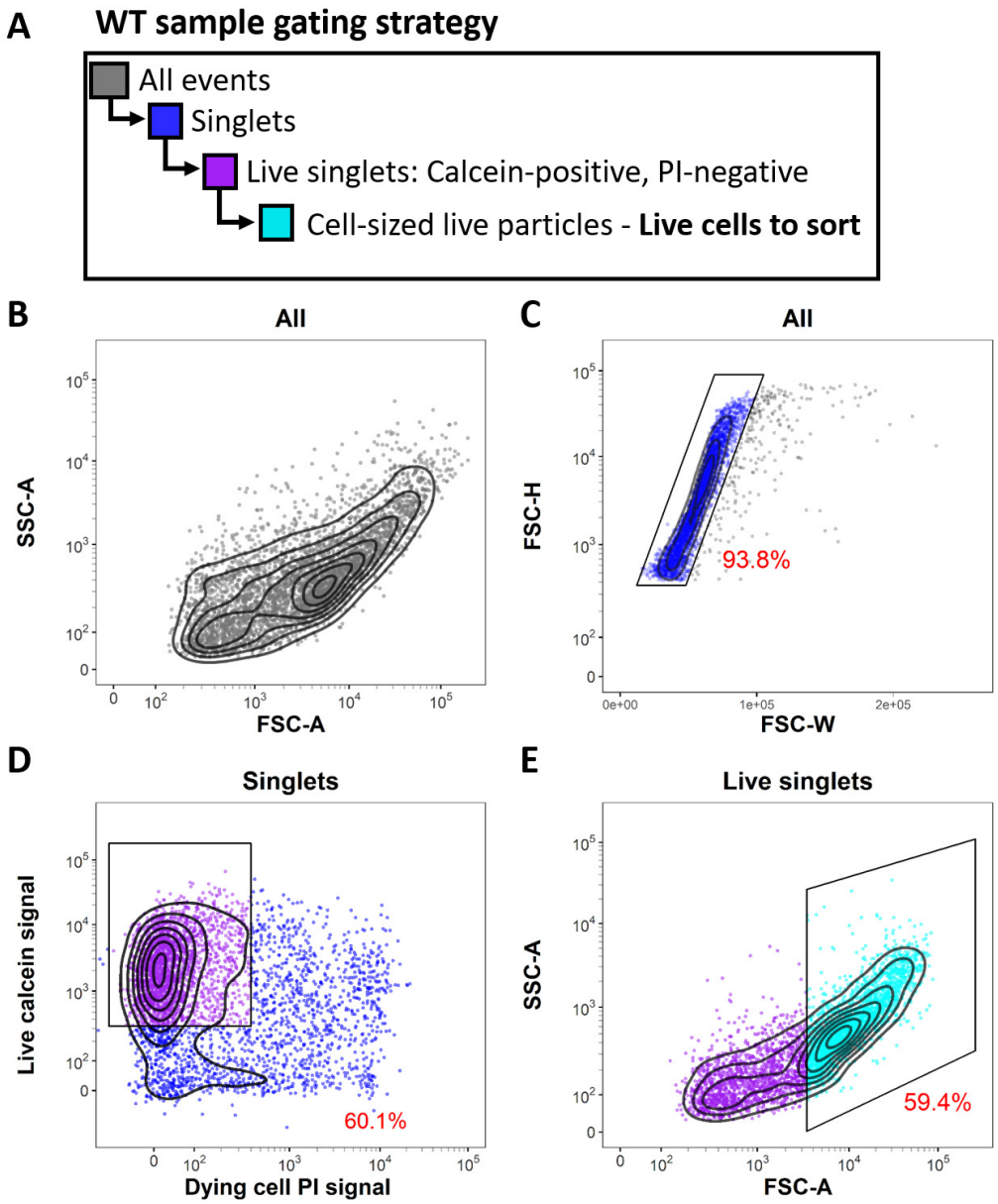

Figure 2.13: FACS gating strategy for live cell-sized singlets with live/dead double-staining

(**A**) Gating strategy to select live single cells from sample stained with calcein AM and propidium iodide (PI). (**B**) Overall events SSC-A vs. FSC-A profile. Note that we are using bi-exponential scaling on both FSC-A and SSC-A because of the size heterogeneities in our samples. (**C**) Singlet gate to exclude multiplets based on FSC-H vs. FSC-W profile. Singlet proportion is usually in the 90–95% range. (**D**) From the singlet population, the live cell gating strategy is based on calcein (live) signal inclusion and propidium iodide (permeabilized dead cells) signal exclusion. (**E**) From the live singlet population, setup of the cell-sized particle gate, to keep only live single cells for sorting. For setting up this cell-sized gate boundaries, see Fig. 2.14G. The proportion of cells that fall within the represented gate is indicated in red.

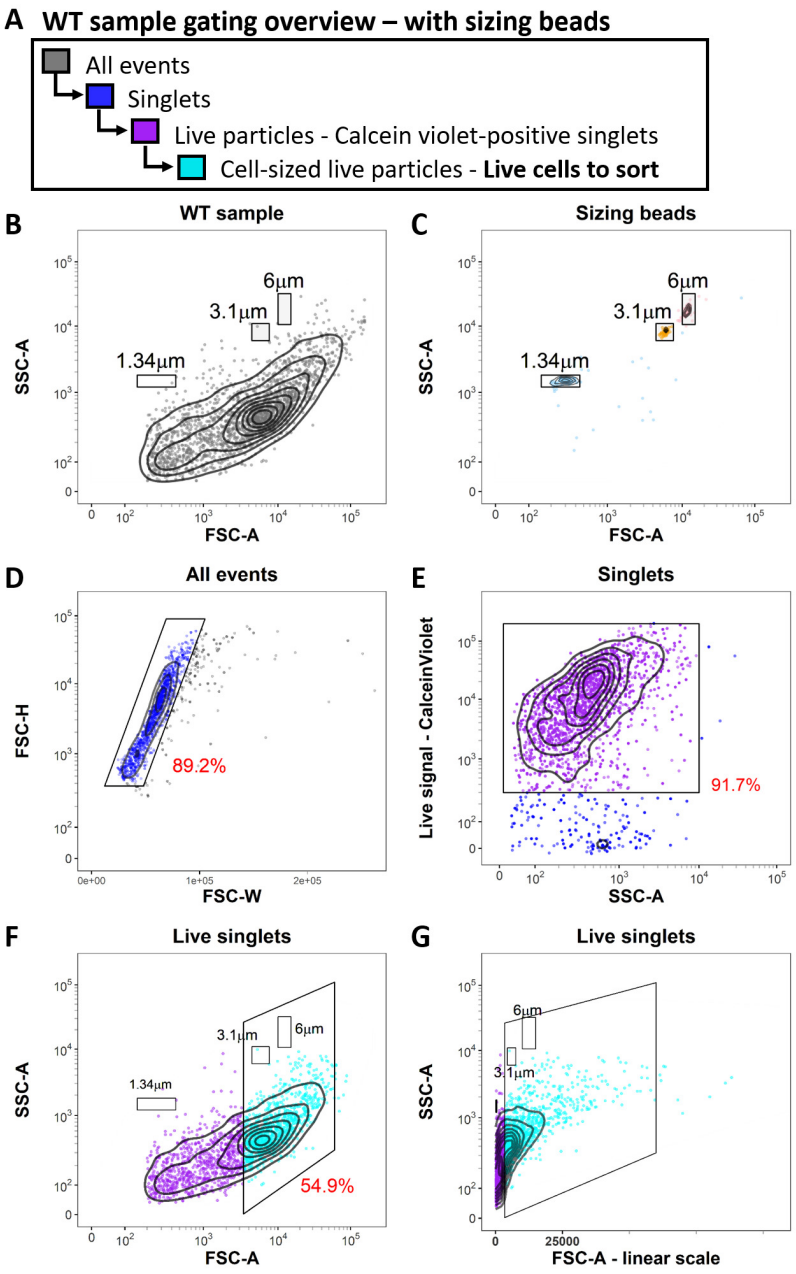

Figure 2.14: FACS gating strategy setup for live cell-sized singlets with absolute size comparison (**A**) Gating strategy to select live single cells. (**B**) All events gate from a whole animal cell suspension (SSC-A vs. FSC-A). Note that we are using bi-exponential scaling on both FSC-A (except in panel g) and SSC-A given the inherent size heterogeneities of *Nematostella vectensis* whole organism cell suspensions. (**C**) Overlay of three sizing beads FCS-A profiles (1.34 μm, 3.1 μm, and 6 μm beads) to set the FSC-A boundary of the cell-sized gate presented in panel G. (**D**) Singlet gate to exclude multiplets based on their FSC-H vs. FSC-W profile. (**E**) Live cell gating strategy based on calcein Violet signal to keep only cells with calcein (live) signal. (**F**) and (**G**) Cell-sized particle gate. Gate boundary is based on beads sizing gate, including all particles 3 μm and above to maximize cells and minimize debris. Representation in a biexponential scale enables clearer demarcation of cell-sized live particles. The proportion of cells that fall within the represented gate is indicated in red.

Single-Cell RNA-Seq Library Generation, QC, and Initial Characterization

1. Process plates through the MARS-Seq library preparation pipe-line with N=17 cycle of PCR amplification at the last step.

2. Check library fragment size distribution using the DNA electrophoresis equipment, according to manufacturer's instructions. MARS-Seq libraries should have a homogeneous fragment size distribution of around 410–440 bp (Fig. 2.15A for an example profile).

3. Check library concentration using the DNA concentration fluorometer, according to manufacturer's instructions. *Nematostella* libraries from a half-plate (190 cells) are routinely in the 1–5 ng/μL range, for a total yield of 20–100 ng per library.

4. Pool libraries equimolarly and sequence using an high output sequencer according to the manufacturer's protocol (see Note 22).

5. Run the MARS-Seq computational pipeline to filter, demultiplex and map sequencing reads onto the *Nematostella* genome according to the user manual (see Note 23).

6. Check that sequencing depth across libraries is homogeneous (Fig. 2.15B). Library sequencing depth depends on the number of libraries pooled and on the sequencing kit used. Here, we sequenced libraries to a median depth of 40,000 reads per well (see Note 24).

7. Check the distribution of UMIs in each library (Fig. 2.15C). Median cell UMI number in *Nematostella* samples is usually around 400–800 UMIs per library, depending on sequencing depth, cell, and sample type. UMI recovery per cell correlates with the overall cell size, with bigger cells yielding higher UMI counts.

8. Check noise level from the recovered UMIs in the control empty wells. UMI recovery from empty well are significantly lower than for wells with cells (red Xs on Fig. 2.15c).

9. Filter out cells based on a UMI threshold. Threshold should be over the noise level. We usually use 100–300 UMIs as a lower boundary. Here, boundaries were set at 120–3000 UMIs (Fig. 2.16A).

10. Cluster cells passing QC using the clustering pipeline, para-meters will depend on cell number and sample cell type diversity. See 2D projection of a pilot experiment on 1200 cells in Fig. 2.16B-D, with a subset of key known markers to propose a cell cluster annotation.

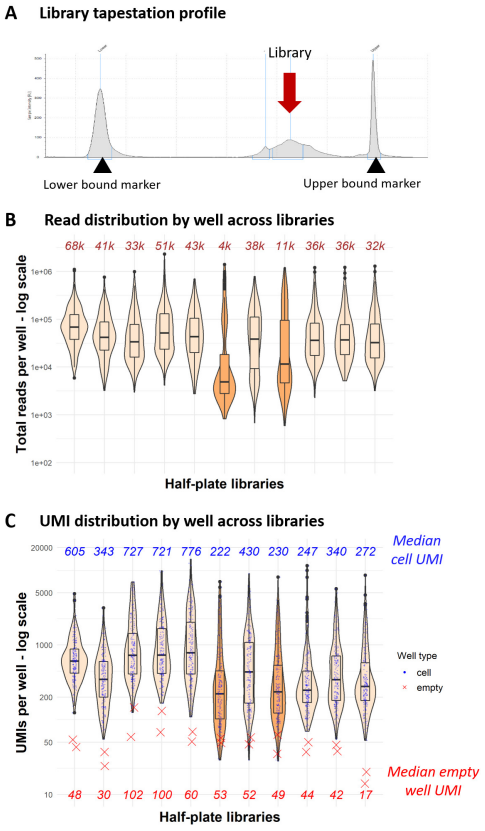

Figure 2.15: Library quality metrics of a sequenced single-cell RNA-seq experiment
(**A**) Example of a typical library tapestation profile of a single MARS-Seq half-plate library (red arrow). Expected fragment size is usually in the 410–440 bp range. Black arrowheads indicate the upper and lower lane markers on a D1000 high sensitivity tape. (**B**) Sequencing depth statistics across sequenced libraries, by half-plate (192 cells) library. Median read depth per library is indicated in dark red at the top. Median sequencing depth target is 40,000 reads/cell when sequencing 32 MARS-Seq libraries on a NextSeq500 high output kit. Note the two libraries flagged in orange whose median sequencing depth is much lower, indicating potentially either a sample issue, or library undersequencing or library quality issues. As their matched other half-plates show the expected sequencing depth, it seems likely to be a technical issue arising during processing or library pooling. The other libraries pass this check. (**C**) Unique Molecular Identifiers (UMI) recovery statistics across half-plate libraries. Same libraries as shown in b. Note that flagged libraries (in orange in b and c) also display a lower UMI distribution compared to their other half-plate. The two empty wells kept as negative controls (red Xs) per half-plate display very low UMI recovery (noise) compared to wells that held one cell (blue dots).



Figure 2.16: UMI distribution in cells passing QC and clustering example from a *N. vectensis* regeneration experiment

(**A**) UMI distribution across all cells. Here, cells from which between 120 and 3000 (red lines) UMIs were recovered were kept for downstream analyses. (**B**) 2D-projection of 1300 cells clustered here using the MetaCell analysis package, colors are used to demarcate individual MetaCell clusters. (**C**) Gene expression (UMIs per 1000 UMIs in the cluster) for known marker genes from[8] are used to annotate the clusters. (**D**) 2D-projection of control and regenerated oral regions at 4/6 days post-amputation. Overall 2D clustering is split by sample origin (control or regenerated), and colored according to a preliminary MetaCell annotation based on known markers published in the *N. vectensis* cell atlas. Most cells are split according to their prospective cell state, with newly regenerated oral side samples falling within all broad cell types.

2.6.5 Notes

1. For cells labeled with a fluorophore overlapping with the calcein AM or propidium iodide (e.g., transgenic fluorescent reporter), it is possible to use the calcein violet (live dye) instead of calcein AM, and SytoxRed (dead cell dye) instead of propidium iodide.

2. For the calcein AM signal: fluorescein (FITC) channel, laser excitation: 488 nm, emission filters: LP 502 nm, and BP 530/30 nm. Propidium iodide (PI) signal: phycoerythrin (PE) channel, laser excitation: 561 nm, emission filters: BP 582/15 nm. If sorting mOrange-labeled cells, the following channel combinations can be used: calcein violet signal (live dye): laser excitation, 405 nm, emission filters: BP 450/40 nm. mOrange signal: phycoerythrin (PE) channel, laser excitation: 561 nm, emission filters: BP 582/15 nm. "Green" autofluorescence signal: fluorescein (FITC) channel, laser excitation: 488 nm, emission filters: LP 502 nm, and BP 530/30 nm. In that case, mOrange-positive cells should be sorted from live singlets from a mOrange signal/Green autofluorescence signal visualization, by selecting mOrange signal-positive cells that are not also GFP-like positive.

3. A detailed step-by-step protocol to efficiently prepare the MARS-Seq capture plates, with lysis buffer and specific poly-T barcoded capture oligos has been published previously[232].

4. Genome file[160] can be downloaded from the JGI website at https://genome.jgi.doe.gov/Nemve1/Nemve1.home.html. Annotation according to [234].

5. We have been following the step-by-step MARS-Seq 2.0 protocol published by the Amit and Tanay's labs[237] to generate sequencing-ready libraries.

6. This protocol uses MARS-seq2.0[237] as provided on the website of Amos Tanay's group;

http : *//compgenomics.weizmann.ac.il/tanay/?page_id*1/4672.

7. We have been using MetaCell, a clustering algorithm based on K-nn graph partitioning[234].

8. A single polyp will yield ample cells to populate all the sixteen 384-well plates for a full MARS-seq run. However, as FACS-sorted plates can be associated with index information, multi-ple replicate polyps can be processed in the same sequencing run.

9. If the animal is not fully relaxed, wait for it to relax before adding $MgCl_2$. If the animal contracts upon $MgCl_2$ addition, wash it out of the $MgCl_2$ solution by moving it into bigger volumes of *Nematostella* medium and use another animal.

10. It is best to cut off the tip of a 3-mL plastic Pasteur pipette to avoid damaging the polyp with the smaller opening while transferring it.

11. 4 days and 6 days post-amputation and control non-amputated animals were used here.

12. For processing a larger amount of starting material, the dissociation volume can be scaled up as well as all the subsequent reagents, or/and samples can be split across several tubes.

13. Liberase activity is optimal at 37°C; however, to avoid heat stress, all steps are carried out at RT (22–24°C), and dissociation usually takes between 10 and 30 min depending on sample density, type, and stage.

14. If the cell suspension is very dense or large fragments of extra-cellular matrix remain, the filter may become clogged. In that case, changing the filter and rediluting the sample prior to refiltering is recommended.

15. Cell concentration, morphology, presence of multiplets, and viability should be checked on a 10 μL aliquot of the cell suspension at a microscope prior to FACS-sorting.

16. It is recommended that the FACS be set up and calibrated prior to starting the dissociation. CS&T beads^{TM} Fluorospheres (BD Biosciences) are used to perform optical QC (delay performance measurements) as well as establish laser delay. For additional settings for single-cell deposition, we have used Accudrop^{TM} fluorescent beads. Sheath fluid (BD FACS Flow) was filtered twice through a 0.2-μm filter. The sheath flow formed stable droplets using a 70-μm nozzle.

17. Single-cell deposition efficiency in each well of the 384-well plate should be controlled prior to experiments using a colorimetric method[239]. Briefly, this uses 3,30,5,5-0-tetramethylbenzidine (TMB) and horseradish peroxidase (HRP) to verify whether a droplet from the cell sorter has successfully reached the fluid in the bottom of the well. The droplet contains the enzyme HRP, which is deposited into a well containing the substrate TMB. When HRP and TMB come into contact, TMB is oxidized and rapidly turns to blue color. Parameters adjusted on the Aria III before single-cell deposition sorting include forward scatter area, side scatter area, and fluorescent area parameters. Forward scatter-area, forward scatter-width, forward scatter-height, side scatter-area, side scatter-width, and side scatter-height are used to exclude multiple cell containing droplets and ensure single cells were deposited. Higher acquisition rates will generally increase the likelihood that droplets will contain multiple cells; therefore, low flow rate is kept constant throughout sorting.

18. Using a biexponential scale for the FSC-A signal is usually more informative because of the very broad size range (ca. 3–25 μm, with a majority of 4–15 μm cells) seen in *Nematostella* cells. The default FSC-A machine threshold may need to be lowered as some Nematostella cells are much smaller.

19. If sorting for a GFP-like fluorophore reporter, whose spectrum overlaps with the calcein AM live cell dye, the live dye can be substituted for calcein violet (e.g., 405 nm, emission

filters: BP 450/40 nm), if there is no spectral overlap. Calcein AM and calcein violet signal are completely coherent (i.e., a double-stained sample will exhibit signal only in the diagonal in the calcein AM/calcein violet visualization plane).

20. Multiple sizing beads (ideally with 1, 3, 6, and 10 μm sizing beads) are used as an absolute size reference for the FSC-A signal, which correlates with particle size. On a FSC-A vs. SSC-A display, most live single cells should be in the 3–20 μm range, with the majority in the 4–15 μm range. Alternatively, a solution of freshly spawned *Nematostella vectensis* mature sperm displays a homogeneous FSC-A vs. SSC-A profile and can be used as a 3 μm threshold reference.

21. Plates of captured sorted cells can be stored over at least several months at -80°C as some samples have been processed and gave similar yields after more than 1.5 years in storage.

22. We routinely used a NextSeq500 high output kit from Illumina but similar coverage can be obtained using several Illumina HiSeq lanes.

23. The pipeline generates UMI count tables used for subsequent clustering as well as a quality control (QC) report whose panels are presented and commented in depth in the MARS-Seq protocol[237].

24. If specific libraries are undersequenced, that could indicate a library pooling issue, or a library quality issue. Repeat DNA electrophoresis analysis of the library to check library profile and re-assess library concentration, if the profile is as expected and the concentration in the expected range for the number of amplification cycles used, use the new concentration to repool and sequence with other libraries. If read number is heterogeneous within a library, that can be due to technical issues during cell sorting or plate processing. Check the distribution of reads or UMIs across the plate, sometimes

plate borders or plate sides display overall lower recovered UMI counts linked to well localization on the plate because of a plate processing issue.

2.7 Conclusion

The tools developed in this chapter are used in chapter 3 to study the cellular lineages underlying adult neurogenesis in *Nematostella*. Beyond methodological advances, initial experiments in the context of amputation-induced oral regeneration provided initial data supporting faithful regeneration and re-specification of all oral cell types, but also very clearly exposed the lack of knowledge about the cellular lineages, source populations and specification pathways that give rise to differentiated cell types in these animals. The absence of this foundational knowledge made it extremely challenging to interpret what is occurring in this regenerative perturbation context, which combines wound healing, spatial patterning, tissue remodeling and morphogenesis in addition to the differentiation of multiple cell types.

To understand adult developmental processes and how animals, like sea anemone, achieve extensive lifelong cellular renewal, we thus elected to study homeostatic cellular renewal in polyps, to remove the confounding effect of global morphogenesis and tissue remodeling at play during both amputation-induced regeneration or early embryogenesis. In such an uninjured context, the body plan is stable, and the animal is growing (as it does lifelong), so there is no gross morphological changes, and this is as close as a steady-state of cell renewal and cell specification as there can be. In chapter 3, we thus investigate the generation of neural cells in the non-perturbed context, to understand the cell sources, trajectories and gene programs associated with acquisition of distinct neuronal identities.

CHAPTER 3

CELLULAR AND TRANSCRIPTIONAL TRAJECTORIES OF NEURAL FATE SPECIFICATION IN SEA ANEMONE UNCOVER TWO MODES OF ADULT NEUROGENESIS

This is an edited version (text and figure modifications) of the accepted article[240]:

Plessier, F., Marlow, H. Cellular and transcriptional trajectories of neural fate specification in sea anemone uncover two modes of adult neurogenesis. Nature Communications (2026).

3.1 Abstract

While neurogenesis is largely restricted to early life stages in animals, some taxa (including cnidarians, planarians and acoels) display lifelong neurogenic abilities. The cellular lineages and transcriptional programs underlying this process remain poorly understood in cnidarians. Combining reporter tracing and single-cell transcriptomics, we identify adult neurogenic trajectories in the sea anemone *Nematostella vectensis*. We uncover two distinct mechanisms: direct differentiation of peptidergic neurons from a multipotent progenitor pool, with identities specified proportionally to existing populations, and a stepwise maturation of cnidocytes (specialized cnidarian neural cells), marked by a transcriptionally distinct intermediate stage. Neural fate commitment is characterized by transient *SoxC* expression, with a homeodomain code associated with neural identities. These modular strategies support continuous neurogenesis and suggest that neural fate specification in cnidarians may use ancestral principles shared with bilaterians. Our study provides a foundational framework for understanding the mechanisms underlying the adult specification of neural cells and their evolutionary diversification across animal lineages.

3.2 Introduction

Neurons are the fundamental units of nervous systems across animals. In most bilaterians (e.g. humans, mice, and flies), neural development is predominantly restricted to early life stages, with limited renewal capacities in adulthood. In contrast, lineages with extensive regenerative abilities and less complex nervous systems, such as cnidarians, acoels, and planarians, maintain ongoing neuronal renewal throughout life[241]. Cnidarians (e.g., sea anemones, corals, hydra, jellyfish), the outgroup to Bilateria, lack a centralized nervous system, with neurons embedded in diffuse nerve nets[119,175]. However, cnidarian and bilaterian neurons share key effectors, including conserved neuropeptide processing machinery, ion channels, and synaptic proteins, as well as early neurogenic transcription factors like proneural basic helix-loop-helix (bHLH) and SoxB genes, pointing to a common evolutionary origin[8,25,50,54,142,171]. Major diversifications of transcription factor (TF) families central to bilaterian neural programs -such as Homeodomain, Sox, and bHLH- occurred early in animal evolution, with substantial cnidarian-bilaterian shared expansions, which may have contributed to the diversification of neural identities across both groups[39,43,46]. Thus, investigating the cell lineages and transcriptional programs involved in neurogenesis in cnidarians could provide valuable insights into the mechanisms of lifelong generation of molecularly diverse neural cells across animals.

Cnidarian neural cells encompass ganglion neurons, sensory cells and lineage-restricted cnidocytes (mechano-sensory neural cells). In the genetically tractable sea anemone Nematostella vectensis, over twenty subtypes have been molecularly characterized[8,101,104,119,178,181,182], and its nervous system shows localized condensation of neurite tracts along the primary body axis, indicating a degree of morphological organization[38,119]. Neurons primarily employ a specialized neuropeptide repertoire[8,98,101,104,178,182,242], along with other neurotransmitters (e.g., glutamate, GABA, but not monoamines)[33,67,200]. Akin to other highly regenerative species such as planarians, acoels, and annelids[243,244], cnidarians like *Nematostella*

exhibit feeding-dependent indeterminate growth and shrinkage, as well as asexual reproduction[158,166]. Notably, a population of peptidergic neurons (*LWamide-like*$^+$) scales with body size[178], suggesting that extensive adult neurogenesis occurs. However, the progenitor cell population(s) which enable adult neurogenesis and the transcriptional trajectories underlying this adult program of neural cell fate specification remain poorly understood.

Adult generation of neurons can occur via two distinct yet non-mutually exclusive mechanisms. Each of these mechanisms is associated with different degrees of fate restriction and nervous system organization. In the first, associated with highly organized centralized nervous systems, as in the rodent olfactory system or the zebrafish telencephalon, pools of lineage-restricted neural stem cells maintained within spatially delineated niches can generate a limited variety of neural subtypes that are integrated into nearby structures[245]. In the second mechanism, associated with more diffuse centralized nervous systems and nerve nets, as in planarians, acoels and hydra, multipotent migratory stem cells can produce diverse somatic lineages, including most to all neuronal subtypes, under both homeostatic and injury-induced conditions[246,247]. In the adult sea anemone *Nematostella vectensis*, mesenterial *Piwi1*$^+$ cells can give rise to cells expressing the conserved pan-neuronal marker *Elav*[202], suggesting that they are a source for new neurons. It remains unclear how these putative adult stem cells contribute to adult neurogenesis in cnidarians, including which neurons they give rise to, whether intermediate lineage restriction occurs and what molecular mechanisms accompany fate restriction.

In animals with continuous neural cell renewal -like cnidarians or planarians-, the simultaneous presence of all cellular states (from stem cells to intermediate and mature neurons) presents a significant challenge for identifying cells along one or multiple lineage trajectories. This inability to distinguish temporal intermediates precludes trajectory inference and examination of their transcriptional maturation. Single cell transcriptomics (scRNA-seq) can generate unbiased static snapshots of gene expression profiles across entire nervous systems

but lineage inference to investigate cell fate acquisition remains challenging[213]. Imaging of genetic drivers can capture cell dynamics through reporter inheritance, but reconstructing complex trajectories requires multiple targeted genetic tools. These issues are compounded in emerging model systems with limited genetic tools and sparse *a priori* knowledge needed for lineage reconstruction. To address these challenges, we integrated temporal reporter tracing and single-cell transcriptome profiling to map the lineages and transcriptional programs underlying adult neurogenesis in the sea anemone *Nematostella vectensis*. We find that distributed neurogenesis scales the nerve net with peptidergic neurons generated directly from non-neural-exclusive progenitors, and cnidocytes maturing through a transcriptional intermediate stage before initiating their secondary neural program. These parallel transcriptional lineage strategies employ bilaterian-shared transcription factor classes (bHLH, homeodomain) across neural identities, suggesting common neural development principles for the generation of neural diversity.

3.3 Results

3.3.1 Distributed peptidergic neuron generation underlies adult nerve net expansion

To support indeterminate growth, an adult neurogenic program involving continuous generation and integration of neurons into the nerve net would be required. To identify such adult-born neural cells *in vivo*, we employed a characterized *Elav* regulatory sequence active in neurons from embryonic to adult stages to drive the expression of a photoconvertible KikumeGreenRed (KikGR) fluorescent reporter[38,214]. We labeled the standing population of *Elav*$^{+}$ cells by photoconverting this reporter (day 0) and analyzed the body wall nerve net after 4 days. New *Elav*$^{+}$ neurons arising from non-expressing cells would lack the photoconverted tracer and display new reporter expression (hereafter referred to as green cells) (Figs. 3.1 and 3.2A).

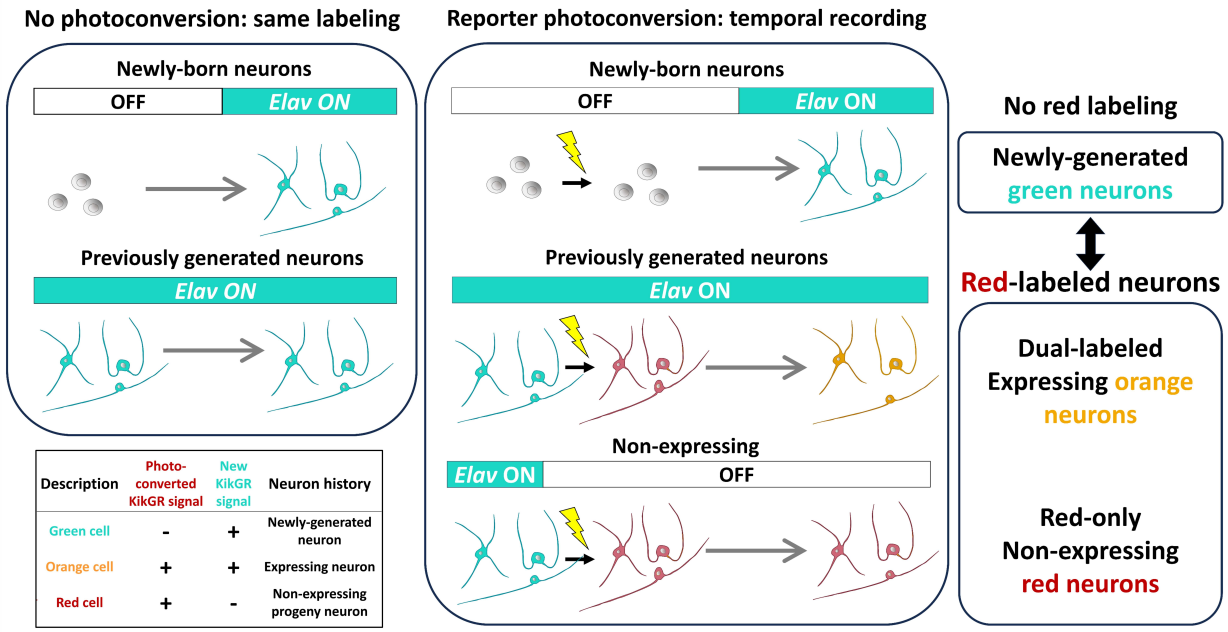

Figure 3.1: Discriminating adult-born neurons by photoconversion-based temporal recording of reporter expression

Expression of a non-photoconvertible fluorophore driven by an *Elav* regulatory region (**left**) does not discriminate between previously expressing cells and newly expressing cells. A photoconvertible reporter (*KikGR*, green to red photoconvertible fluorophore) (**right**) labels all previously expressing neurons at time 0 and then be used to track newly-expressing-only neurons by their presence of the new reporter protein (green signal) and the lack of the photoconverted red labeling. Non-expressing progeny cells of previously expressing cells (red cells here, without newly expressed reporter signal) can be distinguished from expressing progeny cells that contain both the photoconverted red signal and the new green signal (represented as orange cells throughout).

Thus, discriminating green neurons from previously generated stably-expressing neurons (red and green co-expressing, orange cells hereafter) or cells that have ceased expressing the reporter (red cells hereafter), allowed us to investigate the dynamics of neurogenesis *in situ*.

Imaging *Elav*>*KikGR*$^{+}$ animals 4 days post-photoconversion revealed green neurons interspersed with orange neurons in the body wall nerve net (Fig. 3.2), indicating ongoing neurogenesis contributing new neurons to the nerve net. Newly-born neurons are dispersed throughout the epidermal and sub-epidermal layers, including in the neurite tracts at the base of mesenteries (see Supplementary Videos 1-2 available online at doi:10.5281/zenodo.14977133). The majority of detected cells (approximately 80%) are green or orange, with a small fraction

red; the latter category suggests that a subset of differentiated neurons stop expressing the *Elav* driver. Overall, the large fraction of green cells detected 4 days post-photoconversion is consistent with active neurogenesis contributing new neurons to the body wall nerve net through the rapid generation of $Elav^{+}$ neurons from $Elav^{\mathrm{neg}}$ cells (Fig. 3.2).

To evaluate whether newly generated neurons are exclusively incorporated into specific regions of the nerve net, which could point to a spatially restricted growth zone on the oral side of the primary body axis, we examined both the oral and aboral regions. Green neurons are found in both regions at 2, 4 and 7 days post-photoconversion with similar proportions in each region (Fig. 3.2). The synchronous presence of green cells in both regions, coupled with their high fraction, suggest that nerve net expansion occurs through highly active spatially distributed adult neurogenesis, with *Elav*-negative cells differentiating into $Elav^{+}$ neurons within one week.

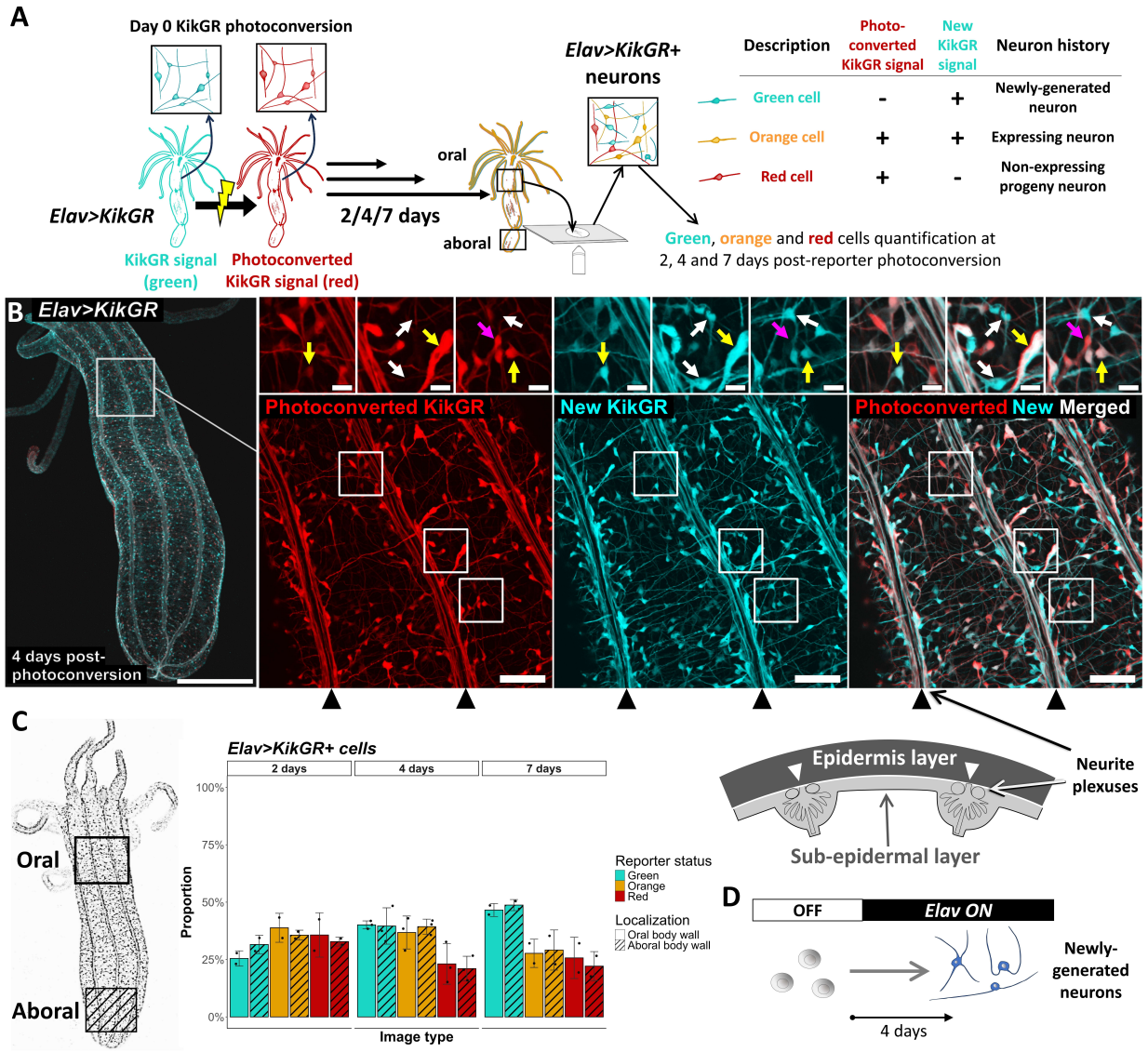

Figure 3.2: Adult neurogenesis across the primary body axis underlies nerve net expansion.

A Active adult neurogenesis highlighted by *Elav>KikGR* reporter photoconversion and *in situ* imaging. At 2,4 and 7 days post-photoconversion, neural cells with the new green reporter signal and lacking the day 0 photoconverted reporter signal are newly-expressing-only cells (green cells). Cells with the photoconverted tracer can display new KikGR expression (ongoing expression, orange cells) or lack new KikGR expression (progeny cells, red cells). **B** *Elav>KikGR*$^{+}$ body wall nerve net at 4 days post-photoconversion (5X view and 40X single-neuron resolution). Photoconverted KikGR tracer signal is shown in red, and new KikGR expression in cyan. Arrows in insets highlight green (in white), orange (in orange) and red (in magenta) neurons. Scale bars are 500 μm (left), 50 μm (right) and 10 μm (insets). **C** Proportion of green, orange and red *Elav>KikGR*$^{+}$ cells detected in the oral and aboral regions at 2, 4 and 7 days post-photoconversion (n=2-3 animals, mean ±S.D.). **D** Generation of *Elav*$^{+}$ neurons from *Elav*-negative cells.

Elav^{+} neurons encompass a diversity of neural cell states[8,38]. To investigate the molecular identity of the newly-generated *Elav*$^{+}$ neural cells, we combined our photoconversion-based temporal recording with single cell RNA-Sequencing (scRNA-Seq) of green, orange and red *Elav>KikGR*$^{+}$ cells Fig. 3.3. We identified 1335 peptidergic neurons from the 3006 *Elav>KikGR*$^{+}$ sorted cells, highlighted by their shared expression of neuropeptide processing genes (prohormone convertases, amidating enzymes). These encompass over a dozen molecularly distinct subtypes, delineated by their restricted expression of neuropeptide precursors and neuropeptide G-protein coupled receptors (GPCR) (Fig. 3.3)[98,101,178]. Notably, we recover all previously described neural populations in the body wall, including a *Prdm14d*$^{+}$ population (Neural 0)[181]. As expected, we do not recover *FoxQ2d*$^{+}$ putative sensory cells and mature cnidocytes, which are not labeled by the *Elav* driver[180]. At 7 days post-photoconversion, we find a majority of the *Elav>KikGR*$^{+}$ neural cells to be newly-generated green cell (58%), consistent with our *in situ* nerve net imaging (Fig. 3.2). Green cells are distributed among all peptidergic neuronal subtypes (Fig. 3.3). Specifically, the number of green cells relative to the combined number of orange and red cells for each cluster was strongly correlated (R^2=0.96, Pearson's ρ). This suggests that *Elav*$^{+}$ neuron subtypes, from the rarer to the more abundant, are generated concomitantly *in vivo*, at a frequency matching their overall proportion in the animal (Fig. 3.3).

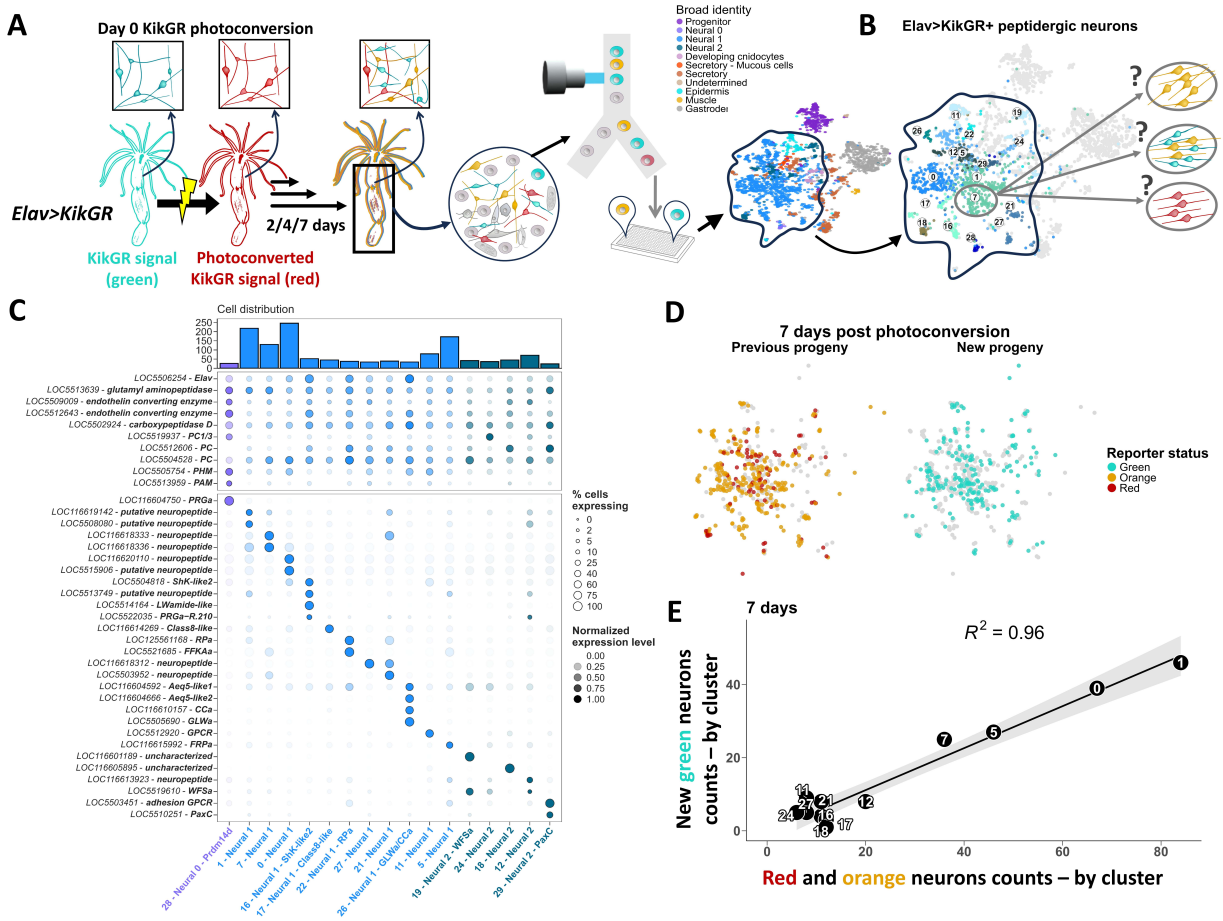

Figure 3.3: Adult neurogenesis generates all *Elav*$^+$ peptidergic neuron subtypes

A Molecular identification of newly-generated *Elav*$^+$ neurons. Plate-based scRNA-Seq (MARS-Seq protocol) allows recording of reporter status post-photoconversion (green, orange or red cell) by flow cytometry for each cell sampled from *Elav>KikGR*$^+$ adult body regions (body wall with mesenteries) at 2,4 and 7 days post-photoconversion. scRNA-Seq transcriptional profiles are then matched to each cell's reporter status color (see Methods). **B** UMAP projection of the 1335 peptidergic neurons from the 3006 *Elav>KikGR*$^+$ sampled cells from 2, 4 and 7 days post-photoconversion with their cluster identifier. **C** Scaled neural marker gene expression across *Elav>KikGR*$^+$ peptidergic neuron clusters. Dot size represents the percentage of cells in this cluster which express the gene. Neural cluster identifiers as in B. **D** *Elav*$^+$ peptidergic neuron scRNA-seq UMAP projection highlighting the overlapping identity distribution between the previously-labeled peptidergic neurons (orange and red cells) (**left**), and the newly-expressing neurons (green) (**right**) at 7 days post-photoconversion. **E** Cell counts by reporter status (green vs. red and orange cells) across peptidergic cell clusters in 7 days photoconverted *Elav>KikGR*$^+$ animals. Line is the fitted linear regression with the 95% confidence interval in shaded grey. Cluster identifiers as in B-C.

3.3.2 Progeny of *FoxL2*$^+$ and *SoxC*$^+$ cells encompass neurons, cnidocytes and secretory cells

Two new drivers (*SoxC* and *FoxL2*) to track neural lineage progression

In silico trajectory inference approaches suggest that in the adult, peptidergic neurons, secretory cells, and cnidocytes (cnidarian-specific neural effector cells) are all part of a shared lineage[53]. However, it is unknown whether these populations derive from a common progenitor pool and what molecular features are associated with neural lineage restriction during adult growth. To better understand neural cell fate acquisition, we generated photoconvertible reporter lines for *SoxC* and *FoxL2*, two transcription factors expressed in putative progenitor cells in adults to track their cellular progeny. Earlier work suggests *SoxC* plays a role in the differentiation of cnidocytes, secretory cells and neurons during embryogenesis, and *SoxC* reporter expression has been noted across these cells in the adult[53]. *FoxL2* is a forkhead box transcription factor highly expressed in progenitor cells and tentacular cnidocytes and lowly-expressed in some neurons[8,248] in *Nematostella*, and in ectodermal neural precursors and neurons in *Hydra*[249], a distantly related hydrozoan cnidarian. Together these two reporters allowed us to probe the trajectories of early undifferentiated cell populations that may give rise to diverse neural cells in the adult.

Characterization of *SoxC>KikGR* expression in juvenile polyps highlights neural progeny cells

We first investigated the temporal activity of *SoxC* in sexually immature adults. Imaging the body wall reveals that *SoxC>KikGR* reporter expression in morphologically diverse cells, across the epidermal and subepidermal layers, including neurons, developing and mature cnidocytes, muscle cells, and secretory cells. To determine whether the neural cells might be part of a *SoxC*$^+$ derived lineage, we assessed the presence of red cells post-photoconversion (Fig. 3.4).

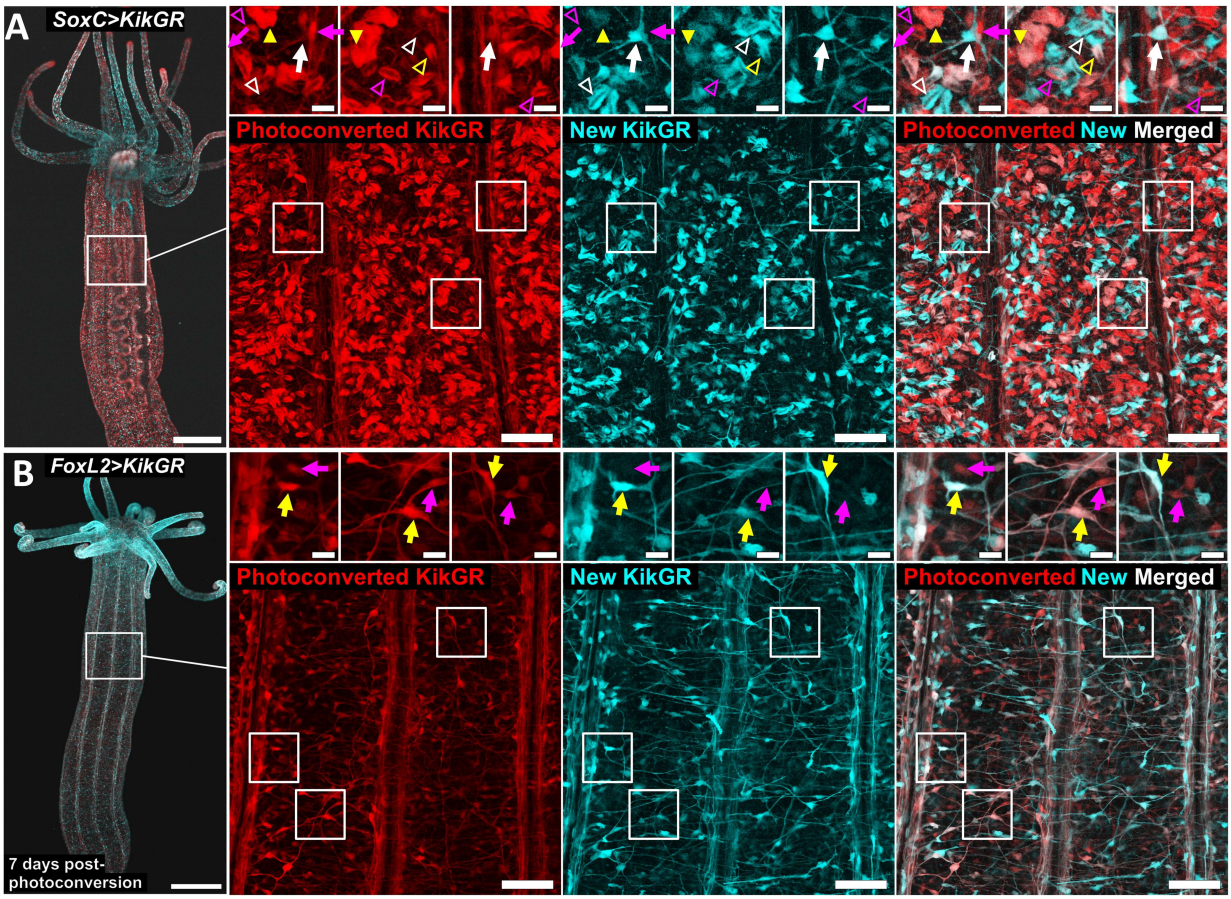

Figure 3.4: Two new photoconvertible drivers uncover neural non-expressing progeny cells.

SoxC>KikGR$^+$ (**A**) and *FoxL2>KikGR*$^+$ (**B**) adult polyp body wall at 7 days post reporter photoconversion. Global field of view (5X, left) and high magnification image (40X extended depth of focus (A) or single image (B), right) with photoconverted KikGR signal shown in red and newly expressed KikGR in cyan. *SoxC>KikGR*$^+$ cells encompass green, orange and red mature cnidocytes, green, orange and red neurons (respectively white, yellow and magenta arrows), and orange secretory cells (yellow filled arrowhead). *FoxL2>KikGR*$^+$ cells encompass orange (yellow arrow) and red (magenta arrows) neurons among other cells. Scale bars are 500 μm (left), 50 μm (right), and 10 μm in the insets.

Thus at 7 days post post-photoconversion, the presence of both red (non-expressing progeny) and green (newly-generated) mature cnidocytes and other neurons shows these are derived from *SoxC*$^+$ cells. The sparse secretory cells detected in the epidermal epithelial layer are mostly orange cells, so these might be mostly actively expressing the driver, in contrast to what is found for neurons and cnidocytes, where most of them are either green

or red cells. Overall, this suggests that $SoxC^{+}$ progenitor cells transiently express the driver early during differentiation and give rise to $SoxC^{\text{neg}}$ derivatives including neurons and cnidocytes. Specifically, red neurons and cnidocytes represent cells generated prior to "day 0" of photoconversion while green cnidocytes and neurons are those that developed within 7 days of the photoconversion "chase". The recovery of primarily red and green cells at earlier time-points (2 and 4 days post-photoconversion,), in similar proportions in oral and aboral regions (Fig. 3.5), further supports the conclusion that early expressing $SoxC^{+}$ cells give rise to $SoxC^{\text{neg}}$ derivatives including neurons and cnidocytes.

Characterization of *FoxL2>KikGR* expression in juvenile polyps highlights neural progeny cells

We next analyzed the expression of the *FoxL2>KikGR* driver. The reporter is expressed in multiple cell types in the body wall, across both epidermal and subepidermal layers, including neurons, mature cnidocytes and epithelial cells, and uncharacterized cells, with sparse secretory and muscle cells (Fig. 3.4). We thus investigated whether these differentiated cells derived from a $FoxL2^{+}$ lineage (presence of red cells) or were actively expressing the reporter (orange cells). At 7 days post-photoconversion, we detect diverse red cells, including neurons and mature cnidocytes, suggesting these are non-expressing cells arising from $FoxL2{>}KikGR^{+}$ cells. The sparse green cells present are mature cnidocytes and morphologically undetermined cells (Fig. 3.6). This shows that $FoxL2{>}KikGR^{+}$ cells give rise to multiple non-expressing derivatives within 7 days, including neurons. Accordingly, a large fraction of red cells (non-expressing progeny cells) are detected at 2,4, and 7 days post-photoconversion (Fig. 3.5). And this adult developmental process is spatially distributed across the primary body axis, as evidenced from the similar fraction of green/orange/red cells across both the oral and aboral regions (Fig. 3.5). Overall, this separately highlights the presence of $SoxC{>}KikGR^{+}$ and $FoxL2{>}KikGR^{+}$ neural progeny cells encompassing both

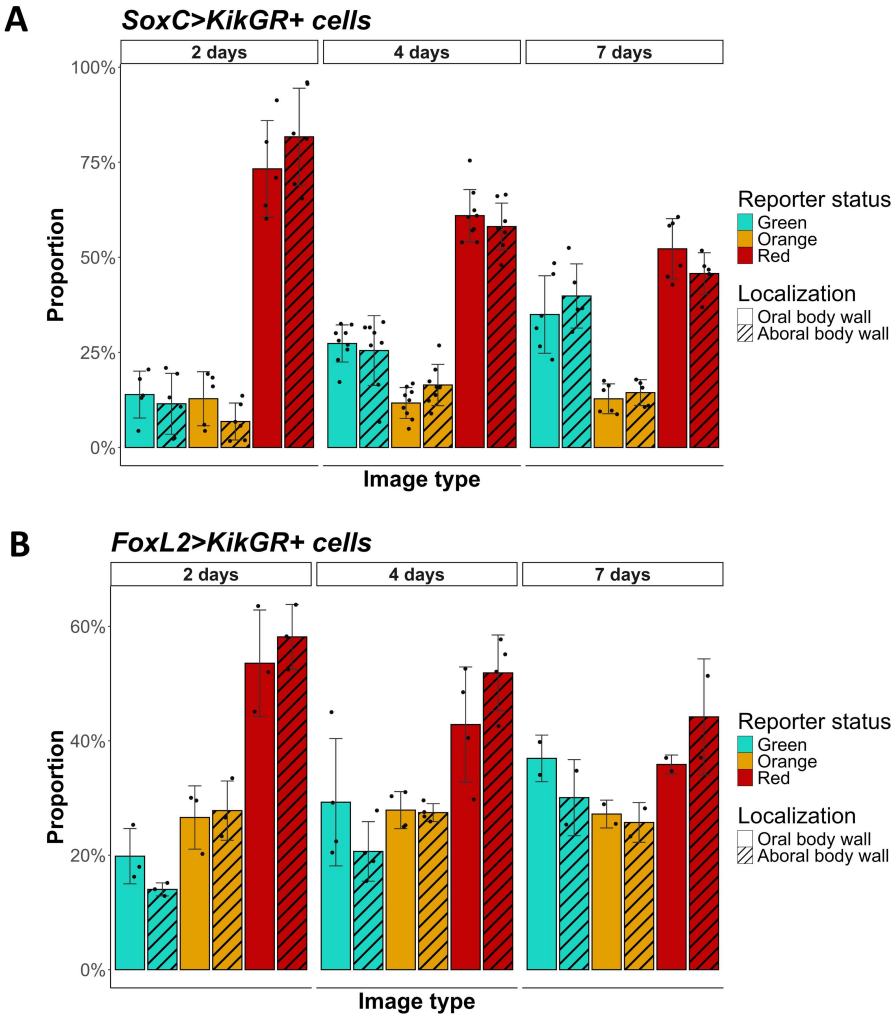

Figure 3.5: Quantification of *SoxC>KikGR*$^+$ and *FoxL2>KikGR*$^+$ cells color distribution at 2, 4 and 7 days post-photoconversion in oral and aboral regions

A-**B** Proportions of detected *SoxC>KikGR*$^+$ (**A**) and *FoxL2>KikGR*$^+$ (**B**) cells in the body wall region by reporter status, time post reporter photoconversion and localization. Cell proportion is normalized to all cells detected within each z-stack image, with the red, orange and green cell fractions. The fraction of green cells increases with the days post-photoconversion across both the oral (no pattern) and aboral (with hachures) areas. n=2-7 animals per reporter and post-photoconversion timepoint, with dots for individual image quantifications (mean ±S.D.).

neurons and mature cnidocytes. These red non-expressing progeny cells thus arise from expressing cells (see below).

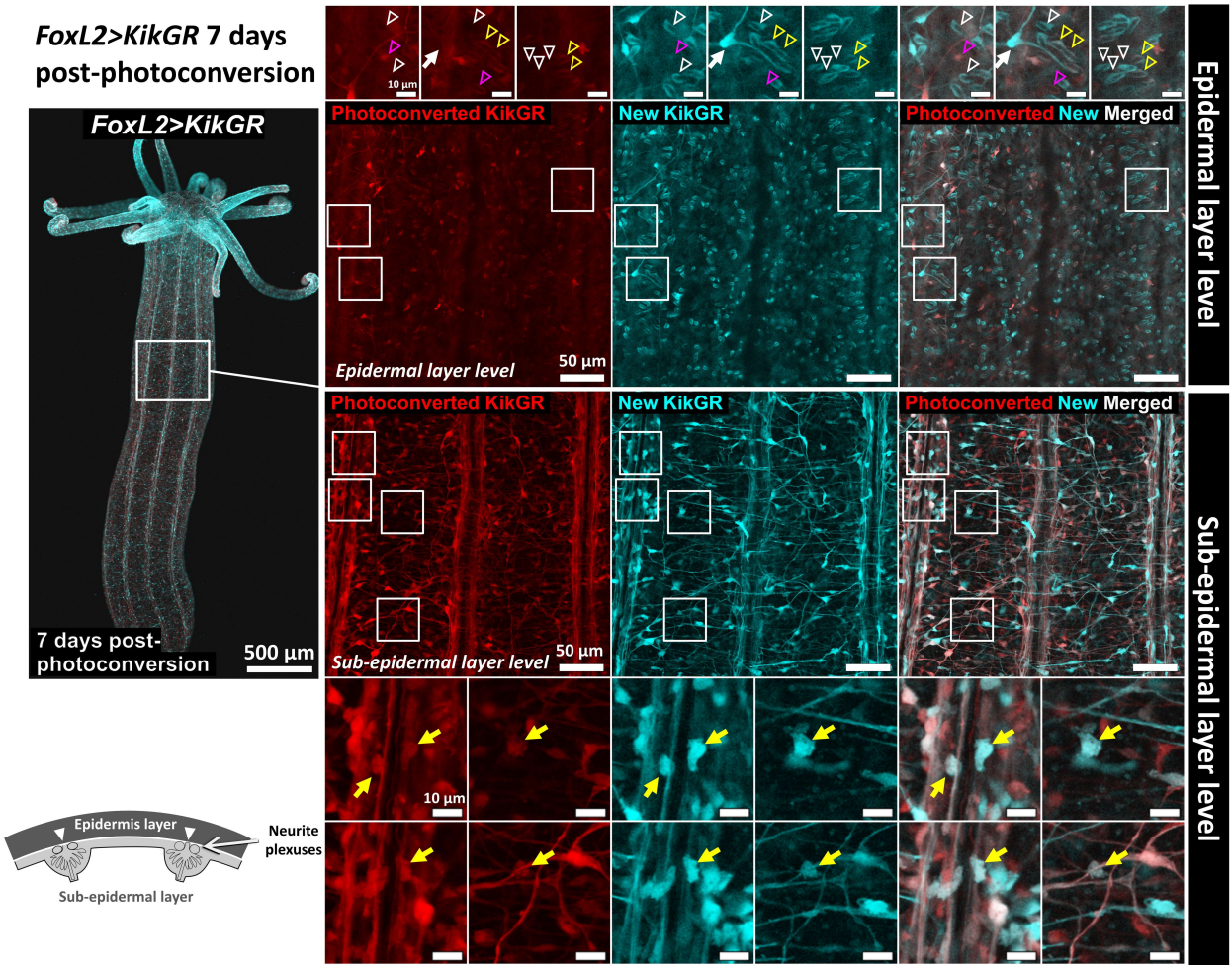

Figure 3.6: *FoxL2>KikGR*$^+$ cells encompass both differentiated (neurons, cnidocytes and secretory cells) and morphologically undifferentiated cells 7 days post-photoconversion

A *In situ* imaging of the 7 day post-photoconversion *FoxL2>KikGR*$^+$ animal shown in Fig. 3.4. Body wall region localization (5X, left), and 40X optical slice at the epidermal layer level (top) and at the sub-epidermal layer level (bottom, same sub-epidermal level as shown in Fig. 3.4) with magnification insets. Differentiated cnidocytes are detected in the epidermal layer, identifiable by their straight intracellular capsule. These cnidocytes include both orange cells (presence of both photoconverted and new KikGR signal, yellow arrowheads) and green cells (only new KikGR signal, white arrowheads), and red cells (only photoconverted KikGR signal, magenta arrowheads). Morphologically undifferentiated orange cells can also be detected in the sub-epidermal layer (yellow arrows). 40X optical slice with magnification insets. Scale bar is 500μm (whole animal overview), 50μm (main panels) and 10μm (insets).

FoxL2>KikGR$^+$ and *SoxC>KikGR*$^+$ cells have broadly overlapping identities

To investigate the molecular signatures of these adult neural developmental lineages, we combined the temporal recording of reporter activity with scRNA-Seq (workflow as in (Fig. 3.3)). We sampled *SoxC>KikGR*$^+$ and *FoxL2>KikGR*$^+$ cells at 2,4 and 7 days post-photoconversion and linked the molecular identity of each cell to its recorded reporter status (green, orange or red cell). We pooled 12,352 KikGR$^+$ cells sorted from the *SoxC>KikGR*, *FoxL2>KikGR* and *Elav>KikGR* drivers across 84 distinct MARS-Seq scRNA-Seq libraries from 27 independent samples and annotated 45 clusters using published marker genes[8,50,53,98,101,132,136,151,180,181] (Fig. 3.7).

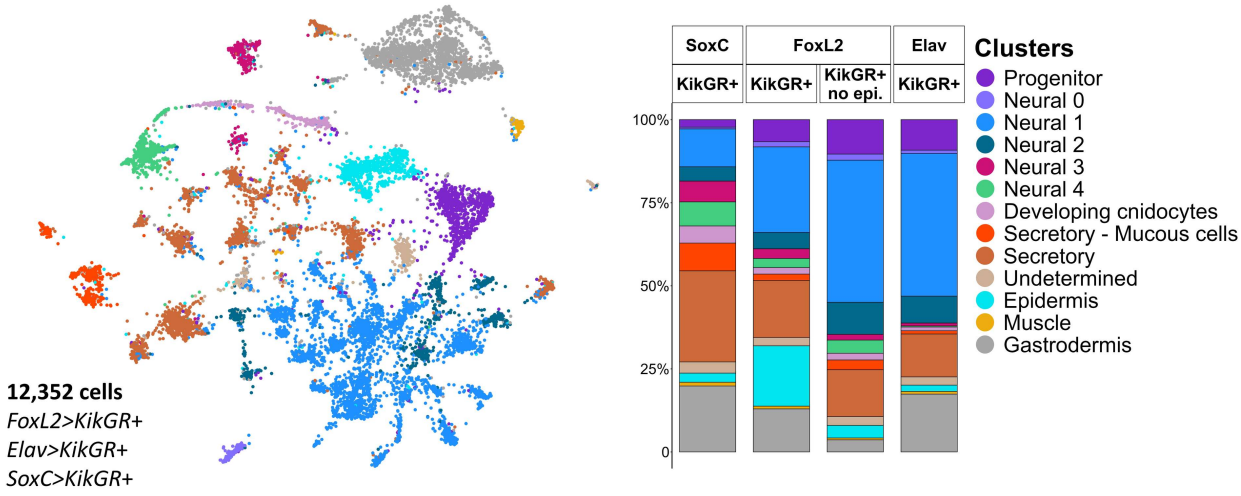

Figure 3.7: Molecular identity of *KikGR*$^+$ cells from all three drivers

(**left**) Global scRNA-Seq UMAP projection of 12,352 KikGR$^+$ cells (2764 *Elav>KikGR*$^+$ cells, 3451 *SoxC>KikGR*$^+$ cells and 6137 *FoxL2>KikGR*$^+$ cells) at 2, 4 and 7 days post-photoconversion, colored by broad cluster identity. (**right**) Cell identity distribution by driver and FACS gating strategy across the *SoxC*, *FoxL2* and *Elav* drivers. For the *FoxL2* driver, the "KikGR+ no epi." FACS gate depleted the epithelial epidermis cell fraction.

Clusters were grouped into 13 broad cell classes, with groups Neural 0 to 4 representing *Prdm14d*$^+$ neurons, *Elav*^{high} peptidergic neurons, *Elav*^{low} peptidergic neurons, and other neural cells including *FoxQ2d*$^+$ sensory cells and *Eag*$^+$ mature cnidocytes[250], respectively.

Overall, our cluster resolution is consistent with subtype-specific gene expression patterns within peptidergic neurons, cnidocytes, and secretory cells. We recover developing cnidocytes which express cnidarian-specific capsule proteins and known regulators, along with mucous cells, and digestive secretory cells (with distinct trypsins, chitinases, and pancreatic lipases)[151]. We identify epidermal epithelial cells, mesenteric retractor and parietal muscle cells[8,173], cnidarian-specific gastrodermis cells, undetermined cells, and putative progenitor cells ($SoxB(2)a^{high}/SoxC^{high}$). Peptidergic neurons are recovered from all three drivers (*Elav*, *FoxL2* and *SoxC*), while cnidocytes, mucous secretory cells and epithelial epidermis cells are mostly absent from the *Elav* driver (Fig. 3.7)), consistent with the *in situ* imaging. The partially overlapping molecular identities of the neural populations suggest that peptidergic neurons represent a more restricted lineage distinct from other neural fates.

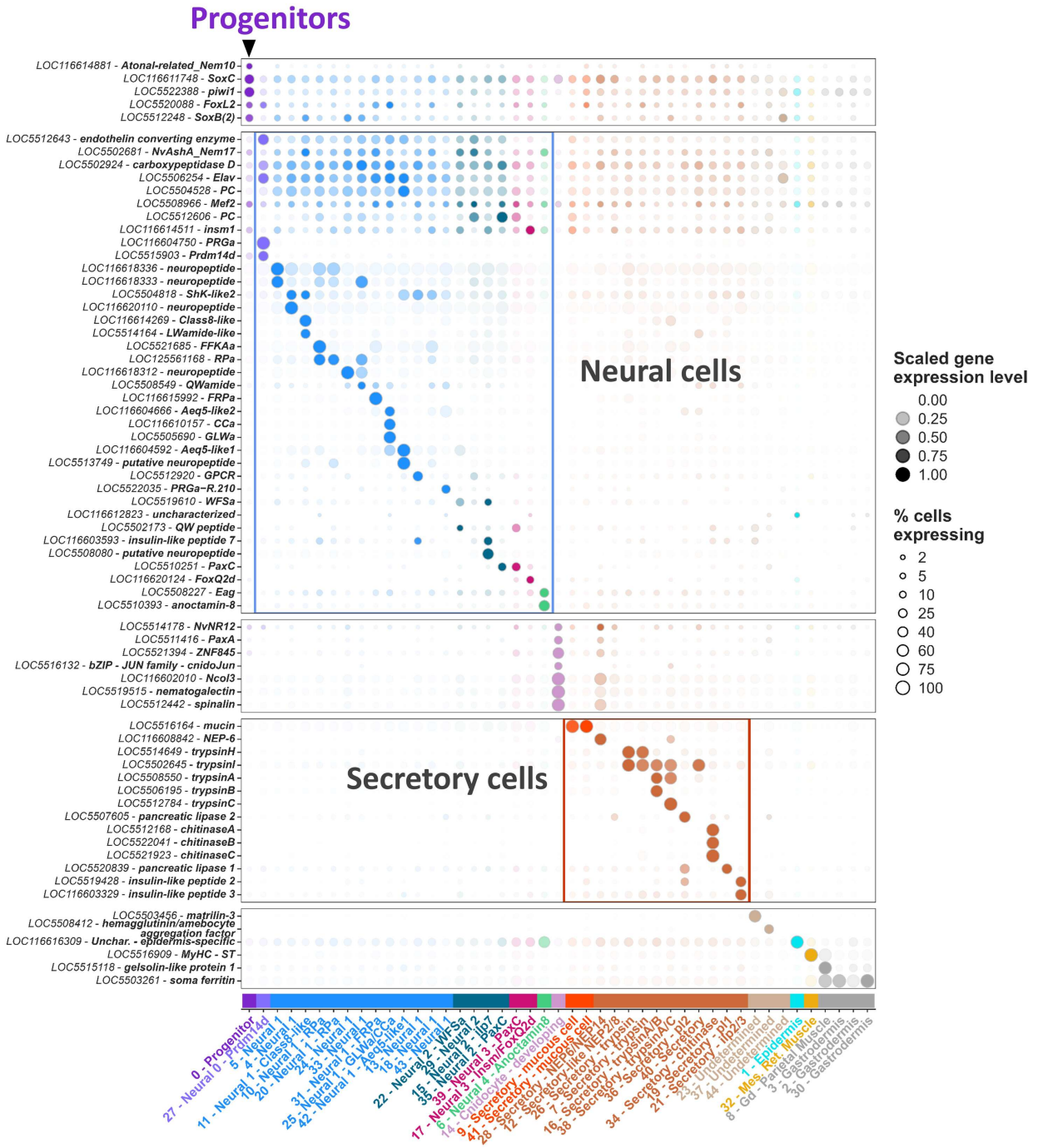

Figure 3.8: Molecular characterization of all KikGR$^+$ cells highlights peptidergic neurons, cnidocytes, secretory cells and putative progenitor cells

Marker gene expression across clusters from the combined *Elav>KikGR*$^+$, *SoxC>KikGR*$^+$, and *FoxL2>KikGR*$^+$ cells clustering from (Fig. 3.7). For each gene, expression is normalized to the highest expression across clusters. Colors correspond to broad cluster annotations, as in Fig. 3.7. Dot size represents the proportion of cells expressing the gene in the cluster.

3.3.3 Tracking of *Piwi1*$^+$ progenitor cells differentiation dynamics

A subset of *Elav*>*KikGR*$^+$, *SoxC*>*KikGR*$^+$ and *FoxL2*>*KikGR*$^+$ cells have a progenitor molecular signature

As putative progenitor cells display the highest *SoxC* levels (Fig. 3.8), and neural cells can arise from unknown *SoxC*>*KikGR*$^+$ cells (Fig. 3.4), we examined these undifferentiated "progenitor" cells recovered from all drivers (in purple in Figs. 3.7 and 3.8) as a source population for neural cell derivatives. These "progenitors" specifically express genes associated with proliferative pluripotent cells across metazoans, such as transcription factors from the bHLH (including *Myc* and atonal-related genes), Sox B and C families, RNA helicases (*PL10*, *vasa1* and *vasa2*), and RNA-associated proteins (*nanos1*, *nanos2*, *piwi1*, *piwi2*)[247].



Figure 3.9: Molecular characterization of progenitor cells

Molecular profiling of the progenitor cell pool recovered from the *Elav*, *FoxL2* and *SoxC* drivers. (**top**) S- phase and G2/M cycling genes module score across all clusters (same clustering and cluster colors as in Fig. 3.8). (**bottom**) Gene expression across all clusters. Dot size represents the proportion of cells in each cluster expressing the gene.

Progenitors also display the strongest enrichment of both S-phase and G2/M cycling

modules (Fig. 3.9). In contrast to progenitors, neurons, cnidocytes and secretory cells have no cell cycle module enrichment, suggesting they are primarily post-mitotic. Epidermis epithelial cells also display a cell cycle module enrichment, consistent with cell division in a subset of epidermis cells[202,222,225].

Photoconversion-based tracking of *Piwi1*$^{+}$ progenitor differentiation dynamics

Recent work on *Nematostella* stem cells showed that *Piwi1*$^{+}$ cells can differentiate into *Piwi1*^{neg} *Elav*$^{+}$ neural cells in the body wall[202]. Here, we show that *Piwi1*$^{+}$ progenitor cells are also recovered from the broadly-neuronal *Elav>KikGR* driver . As *Elav* is highly expressed in peptidergic neurons and at low levels in progenitors (Fig. 3.8), we hypothesized that the *Elav*$^{+}$ progenitors identified here are early differentiation-committed cells. If so, progenitor cells recovered at 7 days post-photoconversion should be green, as the labeled progenitors present at day 0 would have differentiated (Fig. 3.10).

We thus evaluated whether the red and orange cell fractions were depleted in 7 day-post-photoconversion *Elav>KikGR*$^{+}$ progenitors. We find that the majority of *Elav>KikGR*$^{+}$ progenitors at 7 days (88%) are green. Accordingly, earlier timepoints (2 and 4 days) have intermediate red and orange progenitor cells fractions (Fig. 3.10). This is consistent with a scenario in which *Elav*$^{+}$ progenitors are differentiation-committed exit-bound cells (Fig. 3.10). Indeed, lowly-KikGR-expressing *Elav>KikGR*$^{+}$ morphologically undifferentiated green cells can be seen in the sub-epidermal layer by *in situ* live imaging (Fig. 3.11).

Progenitor cells express *SoxC* and *FoxL2* (Fig. 3.8). Our detection of *SoxC>KikGR*$^{+}$ and *FoxL2>KikGR*$^{+}$ non-expressing neural derivatives with overlapping class identities (Fig. 3.5) suggests a common progenitor pool (see below). However, the presence of green neurons labeled by the *SoxC>KikGR* driver but only orange and red neurons with the *FoxL2>KikGR* driver at 7 days post-photoconversion (Fig. 3.4) reveal distinct temporal expression patterns during differentiation. To better understand this neural differentiation process, we evaluated

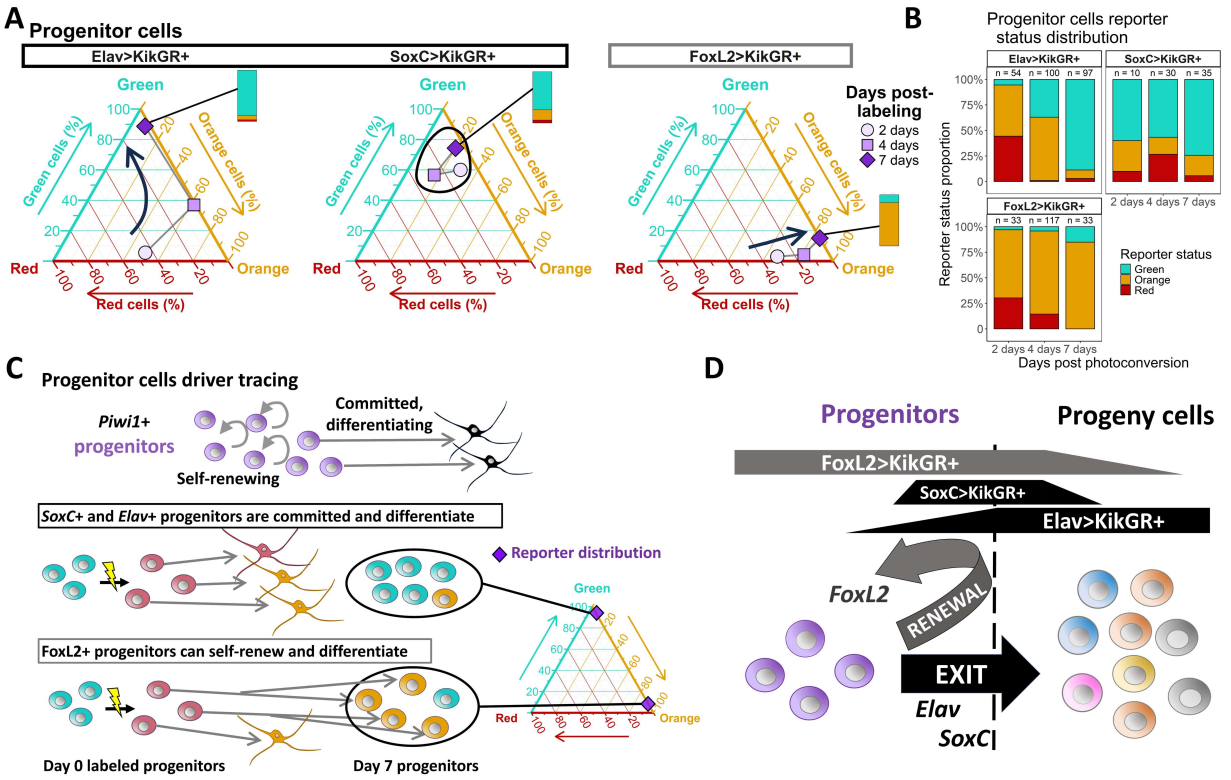

Figure 3.10: A progenitor cell pool that self-renews and differentiates

A,**B** Diverging temporal fates of progenitors. Ternary diagrams (**C**) and barplots (**D**) of the proportion of red, orange and green progenitor cells from the *Elav>KikGR*, *SoxC>KikGR* and *FoxL2>KikGR* drivers at 2,4 and 7 days post-photoconversion. **D** Schematics of *Piwi1*$^{+}$ progenitors giving rise to differentiated cells (*top*), with two expression dynamics: driver expression occurs only in committed, differentiation-bound progenitors (**middle**) or in self-renewing progenitors (**bottom**). For both scenarios, the corresponding progenitor cell color proportions at 7 days post-photoconversion are indicated on the right, with a majority of orange progenitor cells at 7 days indicative of driver expression in self-renewing progenitor cells, while a majority of green progenitor cells at 7 days is indicative of driver expression in differentiating and exit-committed progenitor cells. **E** Integrated model of progenitor cell temporal dynamics.

when progenitors express *SoxC* relative to the initiation of differentiation. Specifically, as endogenous *SoxC* is recovered at a high level in progenitor cells (Fig. 3.8), we assessed whether progenitors stably express this driver. If this were the case, using our photoconversion-based temporal recording, we would expect *SoxC>KikGR*$^{+}$ progenitors to be mostly orange cells. Instead, we find that the majority of *SoxC>KikGR*$^{+}$ progenitor cells are green cells, at 2-, 4-, and 7-day post-photoconversion (Fig. 3.10). This pattern suggests that *SoxC>KikGR*

driver initiation in progenitors is linked to a rapid differentiation initiation (Fig. 3.10).

In marked contrast with the previous differentiation-committed dynamics observed for progenitors with both the *Elav>KikGR*$^{+}$ and *SoxC>KikGR*$^{+}$ drivers, progenitors recovered from the *FoxL2>KikGR*$^{+}$ driver retain their photoconverted label and are either orange or red cells (>87% of progenitors) at 2, 4 and 7 days post-photoconversion (Fig. 3.10). This suggests a stable self-renewing pool of progenitors where the *FoxL2>KikGR* driver is continuously expressed in cells that can differentiate, while expression of *SoxC>KikGR* or *Elav>KikGR* is associated with differentiation onset. This is further supported by the high endogenous expression of *FoxL2* in progenitors (Fig. 3.9), and that, even after 7 days post-photoconversion, *FoxL2>KikGR*$^{+}$ progenitors remain overwhelmingly orange (Fig. 3.10).

Accordingly, at 7 days post-photoconversion, morphologically undifferentiated *FoxL2>KikGR*$^{+}$ orange cells are detected in the body wall (Figs. 3.6, 3.12 and 3.13), whereas undifferentiated *Elav>KikGR*$^{+}$ cells are green (Fig. 3.11), with low expression levels. Thus, the different drivers highlight that the population of progenitor cells is both self-renewing and differentiates into other cell states, consistent with a putative stem cell pool.

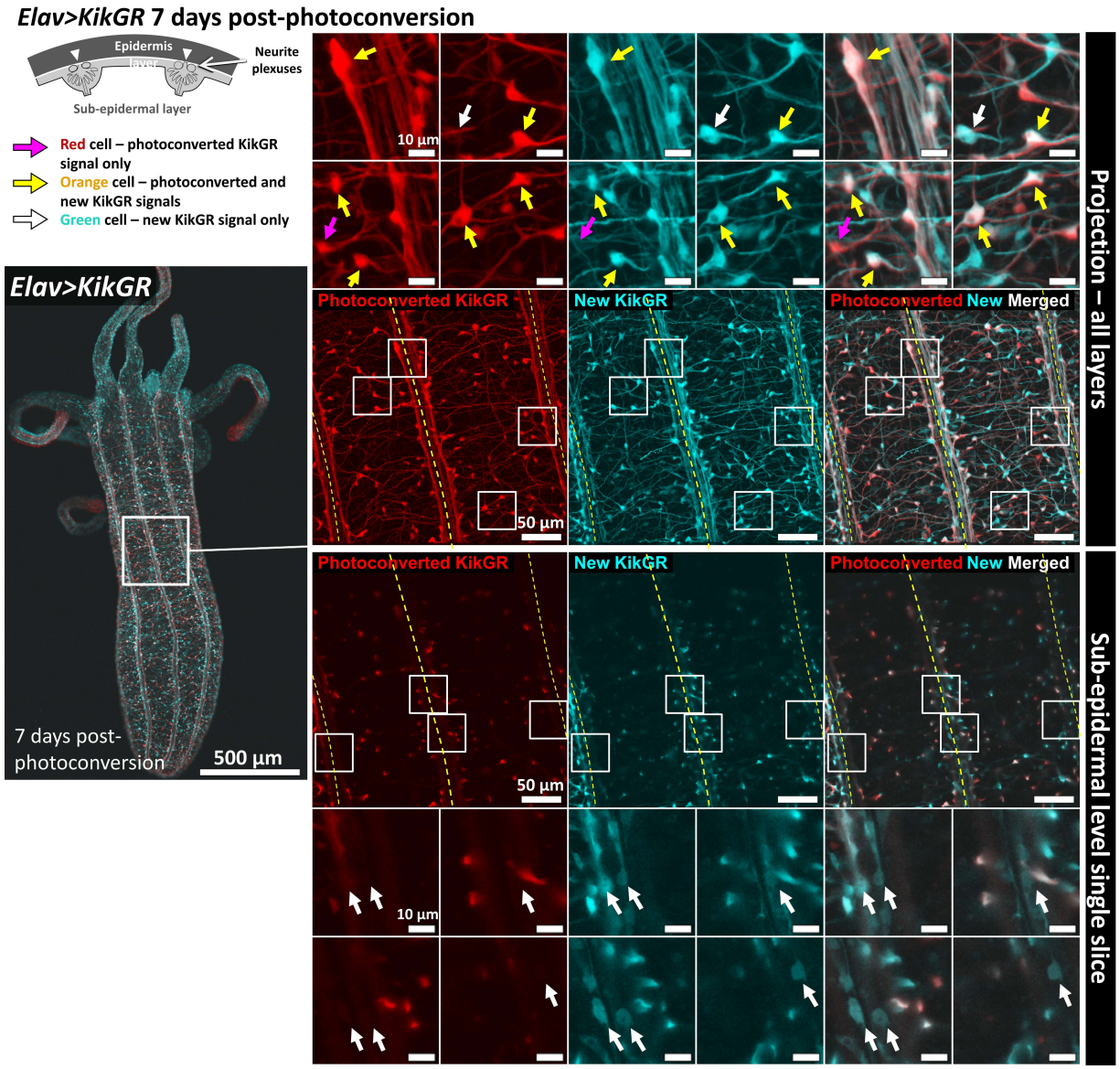

Figure 3.11: Undifferentiated *Elav>KikGR*$^+$ green cells in the body wall at 7 days post-photoconversion alongside green, orange and red neurons

In situ imaging (40X) of the body wall at 7 days post-photoconversion in *Elav>KikGR*$^+$ animals, z-stack projection (**top**) and single sub-epidermal slice (**bottom**). Neurons include green, orange and red cells, indicated by magenta, yellow and white arrows on the projection, respectively, while green while undifferentiated green cells are visible at the sub-epidermal level (white arrow in the bottom panels). The dashed yellow line marks the parietal muscle located between two neurite condensations (white arrowheads in the schematic). Scale bars are 500 μm (left), 50 μm (right), and 10 μm in the insets.

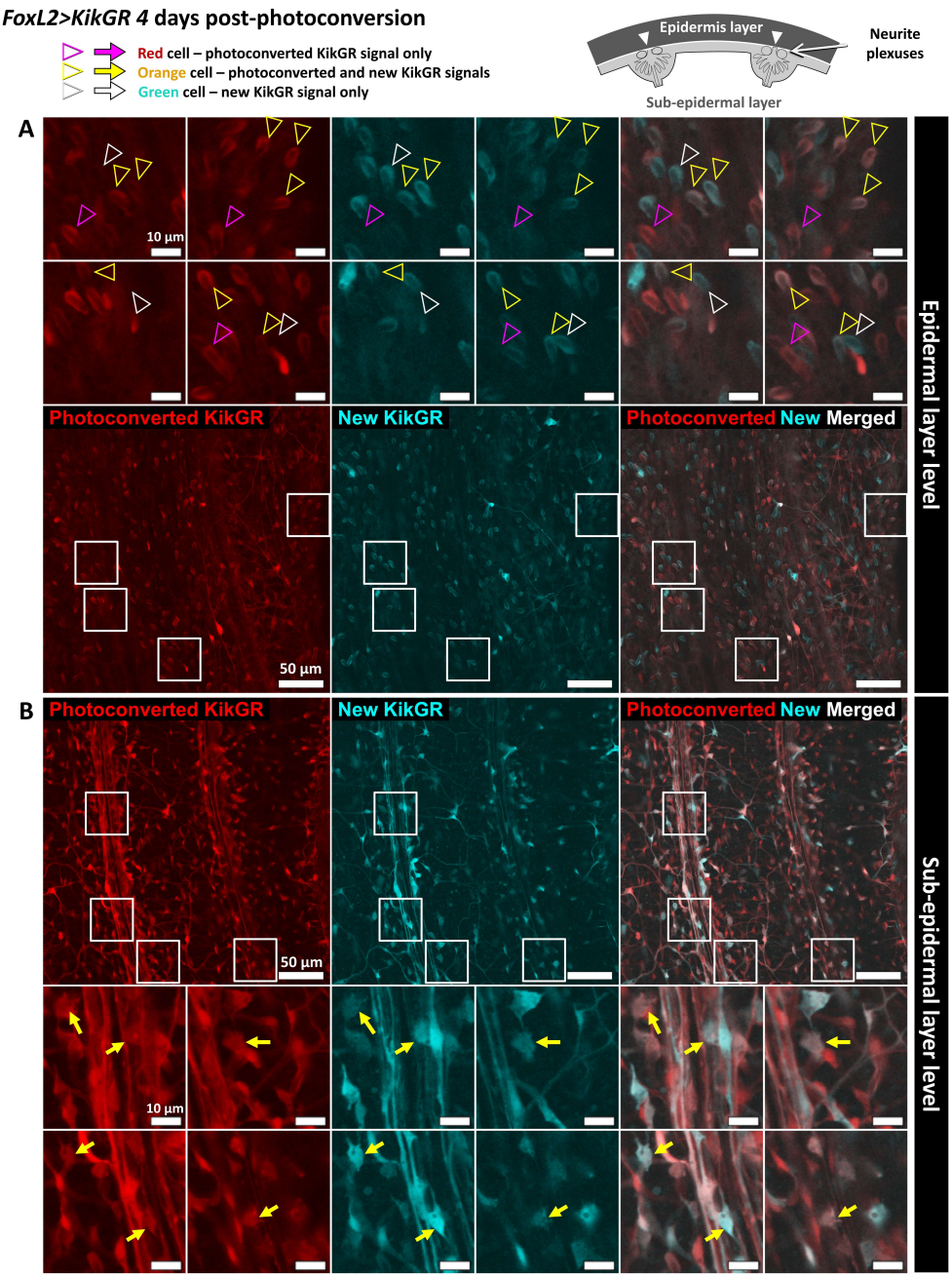

Figure 3.12: Morphologically undifferentiated *FoxL2>KikGR*$^+$ orange cells in the sup-epidermal neurite tracts at 4 days post-photoconversion

In situ imaging (40x) of 4 day post-photoconversion *FoxL2>KikGR*$^+$ animal. Body wall region optical slice at the epidermal layer level (**A**) and at the sub-epidermal layer level (**B**) with magnification insets. Mature cnidocytes present in the epidermal layer can be identified from their opaque intracellular capsule, with both orange cells (photoconverted and new KikGR signal, yellow arrowheads), red cells (only photoconverted KikGR signal, red arrowheads) and green cells (only new KikGR signal, white arrowheads). Morphologically undifferentiated orange cells can be detected at the sub-epidermal layer level (yellow arrows). Scale bar is 50μm (main panels) and 10μm (insets).

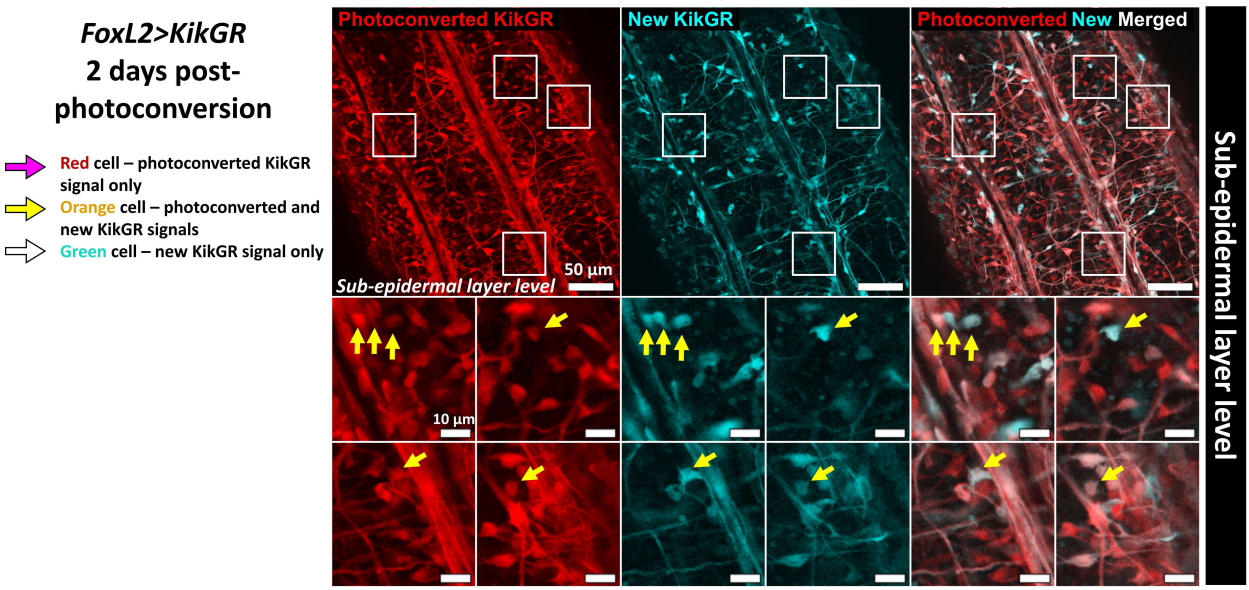

Figure 3.13: Orange undifferentiated *FoxL2>KikGR*$^{+}$ cells at 2 days post-photoconversion

In situ imaging of orange morphologically undifferentiated cells (orange cells) in the subepidermal layer level at 2 days post-photoconversion in *FoxL2>KikGR*$^{+}$ animals (yellow arrows). 40X optical slice with magnification insets. Scale bar is 500μm (whole animal overview), 50μm (main panels) and 10μm (insets).

Progenitor molecular heterogeneity and cell-class-specific precursor states

Subclustering of these progenitor cells uncovered both *Piwi1*^{high} undifferentiated as well as *Piwi1*^{low} putative differentiating states (Fig. 3.14). *Piwi1*^{low} cells also express either peptidergic neural markers, digestive secretory cell markers, or cnidocyte marker genes, respectively in subclusters 7, 9 and 8, termed "precursors". These putative precursor states express lower levels of *Piwi1* relative to the more undifferentiated states (clusters 0 to 6) (Fig. 3.14). A subset of the undifferentiated clusters display a clear cycling signature enrichment (undifferentiated clusters 5, 3 and 6), seen across *FoxL2>KikGR*$^{+}$ and *Elav>KikGR*$^{+}$ cells. The low overall number of *SoxC>KikGR*$^{+}$ cells sampled and the high fraction of putative cnidocytes precursors for this driver limit our analysis. Both undifferentiated and putative precursor cells are labeled by the *Elav>KikGR* and *FoxL2>KikGR* drivers. Putative cnidocyte precursors are absent from *Elav>KikGR*$^{+}$ cells.

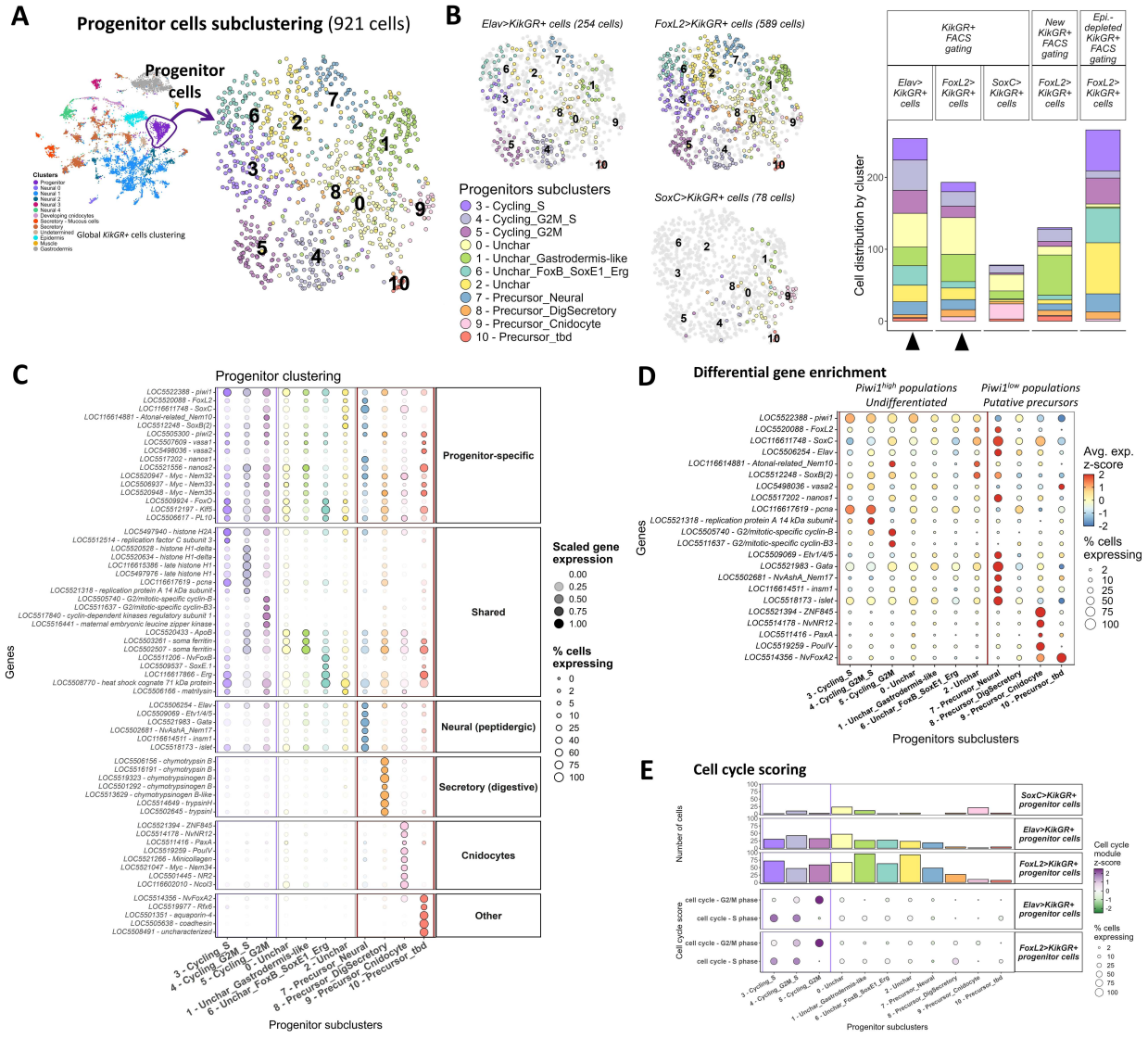

Figure 3.14: Progenitor cell reclustering reveals undifferentiated and putative differentiating states

A UMAP projection of subclustered progenitor cells (921 cells) from the global *KikGR*$^+$ dataset. **B** Distribution of progenitor subclusters by driver on the UMAP and as barplots. *Elav>KikGR*$^+$ and *FoxL2>KikGR*$^+$ progenitor cells encompass overlapping subcluster identities (black arrowheads). see Methods for FACS gating strategy. **C** Gene expression across progenitor subclusters, for progenitor-enriched markers (Fig. 3.9) and other differentially expressed genes. Dot size represents the fraction of cells expressing the gene in that cluster, and color intensity reflects normalized gene expression relative to the highest-expressing cluster. Clusters 7, 8, 9, and 10 express genes characteristic of differentiated cell types (cnidocytes, peptidergic neurons, digestive secretory cells) and may represent intermediate differentiation states (termed putative precursors here). **D** Differential gene expression (z-score) across progenitor subclusters. **E** Cell distribution by line across progenitor clusters (**top**) and cycling score across progenitor subclusters (**bottom**) for *Elav>KikGR*$^+$ and *FoxL2>KikGR*$^+$ cells. Not shown for *SoxC>KikGR*$^+$ due to the small cell numbers.

Undifferentiated *Piwi1*^{high} progenitors are nevertheless recovered across both *Elav>KikGR*$^{+}$ and *FoxL2>KikGR*$^{+}$ cells, alongside *Piwi1*^{low} peptidergic-specific precursor populations among others. However, reporter tracing demonstrates that once the *Elav>KikGR* reporter is activated, progenitor cells no longer maintain the undifferentiated state but instead proceed toward differentiation (Fig. 3.10). Thus, activation of *Elav>KikGR* in undifferentiated progenitors is specifically linked to peptidergic neuronal differentiation, consistent with the absence of cnidocyte precursors among *Elav>KikGR*$^{+}$ progenitor cells (Fig. 3.14), and the lack of mature *Elav>KikGR*$^{+}$ cnidocytes (also consistent with[38]). Overall, this uncovers a common pool of *Piwi1*$^{+}$ progenitors with a stem cell-like molecular identity that self-renews and differentiates into multiple cell types.

3.3.4 Direct differentiation of peptidergic neurons

Peptidergic neurons arise from transiently expressing *SoxC*$^{+}$ cells

We then examined the adult developmental lineage of peptidergic neurons, the largest neural class, from these *Piwi1*$^{+}$ progenitors. *Elav*$^{+}$ peptidergic neurons are recovered from all three drivers (*Elav/KikGR*, *SoxC>KikGR* and *FoxL2>KikGR*) (Neural 1 in Fig. 3.7 and Fig. 3.8)). To assess whether *Piwi1*$^{+}$/*SoxC*$^{+}$/*Elav*^{low} progenitors differentiate into peptidergic neurons, we tracked the derivatives of this population using the *FoxL2>KikGR* driver Fig. 3.15, as *FoxL2>KikGR*$^{+}$ progenitors carry the photoconverted label even at 7 days post-photoconversion. We found that over 90% of peptidergic neurons recovered from the *FoxL2* driver are either red or orange cells at 2, 4 and 7 days post-photoconversion Fig. 3.15. Thus, the presence of the photoconverted tracer label in peptidergic neurons after 7 days post photoconversion, together with the large fraction of red cells, are consistent with our proposed model of active neuron generation from *FoxL2*$^{+}$/*Piwi1*$^{+}$ progenitor cells.

As *SoxC* expression in *FoxL2*$^{+}$/*Piwi1*$^{+}$ progenitors is transient and linked to differentiation initiation, if these *Piwi1*$^{+}$/*SoxC*$^{+}$ cells develop into peptidergic neurons, we predicted

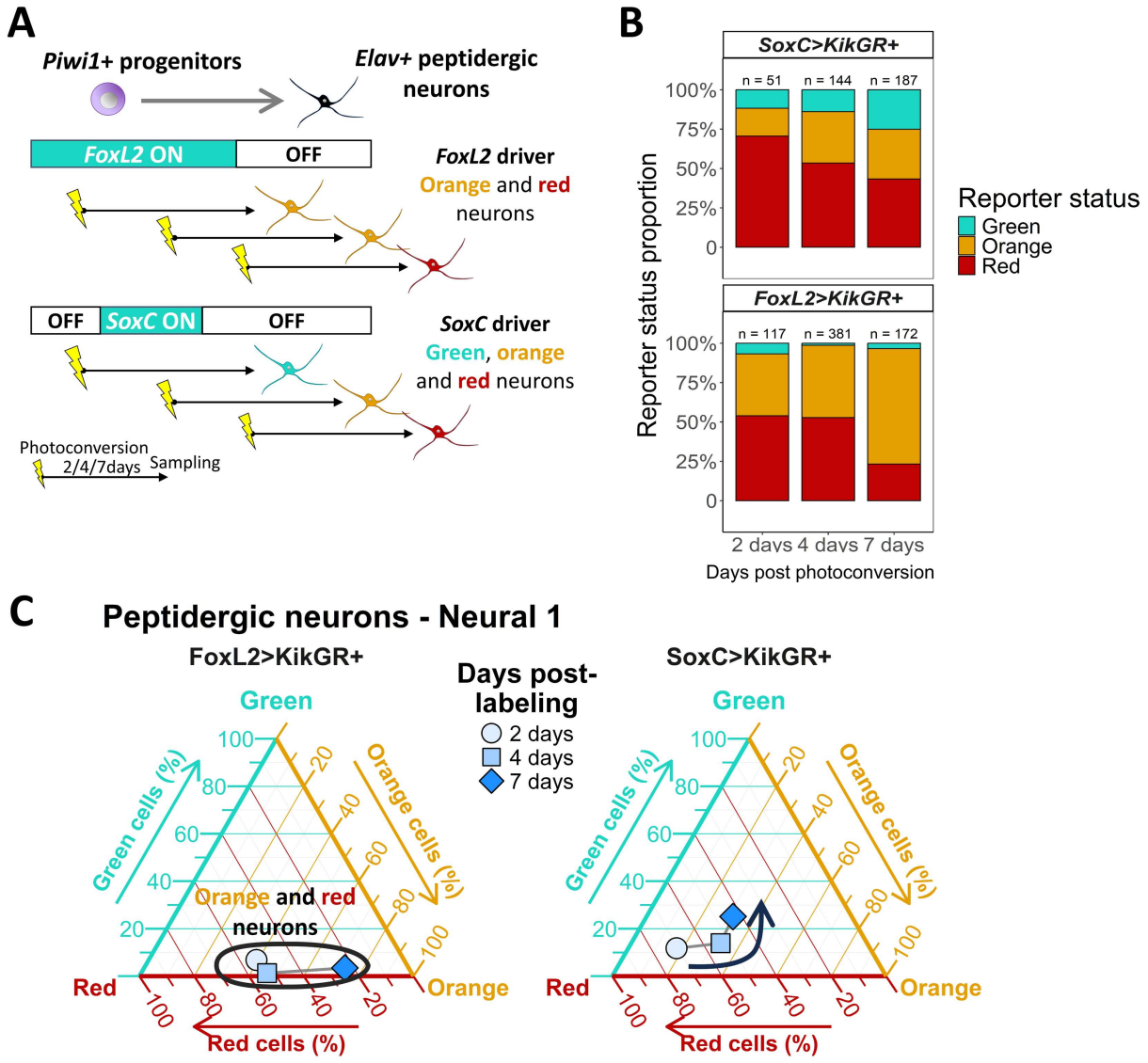

Figure 3.15: Generation of *Elav*$^+$ peptidergic neurons subtypes from *FoxL2*$^+$/*SoxC*$^+$ progenitors

A Lineage tracing of peptidergic neurons from continuously labeled *FoxL2>KikGR*$^+$ *Piwi1*$^+$ progenitors. These progenitors give rise to orange and red *FoxL2>KikGR*$^+$ peptidergic neurons at 2, 4 and 7 days post-photoconversion. Transiently expressing *SoxC>KikGR*$^+$ progenitors generate peptidergic neurons with green, orange, and red reporter states present at 2, 4 and 7 days-photoconversion, reflecting predicted differences between continuous (*FoxL2*) and transient (*SoxC*) driver expression in the peptidergic lineage. **B**,**C** Proportion of green, orange and red *Elav*$^+$ peptidergic neurons (Neural 1 class) recovered from *FoxL2>KikGR*$^+$ and *SoxC>KikGR*$^+$ cells at 2, 4 and 7 days-photoconversion.

that the peptidergic neurons recovered from the *SoxC>KikGR* driver should encompass both red cells (cells generated prior to photoconversion), and an increasing fraction of orange and green cells (newly generated cells) at later timepoints Fig. 3.16. Therefore, we examined whether we could recover both green, orange and red peptidergic neurons from the SoxC driver at all days post-photoconversion sampled Fig. 3.16. Overall, the simultaneous presence of red, orange, and green peptidergic neurons from the SoxC driver, consistent with our *in situ* nerve net imaging data Fig. 3.4, supports our model of peptidergic neurons differentiating from $Piwi1^{+}/SoxC^{+}$ progenitors, with a transient early expression of *SoxC*.

Peptidergic neuron clusters do not encompass clear intermediate states

To identify the neural lineage(s) giving rise to peptidergic neuron subtypes, we next investigated whether the distinct peptidergic neurons Fig. 3.3 are terminally differentiated, or whether some are temporal intermediates. We thus compared the distribution of the newly generated cells (green and orange cells from the *SoxC>KikGR* driver) to the standing cell pool (red cells) across peptidergic neuron clusters (Fig. 3.16). As the *SoxC* driver is transiently expressed during differentiation, intermediate cell states would be mostly orange or green cells, while terminally differentiated end states would be mostly red cells, with increasing green/orange fraction as newly generated cells accumulate (Figs. 3.16 and 3.17, scenario 1). In contrast, if clusters are distinct terminally differentiated states without clear intermediate states among them, they would share a similar green+orange to red cells proportion (Fig. 3.16, scenario 2).

Thus, the ratio of green+orange to red cells across clusters within a cellular lineage can discriminate between intermediate versus terminal states. At 7 days post-photoconversion, longer than the peptidergic maturation timeline, we found a strong linear relationship between newly and previously generated neurons across Neural 1 peptidergic subtypes (R^2= 0.92, Pearson's ρ, Fig. 3.16, scenario 2). Although the number of sampled cells for the less

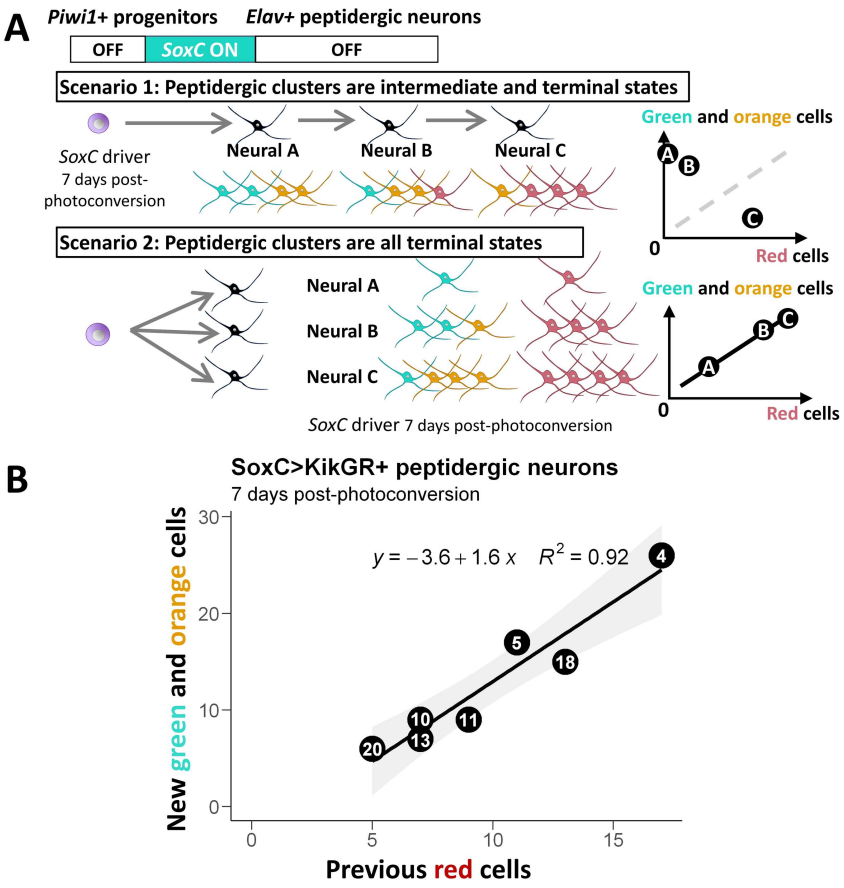

Figure 3.16: Peptidergic neuron subtypes are end states arising from *SoxC*$^+$ cells
A Peptidergic neuron development: (top, scenario 1) neuron clusters encompass both differentiation intermediates and end states or (bottom, scenario 2) neuron clusters are distinct terminally differentiated subtypes generated in parallel. *SoxC>KikGR* labels early intermediate states (green or orange cells, no red cells), and late differentiated states accumulate red cells (scenario 1). The ratio of green+orange to red cells across cell clusters distinguishes between intermediates or concurrent differentiated states within a cellular lineage. **B** Distribution of *SoxC>KikGR*$^+$ peptidergic neuron clusters at 7 days post-photoconversion shows similar green+orange to red cells proportions across clusters, supporting that they are terminally differentiated neurons (scenario 2 in **A**) without intermediates among these. See also Fig. 3.17.

abundant subtypes limits conclusive findings across all clusters, our analysis of the more abundant clusters supports peptidergic neurons as distinct terminal states, rather than intermediate states (Fig. 3.17).

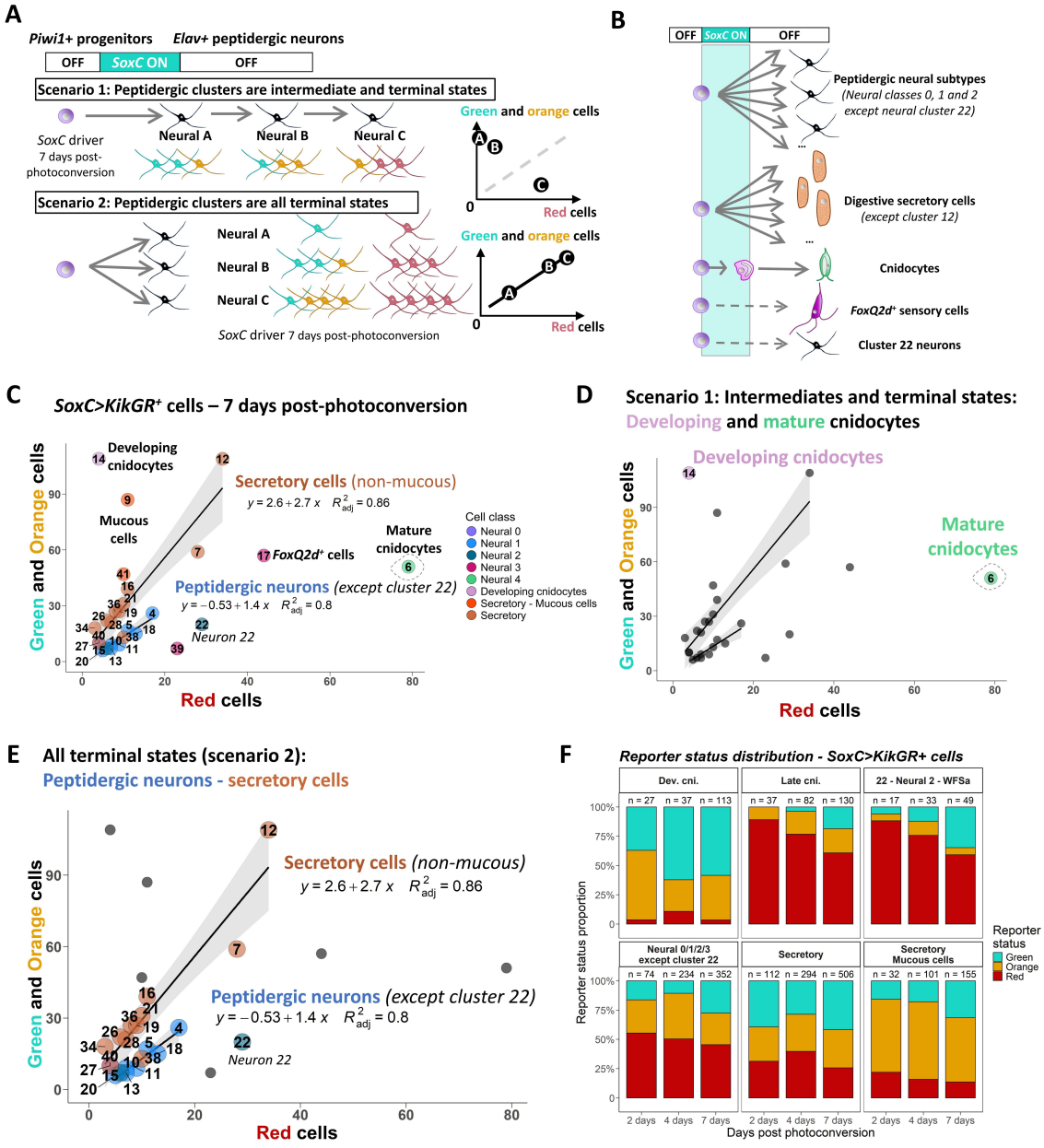

Figure 3.17: Transient *SoxC>KikGR* expression in differentiating peptidergic neurons, secretory cells, and cnidocytes

A Identification of intermediate cell states from *SoxC>KikGR* reporter status. Intermediate cell states encompass orange and green cells (scenario 1), while later states have green, orange and red cells (scenario 2). **B** Transient *SoxC* expression in early differentiating cnidocytes, peptidergic neurons, sensory cells, and secretory cells. **C-E** Ratio of newly generated (green+orange) to non-expressing (red) *SoxC>KikGR*$^{+}$ progeny cells at 7 days post-photoconversion. **D** Developing cnidocytes are an intermediate state (scenario 1). **E** Peptidergic neurons have the same green+orange to red ratio fitting scenario 2 (in blue, excluding cluster 22), as well as digestive secretory cells (in brown) (**E**). **F** Reporter status distribution across neural and secretory cells types.

A restricted peptidergic neuron-committed lineage

Overall, our data are consistent with closely related peptidergic subtypes being generated from the *Piwi1*$^{+}$ multipotent population, with their relative ratio fixed at an early peptidergic-committed stage as supported by multiple independent lines of evidence. First, the proportional ratio of old to new (red cells to green+orange cells) *SoxC>KikGR*$^{+}$ peptidergic neurons (scenario 2 in Figs. 3.16 and 3.17) is specific to these cells and not seen across digestive secretory cells, which display a separate proportional relationship (Fig. 3.17). Furthermore, if there is a shared peptidergic-committed lineage, as the *FoxL2>KikGR*, *SoxC>KikGR* and *Elav>KikGR* drivers label the *Piwi1*$^{+}$ source population, and thus also its derivatives, then the relative abundance between peptidergic subtypes would be constrained by this shared lineage and would thus correlate across drivers, but not for cells generated through distinct cellular lineages. Indeed, we find that the relative abundance of peptidergic cells subtypes is correlated across drivers and that the relationship is specific to peptidergic neurons (excluding neural cluster 22, see below); while *FoxQ2d*$^{+}$ sensory cells, cnidocytes and secretory cells do not share the same relationship (Fig. 3.17). This early peptidergic-specific lineage restriction is further independently supported by reanalysis of RNA-Seq data from a *Prdm14d>GFP* reporter line, which labels both *Prdm14d*$^{+}$ neurons (Neural 0 in Fig. 3.8), and uncharacterized progenitors at low levels[181]. Genes enriched in *Prdm14d>GFP*$^{+}$ cells are expressed here in *Prdm14d*$^{+}$ neurons, as well as across non-*Prdm14d*-expressing peptidergic subtypes; and not in secretory cells or other neurons (Fig. 3.18), supporting that peptidergic-committed *Piwi1*$^{+}$/*Prdm14d*^{low} progenitors are restricted to peptidergic neural fates. Accordingly, genes enriched in *Prdm14d>GFP*^{neg} cells are expressed across secretory cells and other neural cells including cluster 22 neurons, cnidocytes (developing and mature) and *FoxQ2d*$^{+}$ sensory cells, further supporting that these cells are not part of the shared peptidergic lineage.

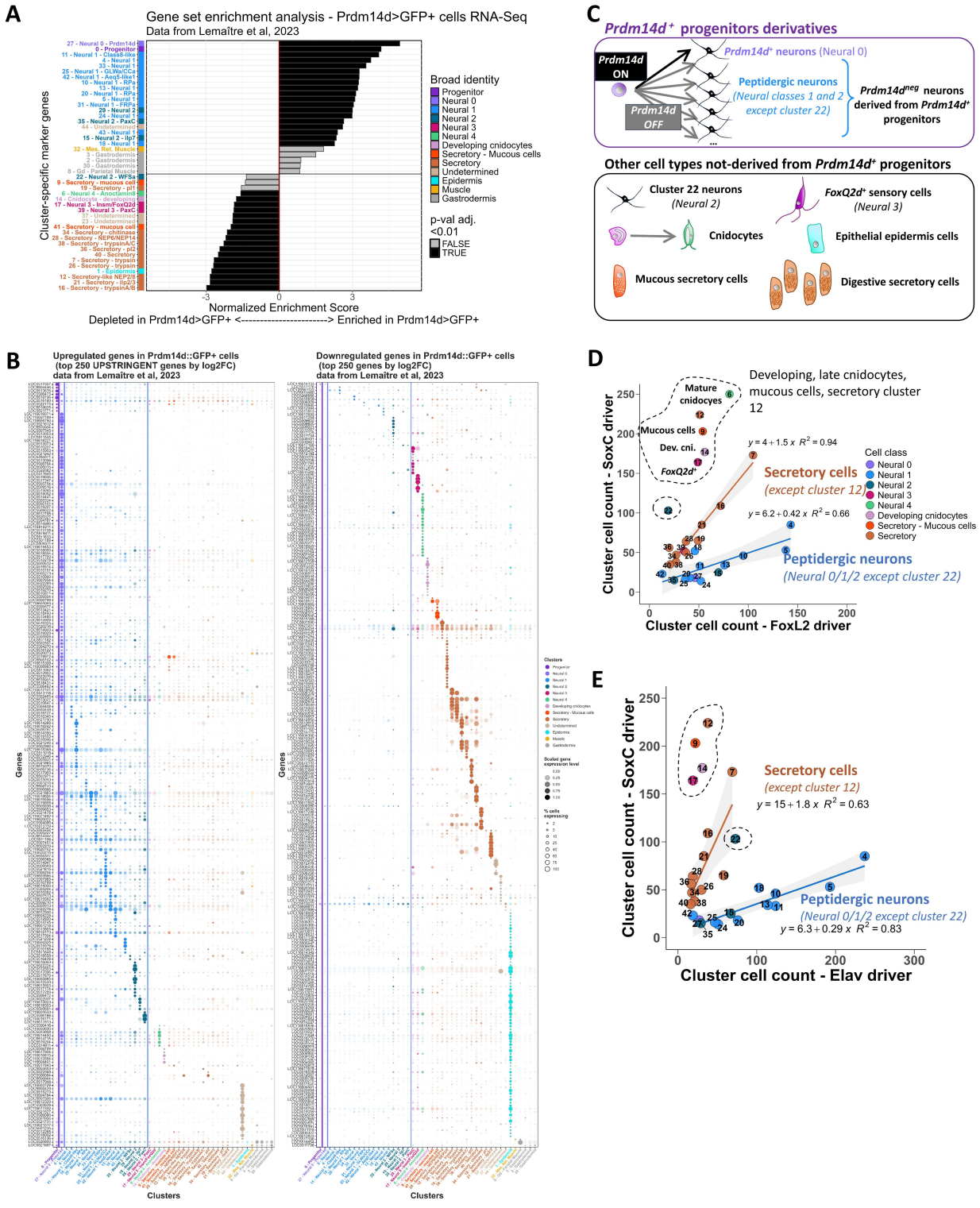

Figure 3.18: Reanalysis of *Prdm14d* driver RNA-Seq data supports a peptidergic-restricted progenitor lineage

Figure 3.18: Reanalysis of *Prdm14d* driver RNA-Seq data supports a peptidergic-restricted progenitor lineage *(continued)*
A Gene set enrichment analysis (GSEA) of *Prdm14d>GFP*$^{+}$ RNA-Seq data[181] using cluster-specific marker genes. Genes enriched in *Prdm14d>GFP*$^{+}$ cells (Normalized Enrichment Score: NES >0) overlap with marker genes from peptidergic clusters (*Prdm14d*^{neg}), and genes depleted in *Prdm14d>GFP*$^{+}$ cells (NES <0) overlap with markers of other differentiated cell types, including secretory cells and cnidocytes. **B** Expression of the top 250 upregulated (***left***) and top 250 downregulated genes (***right***) from *Prdm14d>GFP*$^{+}$ cells across clusters. Upregulated genes are expressed in progenitors (*Prdm14d*$^{+}$) and in both *Prdm14d*$^{+}$ peptidergic neurons (Neural 0) and *Prdm14d*^{neg} peptidergic neurons (Neural 1 and 2). Genes downregulated in *Prdm14d>GFP*$^{+}$ cells relative to *Prdm14d>GFP*^{neg} cells are enriched in other differentiated cell types, including cnidocytes, secretory cells, epidermal epithelial cells. These findings support a *Prdm14d*$^{+}$ progenitor lineage restricted to peptidergic neurons. **C** Schematic of cell types derived from *Prdm14d*$^{+}$/*Piwi1*$^{+}$ progenitors (top), including *Prdm14d*$^{+}$ peptidergic neurons (Neural 0) and *Prdm14d*^{neg} peptidergic neurons (Neural 1 and 2), contrasted with differentiated lineages (bottom, cluster 22 neurons, cnidocytes, digestive secretory cells) that do not arise from this lineage. *Prdm14d* is endogenously expressed in *Piwi1*$^{+}$ progenitors and *Prdm14d*$^{+}$ peptidergic neurons (Neural 0), but absent from Neural 1 and 2 peptidergic neurons (see Fig. 3.8). Neural 1 markers are expressed in *Prdm14d>GFP*$^{+}$ cells (see **A** and **C**). **D-E** Distribution of cell clusters across driver lines. **D** *SoxC>KikGR*$^{+}$ and *FoxL2>KikGR*$^{+}$ cells. **E** *SoxC>KikGR*$^{+}$ and *Elav>KikGR*$^{+}$ cells. Peptidergic neurons (blue) and digestive secretory cells (orange) are proportionally represented across drivers. Other neural and secretory cell types that deviate from this proportionality are outlined by dashed lines, including cnidocytes, mucous cells, neural cluster 22 and *FoxQ2d*$^{+}$ sensory cells.

And digestive secretory cells (excluding cluster 12) also display correlated abundance across drivers (Fig. 3.17), with a proportional ratio distinct from that of peptidergic neurons, further supporting a separate digestive-secretory-cell specific lineage.

Taken together, the data from multiple experimental lines of evidence supports a model in which direct rapid generation of distinct *Elav*$^{+}$ peptidergic neuron subtypes from peptidergic-restricted *Piwi1*$^{+}$/*SoxC*$^{+}$/*Elav*^{low} cells underlies homeostatic nerve net scaling (Fig. 3.19).

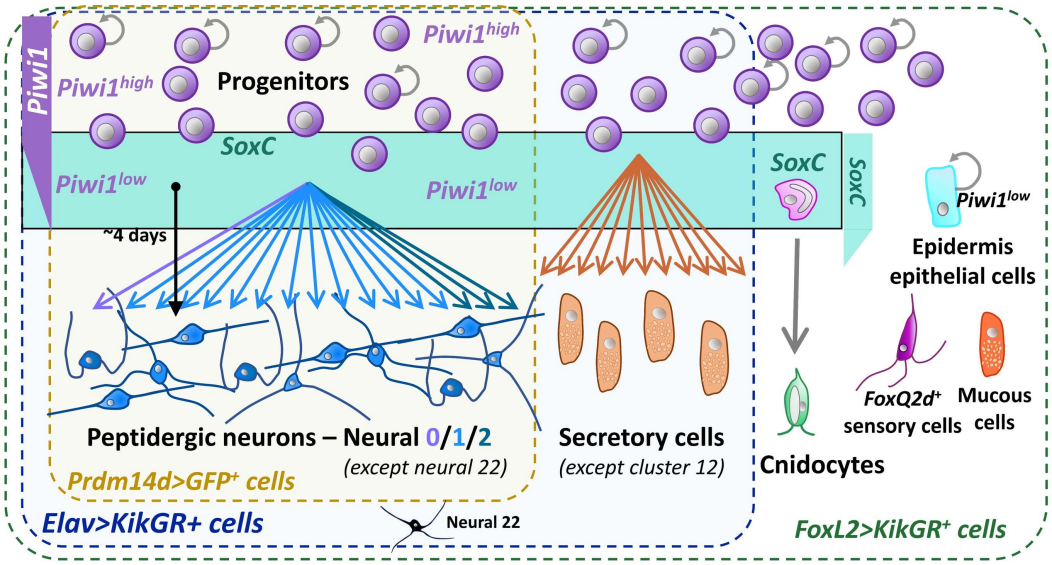

Figure 3.19: Integrated model of adult neurogenesis

Peptidergic-committed *Piwi1*$^{+}$ multipotent cells differentiate into all peptidergic subtypes within 4 days in adults. Elavlow/Piwi1+ progenitors give rise to *Elav*^{neg}/*Piwi1*^{neg} digestive secretory cells.

3.3.5 Two-step transcriptional maturation of cnidocytes

Sea anemone neural cells encompass peptidergic neurons, putative sensory cells, and cnidocytes, cnidarian-restricted mechano-sensory neural cells[47,180,251]. Having addressed the progression of peptidergic *Elav*$^{+}$ neurons, from *Piwi1*$^{+}$/*SoxC*$^{+}$/*FoxL2*$^{+}$/*Elav*^{low} progenitors to differentiated cells, we examined whether neurons and cnidocytes share a similar differentiation trajectory.

We find that cnidocytes, like peptidergic neurons, develop from transiently-expressing *SoxC*$^{+}$ cells (Figs. 3.20 to 3.22). Indeed, only orange and green developing *SoxC>KikGR*$^{+}$ cnidocytes are detected *in situ* (Figs. 3.21 and 3.22) and in the scRNA-Seq data at 2, 4 and 7 days post-photoconversion (Fig. 3.20). Mature cnidocytes do not actively express *SoxC* (Fig. 3.8). Accordingly, we find red mature *SoxC>KikGR*$^{+}$ cnidocytes at 2 days post-photoconversion, with newly generated orange and green mature cnidocytes appearing at 4 and 7 days, respectively (Figs. 3.20 to 3.22).

Thus, both peptidergic neurons and cnidocytes transiently express *SoxC* during early

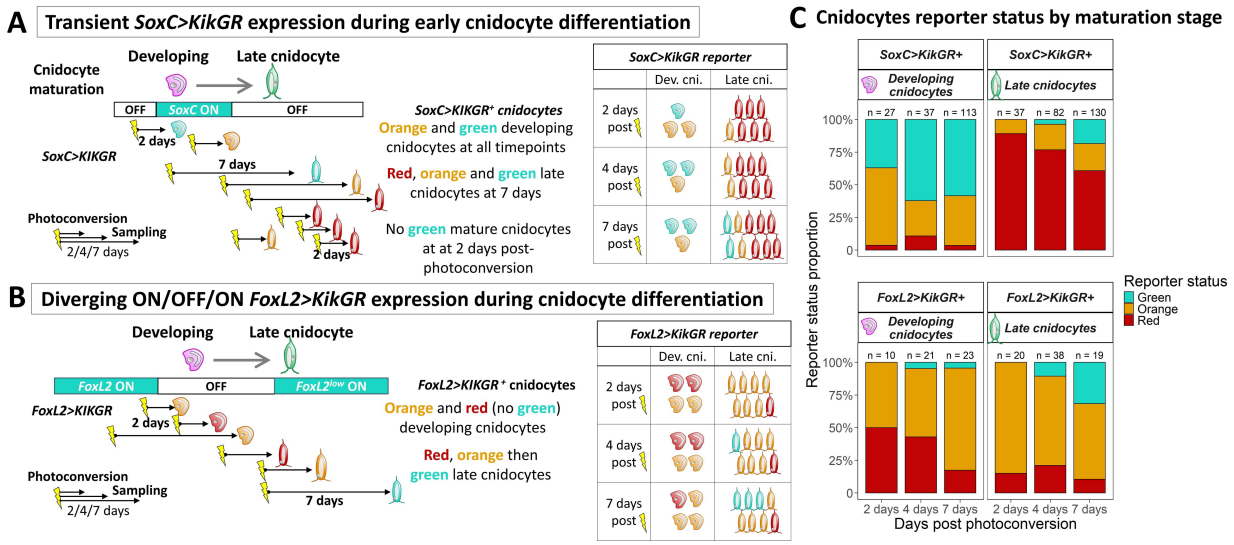

Figure 3.20: Cnidocytes undergo a biphasic maturation

A Transient *SoxC>KikGR* expression during early cnidocyte differentiation and predicted reporter status (green, orange or red cell) of developing and mature cnidocytes at 2, 4 and 7 days post-photoconversion. Developing cnidocytes remain green or orange at all timepoints. Data in **C** and independently validated by *in situ* live imaging in Figs. 3.21 and 3.22. **B** Dynamic ON/OFF/ON *FoxL2>KikGR* expression during cnidocyte differentiation and predicted reporter status at 2, 4 and 7 days post-photoconversion. *FoxL2>KikGR* is active in progenitors, downregulated during early cnidocyte maturation and reactivated at later stages. Data shown in **C** and shown by *in situ* live imaging in Fig. 3.6. **C** Distribution of reporter states for *SoxC>KikGR*$^{+}$ and *FoxL2>KikGR*$^{+}$ cnidocytes during differentiation. Cell numbers are indicated above bars.

differentiation. Furthermore, while developing cnidocytes do not express *FoxL2* (Fig. 3.8), they arise from undetermined *FoxL2>KikGR*$^{+}$ cells, as evidenced by the presence of red and orange developing cnidocytes at 2, 4, and 7 days post-photoconversion (Figs. 3.4, 3.6, 3.12 and 3.13). We also note that *FoxL2>KikGR* is secondarily re-activated in mature cnidocytes, consistently with the presence of green mature cnidocytes at 4 and 7 days post-photoconversion alongside the majority of red and orange ones (Figs. 3.4, 3.6, 3.12 and 3.20).

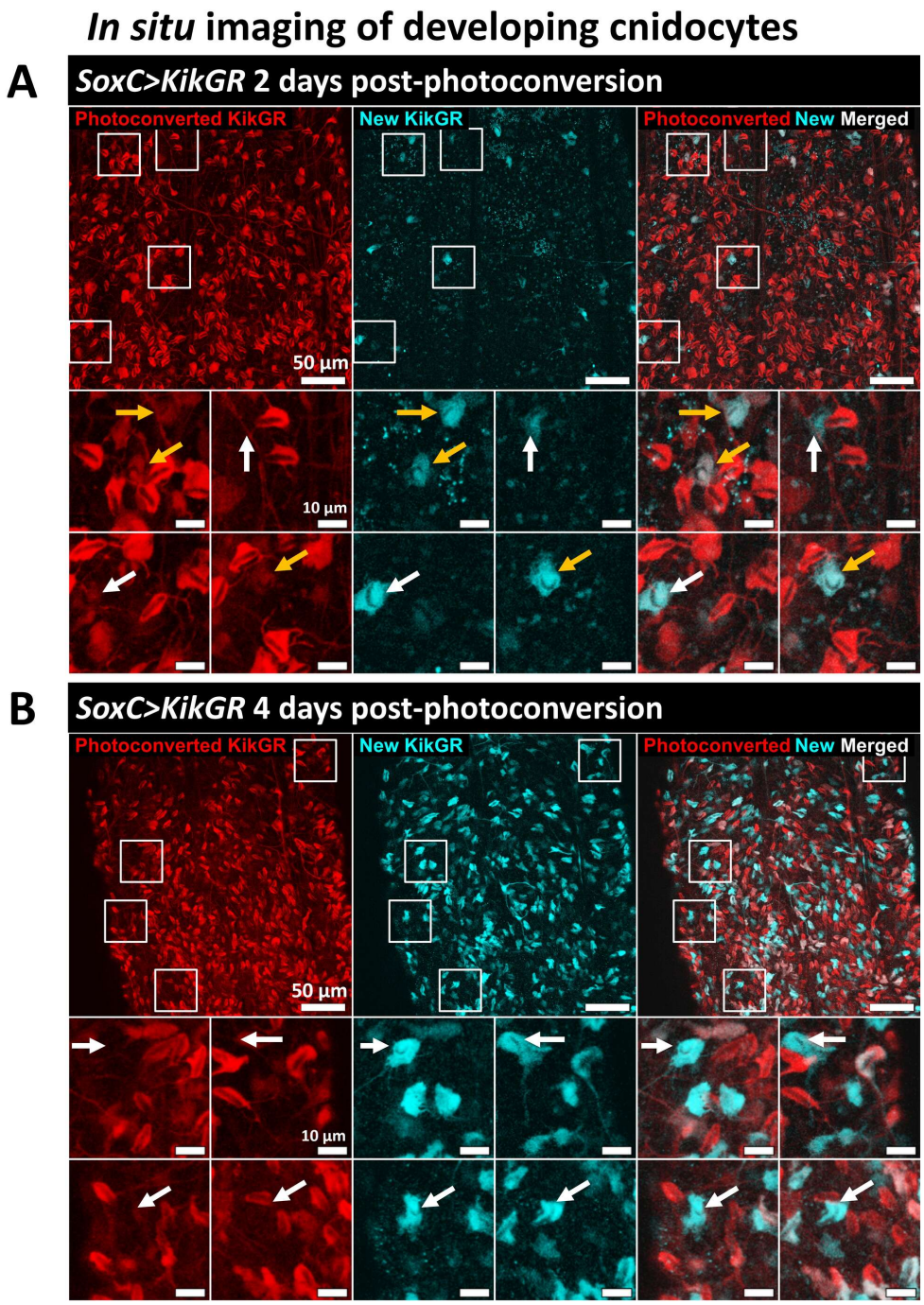

Figure 3.21: Developing cnidocytes transiently express *SoxC>KikGR* at 2 and 4 days post-photoconversion

A-C *In situ* body wall imaging of *SoxC>KikGR*$^{+}$ developing and mature cnidocytes at 2 and 4 day post-photoconversion. 40X projection with insets showing the green developing cnidocytes (white arrows) at 2 and 4 days post-photoconversion, as well as orange developing cnidocytes (orange arrows). Scale bar is 50μm and 10μm in insets. Red developing cnidocytes are absent at all times post-photoconversion while red mature cnidocytes are always present.

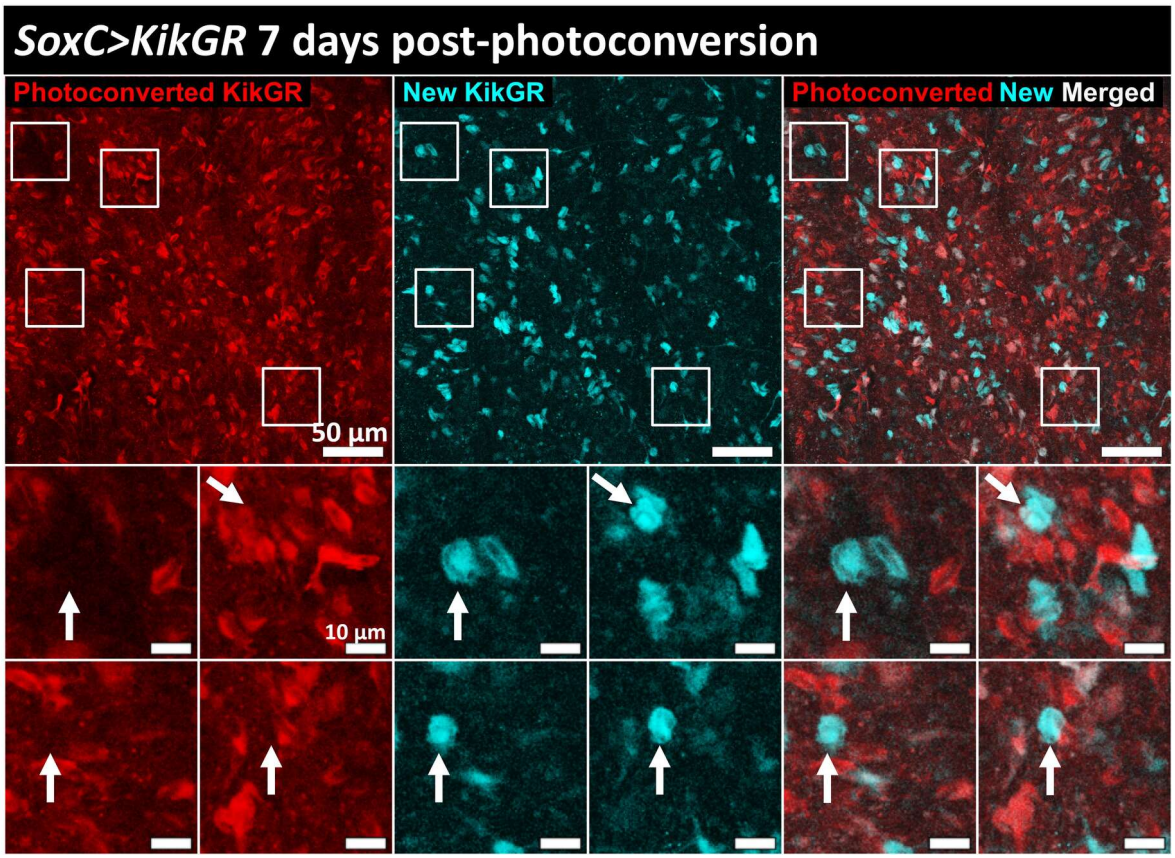

Figure 3.22: Green *SoxC>KikGR*$^+$ developing and late cnidocytes at 7 days post-PC

A-C *In situ* body wall imaging of *SoxC>KikGR*$^+$ developing and mature cnidocytes at 7 day post-photoconversion. 40X projection with insets showing the green developing cnidocytes (white arrows) at 7 days post-photoconversion. Scale bar is 50μm and 10μm in insets. Red developing cnidocytes are absent at all times post-photoconversion while red mature cnidocytes are always present.

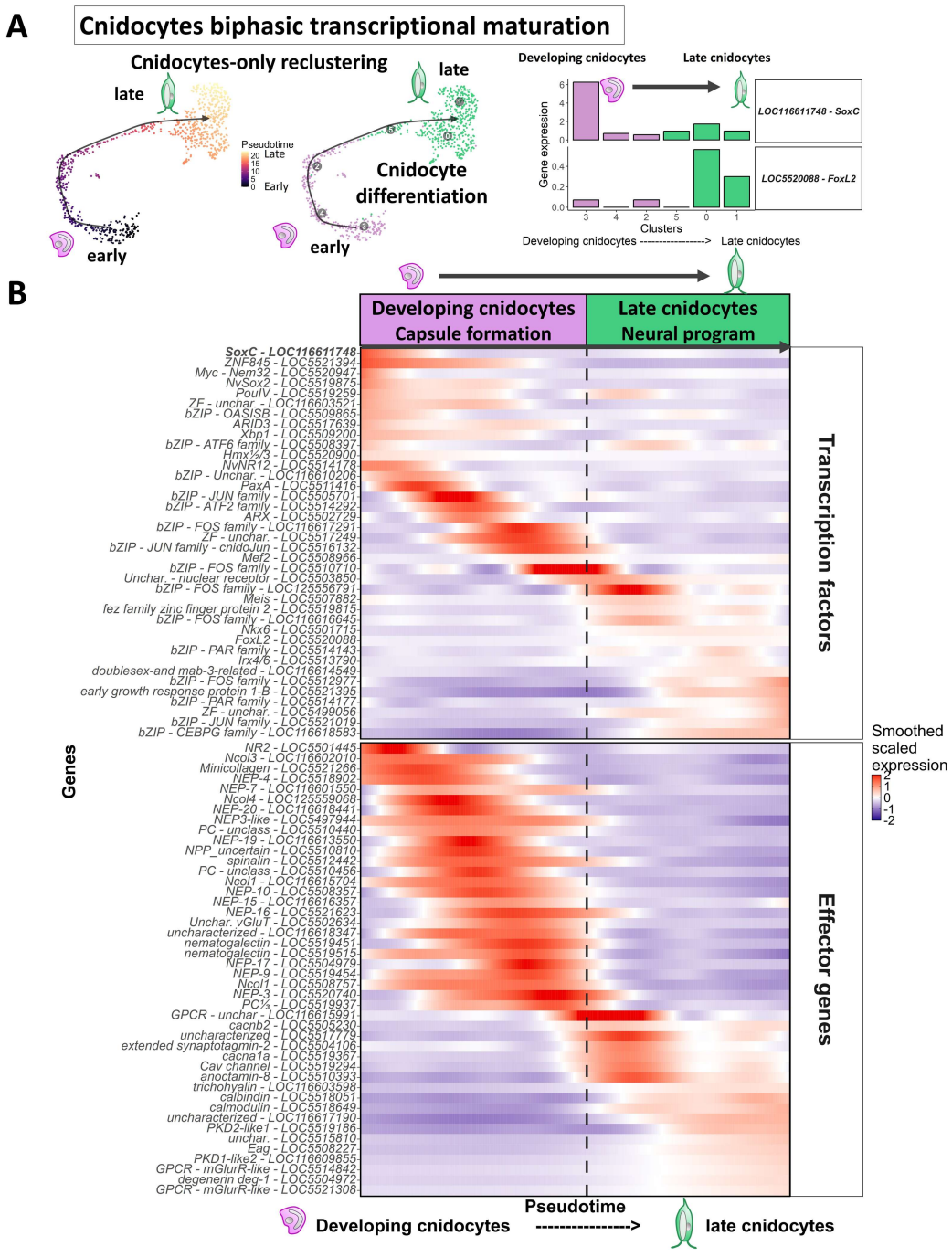

Figure 3.23: Biphasic transcriptional maturation of cnidocytes with a primary capsule program followed by a secondary neural program delayed after capsule formation

A (***left***) Cnidocyte-specific UMAP projection with inferred pseudotime. (***right***) *SoxC* is highly expressed at early differentiation stages while *FoxL2* is not expressed and shows weak re-expression in late cnidocytes. **B** Gene expression dynamics across pseudotime highlight an initial capsule formation program (*ZNF845*, *NvNR12*, *PaxA*, *PouIV*), that is downregulated as a secondary program emerges, marked by distinct effector and transcription factors (several bZIP, *nkx6*, *FoxJ1*, a *Myc*, *PouIV*, and an uncharacterized *Sox* gene).

Thus, while the cellular origin(s) of cells that give rise to developing cnidocytes is unclear, with isolated developing cnidocytes present at the base of the epidermis layer in the body wall, these cells also develop from *FoxL2*$^+$ cells, and developing cnidocytes transiently express *SoxC* during their earliest stage of maturation Fig. 3.23. We also note that a distinct subset of progenitor cells show specific expression of known cnidocyte regulators ("cnidocytes precursors" in Fig. 3.14), and could indicate that *Piwi1*$^+$ progenitors also give rise to developing cnidocytes.

However, while peptidergic neurons (Neural 0/1/2 in Figs. 3.5 and 3.8) are recovered across all drivers (*FoxL2>KikGR*, *Elav>KikGR*, *SoxC>KikGR*), the late cnidocytes (Neural 4 class) are not recovered from *Elav>KikGR*$^+$ sorted cells (Fig. 3.5, also in[38]). Thus, *SoxC* is transiently expressed in developing cnidocytes and separately in precursors of peptidergic neurons, and our findings are incompatible with a shared *SoxC*$^+$ neural-specific bipotent intermediate state (see also Fig. 3.18).

We then investigated the transcriptional signatures of neural fate acquisition in peptidergic neurons and cnidocytes, two distinct cellular lineages. Cnidocytes mature through two sequential transcriptional phases Fig. 3.23. Initially, during the developing cnidocyte stage, cnidarian-specific genes are used to build their specialized capsule organelle (intracellular harpoon-like structure). Subsequently, developing cnidocytes initiate a neural transcriptional program, which includes the expression of neural effector genes, such as the *Eag* (Ether-à-go-go family) voltage-gated K^+ channel[250] and a voltage-dependent L-type calcium channel subunit beta-2 (*cacnb2*)[251]. This biphasic cnidocyte effector program is characterized by a shift in transcription factors usage. Early in development, cnidocytes express *PaxA*, *cnido-Jun*, *NR12*, *ZNF845* and other zinc finger and basic leucine zipper (bZIP) transcription factors, which are associated with capsule formation[136,188,194]. In contrast, the secondary stage is marked by a distinct regulatory signature defined by expression of *nk6*, uncharacterized *Sox* and *Myc* genes and additional bZIP genes Fig. 3.23, indicative of the onset of

the neural program.

Thus, cnidocytes undergo two distinct transcriptional phases, with the sequential deployment of a cnidarian-specific capsule program, followed by a neural program. This two-phase transcriptional process is supported by reanalysis of RNA-Seq data from a *PouIV* knockout (KO) mutant[120], in which immature cnidocytes accumulate. Thus, genes upregulated in *PouIV* KO juvenile polyps correspond to those expressed in developing cnidocytes whereas *PouIV* KO downregulated genes overlap with those expressed in both developing and late cnidocytes (Fig. 3.24). These findings support a biphasic transcriptional maturation and suggest that when the developing cnidocyte program is blocked, the secondary neural program is not initiated. Thus, this sequential maturation is a two-step process with a clear transcriptional intermediate program that is deployed, then cleared prior to the initiation of the neural program, revealing a different neurogenic strategy from that of peptidergic neurons.

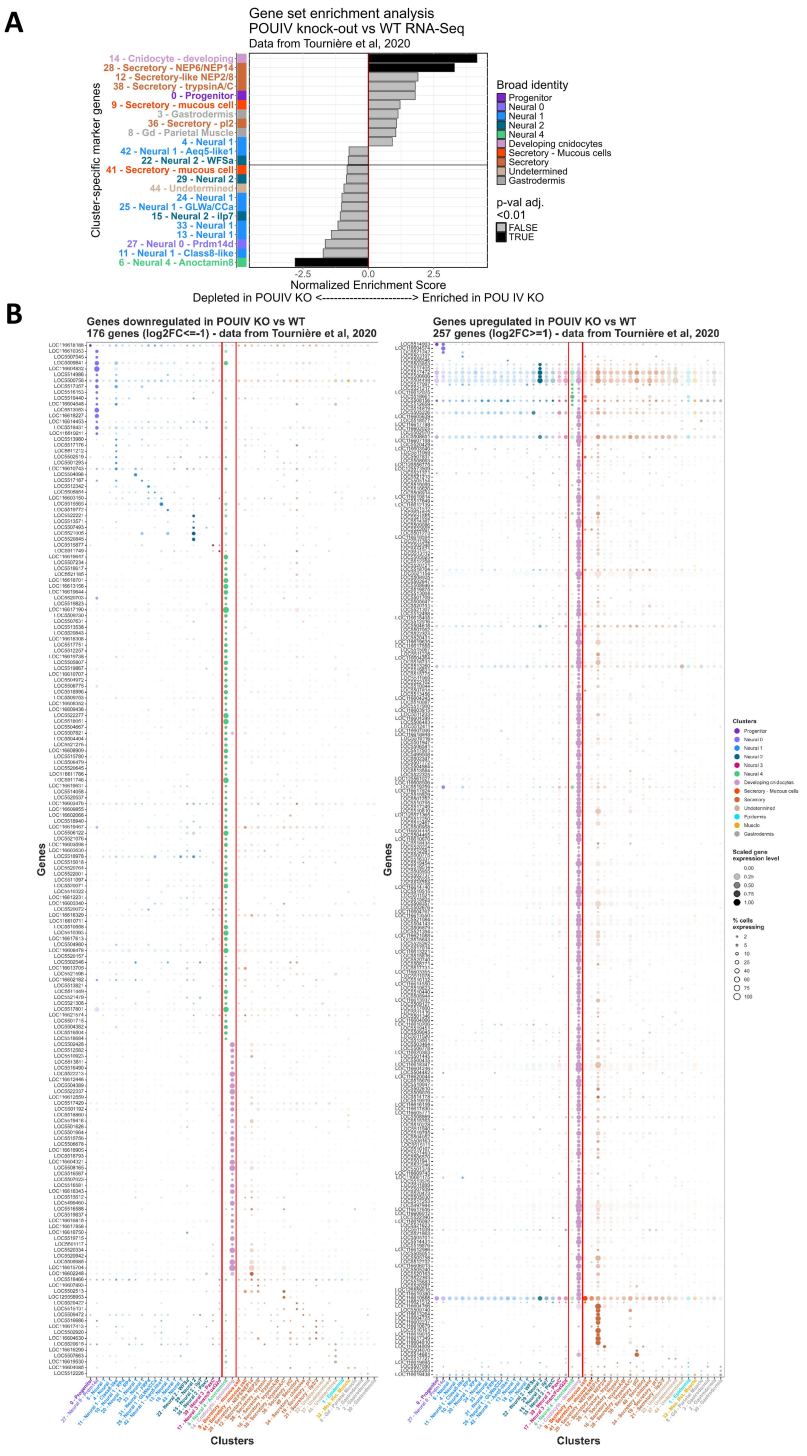

Figure 3.24: Reanalysis of a *PouIV* KO RNA-Seq dataset supports a biphasic transcriptional maturation of cnidocytes

A Gene Set Enrichment Analysis (GSEA) of cluster marker genes in *PouIV* KO RNA-Seq data[120]. **B** Expression of top downregulated (log2FC≤-1, left) and upregulated (log2FC≥1, right) genes in the *PouIV* KO RNA-Seq data across our global KikGR$^+$ clustering (as in Fig. 3.8).

3.3.6 Neural cells arising from different lineages deploy distinct TF-effector gene modules

Since mature cnidocytes and peptidergic neurons acquire their neural identities through distinct processes, we further examined whether they converge on a shared neural transcriptional signature. We find that the expression profiles of neural effector genes are distinct (Fig. 3.25). Neuropeptide precursors, processing enzymes and neuropeptide GPCRs are expressed in peptidergic neurons but not in late cnidocytes, suggesting that peptidergic transmission is not utilized by cnidocytes. Other neural-expressed GPCRs and ion channels are specific to either late cnidocytes or peptidergic neurons, with only a limited overlap between the two. Overall, these results demonstrate that mature cnidocytes and peptidergic neurons have separate neural effector expression profiles.

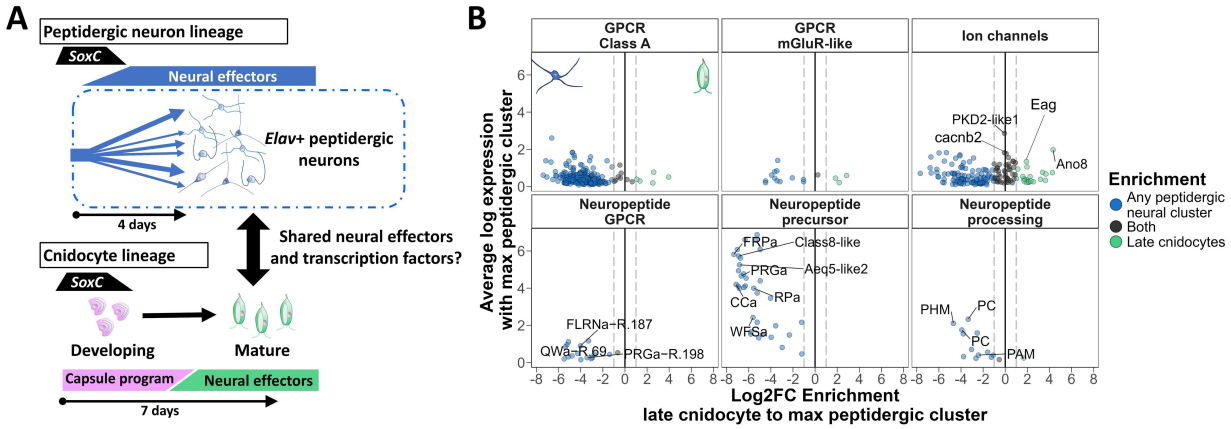

Figure 3.25: The two neural lineages deploy distinct neural effector genes.
A The two main classes of neural cells: *Elav*$^{+}$ peptidergic neurons (**top**) and cnidocytes (**bottom**) acquire their mature neural identities through distinct cellular trajectories with an initial transient expression of *SoxC*. **B** Neural effector differential expression between *Elav*$^{+}$ peptidergic clusters and late cnidocytes. Average expression log2 fold-change enrichment between the late cnidocyte cluster and the maximum expressing peptidergic cluster. The dashed line represents a 2-fold change in expression levels between the peptidergic cluster with the highest expression and late cnidocytes.

Despite the distinct differentiation trajectories and effector profiles of these neural cells, we examined whether they shared pan-neural transcription factors. Using weighted gene co-expression network analysis (WGCNA)[252] to generate modules of genes whose expres-

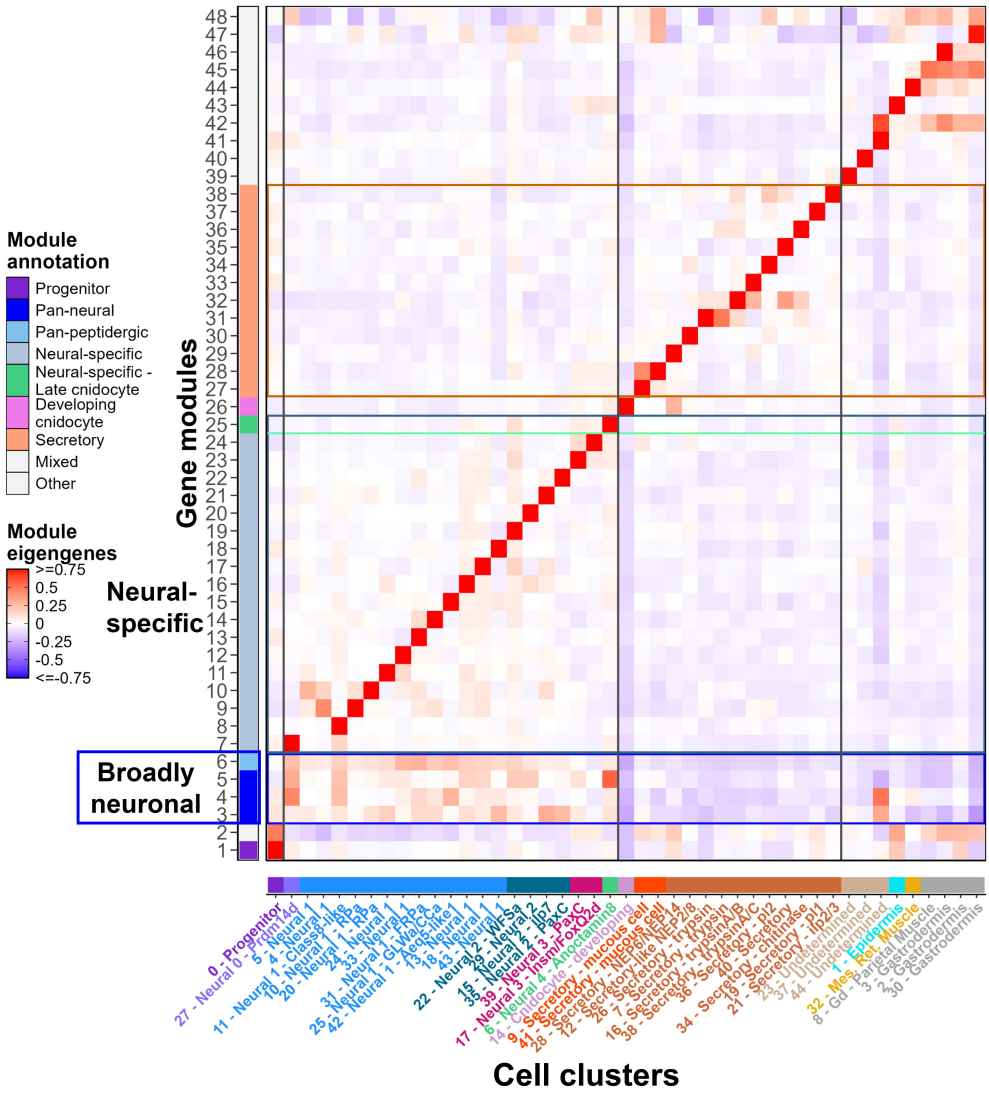

Figure 3.26: Peptidergic neurons and cnidocytes deploy distinct TF-effector gene modules

WGCNA gene modules eigengene value across cell clusters. Gene modules (rows) are annotated as neural-specific, broadly neural or secretory-specific. Cell clusters (columns) colors as in Fig. 3.7. Peptidergic neurons share gene module 6, and late cnidocytes share gene module 5 with the other neurons; while developing cnidocytes use a distinct gene program (module 26).

sion is highly correlated, we identify 48 cell-type specific and broader class-specific modules across all cell types (Fig. 3.26). Each neural identity combines a cell-type specific module (modules 7-21 and 25) with shared neural module(s) (modules 3-6) (Fig. 3.26). This is specific to neural cells, and contrasts with the lack of a shared pan-secretory module for the

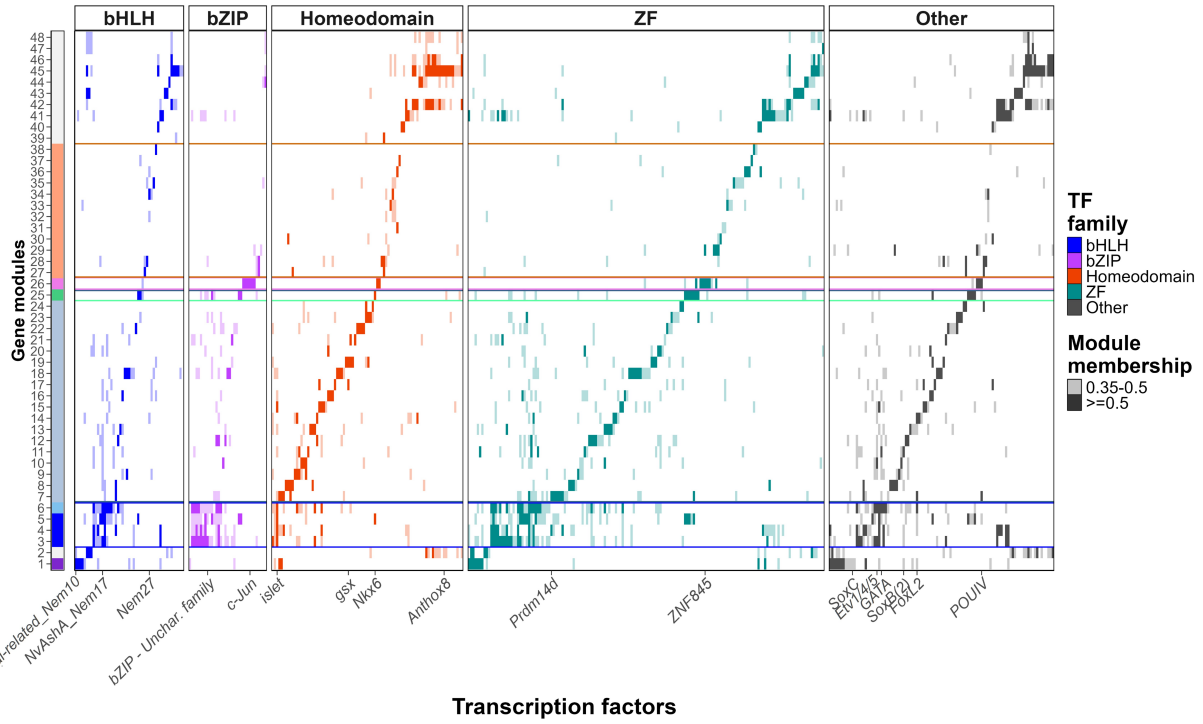

Figure 3.27: Peptidergic neurons and cnidocytes use distinct TFs with bHLH and bZIP factors for broad neural identity and a subtype-specific homeodomain code.

Transcription factor membership across WGCNA modules (same rows and module identifiers as in Fig. 3.26) for bHLH, bZIP, homeodomain, zinc finger and other transcription factor families. Each transcription factor is a column, with low (semi-opaque, membership >=0.35) or high (opaque, membership >=0.5) module membership grades to highlight TF associated to multiple modules. Horizontal lines demarcate WGCNA modules associated with broad or restricted neural identities, mature cnidocytes, developing cnidocytes or secretory cells (respectively bottom to top).

dozen molecularly distinct secretory subtypes. Among the shared neural modules, we first identify a pan-peptidergic module (Module 6) for peptidergic neural identity. This module encompasses pan-peptidergic effectors and transcription factors including a GATA factor, a ETS factor, and bHLH factors such as *NvAshA/Nem17*, an achaete-scute-related gene involved in embryonic peptidergic specification[39,52] (Figs. 3.26 to 3.29). Thus, we propose that this set of transcription factors underlies early peptidergic maturation, consistent with our model of *Elav*$^{+}$ peptidergic neurons arising from a shared cellular lineage. These genes are also specifically enriched in putative peptidergic precursors in the progenitor-specific subclustering (Fig. 3.14). Top co-expression relationships between modules further highlight

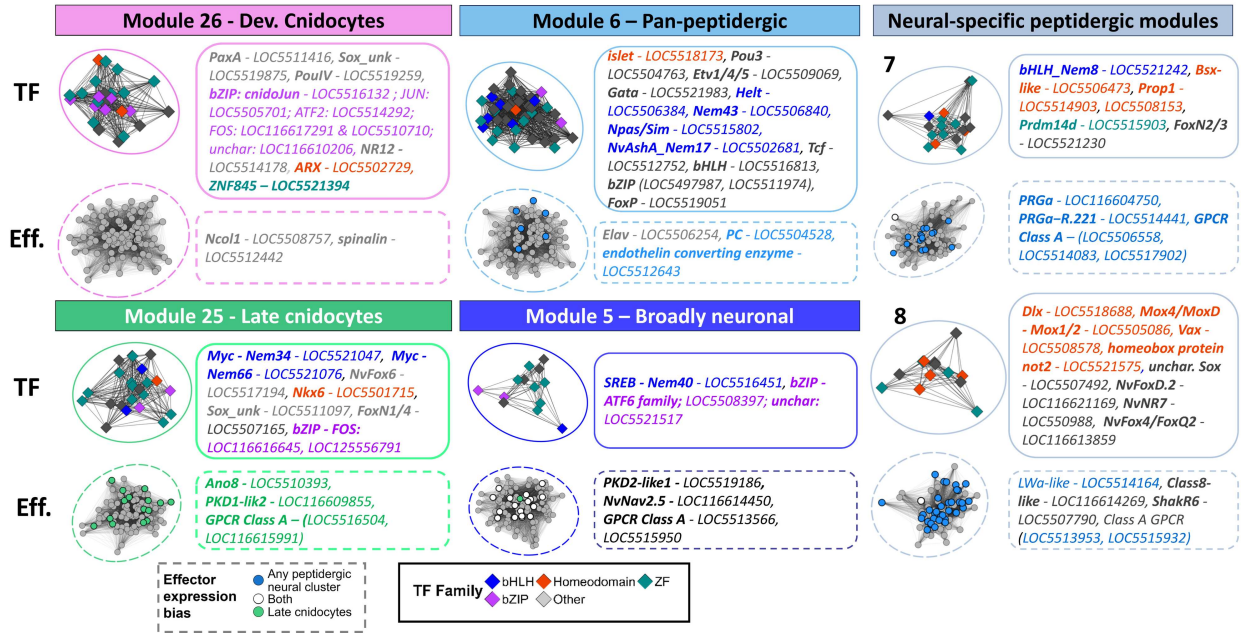

Figure 3.28: Cell class-specific and subtype-specific modules encompass known regulators and effector genes.

WGCNA modules from Fig. 3.26 showing co-expression networks and highlighting transcription factors (**top**) and effectors (**bottom**) genes among pan-peptidergic and neural subtype-specific gene modules.

this pan-peptidergic module 6 as a "hub" for peptidergic subtype-specific gene modules 7 to 21 Fig. 3.29. Secondly, we find a smaller pan-neural module partially shared between late cnidocytes and peptidergic neurons (Module 5). It includes few neural effectors and transcription factors, such as the SNAIL family C2H2 zinc finger transcriptional repressor *Scratch*, also used in vertebrate neural maturation[253,254]. While this module spans both late cnidocytes (neural phase) and peptidergic neurons, it is not utilized in the first stage of cnidocyte maturation (capsule phase) and thus does not correspond to a common lineage intermediate. Furthermore, as our early peptidergic neural fate specification model suggests, progenitor Module 1 includes the atonal-related bHLH transcription factor *Nem10* (*NvAth-like*/*NvArp3*), an embryonic neurogenesis regulator[39,51,171] (Fig. 3.26). In contrast, the developing cnidocyte module (Module 26) encompasses all known embryonic regulators of early cnidocyte development (*PaxA*, the bZIP *cnidoJun*, *NR12*, and zinc finger factors like

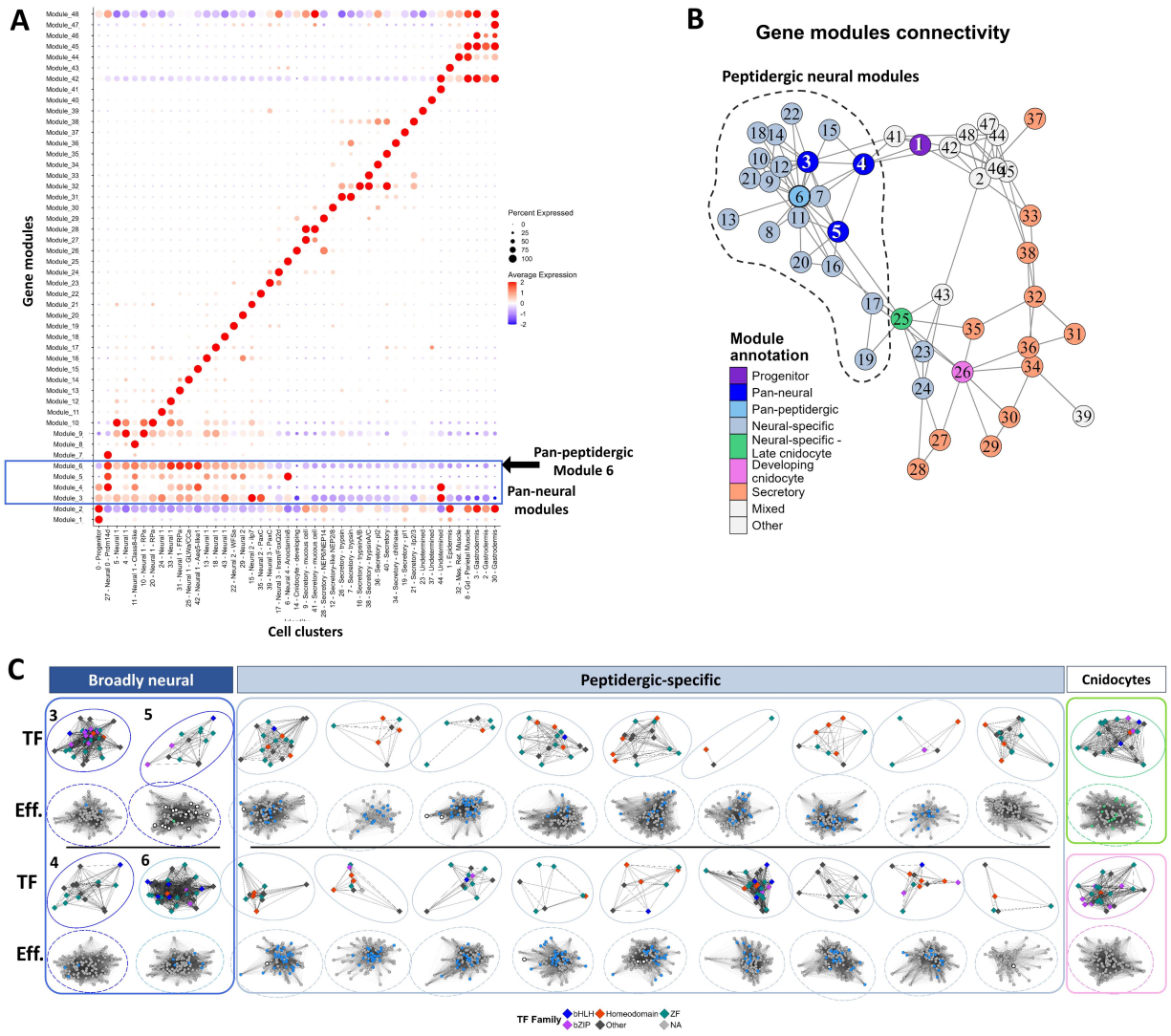

Figure 3.29: Neural class-specific and cell-type specific WGCNA gene modules

A Module expression scores across clusters with pan-neural (Modules 3-6) and neural subtype-specific modules (Module 7-22 and 26). **B** Gene modules top co-expression relationships with the gene module identifiers, same module colors as in C. Broadly neural gene modules are in light blue. **C** Neural WGCNA modules transcription factor (**top**) and top effectors (**bottom**) co-expression networks. Broadly neural modules (Modules 3-6) encompass bHLH (blue diamonds), bZIP (purple diamonds) and zinc finger (dark green diamonds) transcription factors. Peptidergic and late cnidocyte shared neural effector genes (in black in Fig. 3.25) are in shared module 5 (in white circles). Each peptidergic subtype module (Modules 7-22) has homeodomain transcription factor(s) (red diamonds), zinc finger factors (dark green diamonds), and peptidergic-specific neural effector genes (blue circles, same genes as in Fig. 3.25).

ZNF845[136,188,193,194]. These genes are also specifically enriched in putative cnidocyte precursors in the progenitor-specific subclustering (Fig. 3.14). While cnidocyte and peptidergic

neural programs exhibit limited transcriptional convergence and rely on distinct maturation strategies, there are a few shared factors which suggest some pan-neural shared regulation.

Overall, we find classes of transcription factors systematically associated with pan- (bHLH, bZIP classes) and subtype-specific (Homeodomain, zinc fingers classes) neural modules. bHLH transcription factors play proneural roles in bilaterians, and cnidarian orthologs of the *achaete/scute* and *atonal/NeuroD*[40,50–52,171] are associated here with shared pan-neural modules and broad peptidergic and cnidocyte neural fate, along with *Myc* family genes (Figs. 3.27 and 3.29). Sets of bZIP factors are enriched in pan-peptidergic, pan-neural, developing and mature cnidocytes modules (respectively modules 6, 3-5, 26 and 25). In addition to class-specific neural bHLH and bZIP factors, each neural subtype is defined by a unique homeodomain and zinc finger code (Figs. 3.27 and 3.30). Distinct homeodomain factors, with minimal overlap, specify peptidergic neurons, late cnidocytes, and sensory cells; a strategy exclusive to neural cells and absent in secretory cells. This homeodomain code is associated with subtype-specific effectors, including ion channels, neuropeptide precursors, and receptors. Developing and late cnidocytes also exhibit unique homeodomain and ZF transcription factors, with only *PouIV*, a BTB (*Broad complex*, *Tramtrack*, *Bric-à-brac* family), and an H2C2 ZF transcription factor shared between the two phases.

Together, these analyses highlight that adult neural identity integrates three transcriptional modules: (1) a pan-peptidergic module defining peptidergic neural identity, (2) a small, broadly neural module partially shared between late cnidocytes and peptidergic neurons, and (3) subtype-specific programs with a homeodomain code. Cnidocytes undergo a two-step differentiation process, each phase employing distinct transcription factors separate from those in peptidergic neurons. While peptidergic neurons and cnidocytes follow independent cellular trajectories and transcriptional programs to establish their mature neural identities, they utilize similar transcription factor classes -bZIP and bHLH factors act broadly within neural classes, whereas homeodomain and zinc finger factors are subtype-specific. These modu-

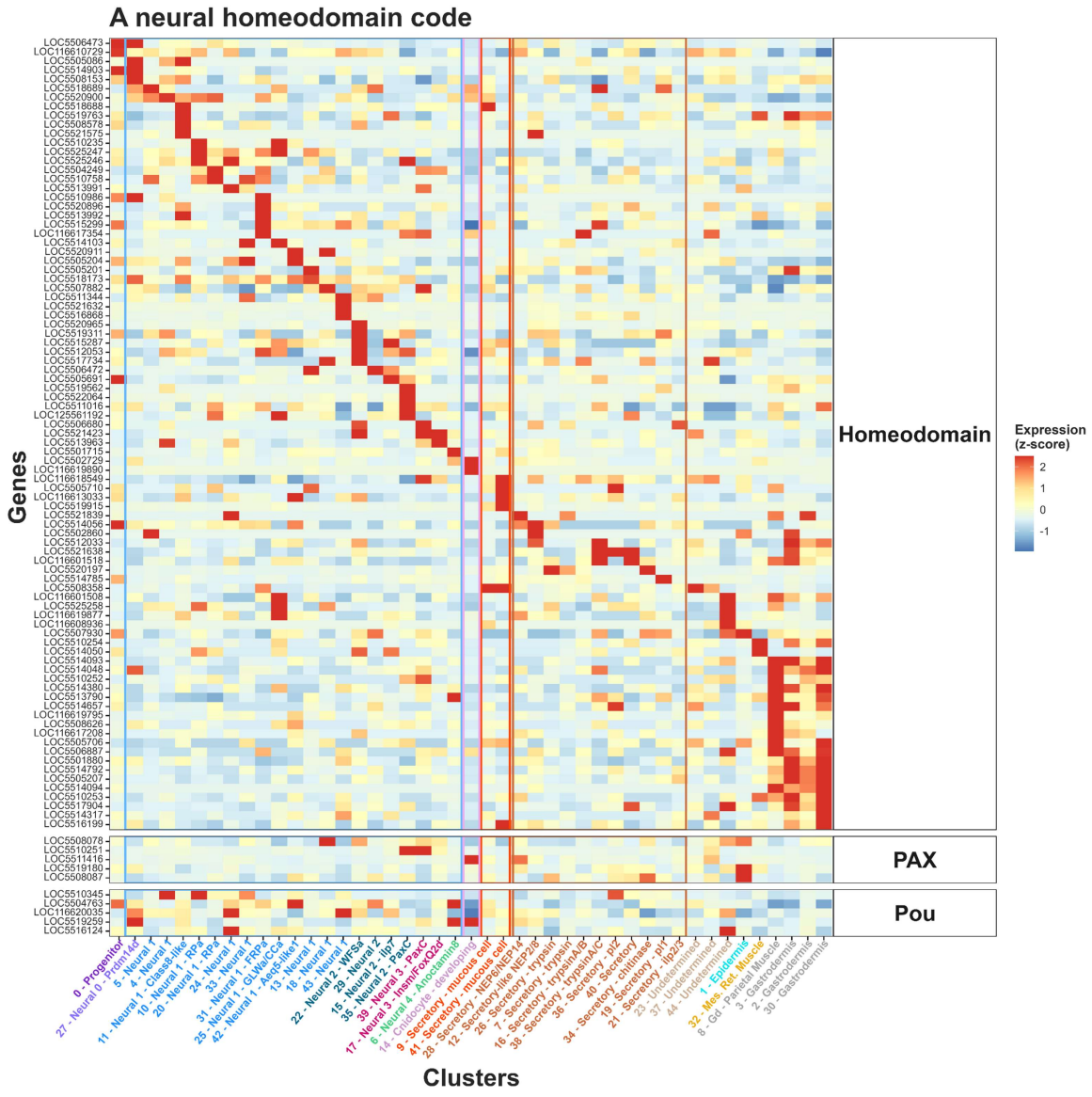

Figure 3.30: A neural homeodomain code

Expression of Homeodomain proteins across all clusters highlights neural subtype-specific expression, with each cluster expressing a unique set of transcription factors.

lar neural renewal strategies are unique to neural cells and are not employed in digestive secretory cells, despite their shared cellular mode of renewal with peptidergic neurons.

3.4 Discussion

A diversity of animals, including cnidarians, planarians and acoels, display indeterminate growth, necessitating lifelong (re)generation of multiple neuron types[241]. Here, we elucidate the cellular and transcriptional trajectories underlying adult neurogenesis in the sea anemone *Nematostella vectensis*. Using reporter tracing, nerve net imaging and single-cell transcriptomics, we show that active neurogenesis expands the nerve net along the primary body axis through two distinct pathways: (1) peptidergic neurons, the major neural cell class, arise directly from multipotent progenitors, with subtype identities specified in proportion to existing peptidergic populations, and (2) cnidocytes undergo a two-stage maturation process characterized by a transcriptionally distinct intermediate phase, marked by phase-specific transcription factors. In the context of indeterminate growth, our observations show that neurogenesis occurs across the primary body axis, without a restricted growth zone, reminiscent of the distributed processes described in planarians and acoels[241,244]. The neurogenic rate uncovered -more than one new cell for two existing ones in the body per one-week period, across both peptidergic neurons and cnidocytes, is consistent with the exponential growth rates observed in fed adults[166]. Importantly, the concomitant generation of neural subtypes in fixed proportions occurs independent of injury. This suggests that regenerative capacities may be rooted in continuous cell renewal rather than exclusively as a damage response. The occurrence of homeostatic neural renewal is tightly associated with extensive injury-induced regeneration and asexual reproduction abilities across metazoans. The reoccurrence of the known peptidergic, cnidocyte and pan-neural embryonic regulators in the corresponding adult gene modules further suggests that the molecular programs governing neural and cnidocyte specification may be deeply conserved across developmental stages.

We show the continuous differentiation of cells from the adult stem-like progenitor compartment[202,203,255] into multiple cell types, including neural and secretory cells. Our results suggest potential self-renewal within a subset of the progenitor pool; however, this has not yet been directly demonstrated. Thus, whether some of these cells are bona fide stem cells and whether they might correspond to the dispersed slow-cycling cells reported in the literature[203,222,225,256] remains unresolved. Such definitive identification of a self-renewing stem cell pool awaits future advances in genetic lineage tracing in this species. The transcriptional signature of this progenitor pool resembles that of adult multi- to pluripotent stem cells found in planarians, acoels and in *Hydra* and *Hydractinia*[246,247,257,258], pointing to shared molecular mechanisms of stem cell potency in animals. Notably, our study shows that the *Elav>KikGR*$^{+}$ expression dynamics reveal that *Piwi1*$^{+}$ progenitor cells, despite retaining undifferentiated transcriptional profiles, are already committed to differentiation. This indicates that the decision to differentiate (and the fate specification) occurs early, in cells with stem-cell-like molecular identity, rather than being limited to progenitors that have begun expressing cell-type-specific genes (putative precursors). This further suggests that fate decisions can be made alongside the cell cycle, as proposed for planarian neoblasts[259]. This contrasts with neural development in fixed-growth animals, where transit-amplifying intermediates, such as Drosophila neuroblasts, or mammalian apical radial glia, sequentially generate neurons during embryonic neurogenesis[260]. In *Nematostella*, direct neurogenesis may enable rapid neuron production to sustain indeterminate growth, with all body wall peptidergic neuron subtypes arising from undifferentiated progenitors per four-day period. However, this could constrain molecular diversity, producing fewer neuron types compared to fixed-growth bilaterians.

Our study further uncovers lineage biases within the stem-like progenitor pool. Although all three drivers label the progenitor compartment, their resulting progeny show only partial overlap, with differential neural and secretory cell class recovery. This suggests two possible

models for neural cell fate determination. One model proposes a shared "neuroglandular" lineage, where cells exiting the progenitor state are committed to peptidergic and digestive secretory cell fates[53]. Alternatively, a second model proposes that commitment to these fates occurs earlier within the progenitor compartment, with distinct cellular lineages forming upon exit. Our data, along with the reanalysis of a *Prdm14d* driver[181], strongly support the latter scenario. This is consistent with *insm1* reporter data[199], expressed in progenitors, secretory cells and neurons in adults. Furthermore, the concurrent generation of peptidergic neurons and digestive secretory cells, in proportions that reflect their relative subtype abundance, suggests that subtype identity is pre-determined at exit from the progenitor state, consistent with the distinct putative peptidergic and secretory-specific progenitor subsets.

In contrast to the direct generation of peptidergic neurons, cnidocytes, the hallmark stinging cell of the phylum, undergo a distinct two-step transcriptional maturation. They transition from a cnidarian-specific capsule-forming program to a bilaterian-like neural program, each associated to distinct sets of transcription factors. The secondary neural phase of cnidocyte development shares some transcription factors with peptidergic neurons, suggesting limited transcriptional convergence across neural classes. Cnidocytes are conserved across all cnidarians[132,261,262], but their renewal mechanisms vary. In hydrozoans, which diverged from anthozoans over 500 million years ago, such as the jellyfish *Clytia hemisphaerica*, they arise from stem cell niches in tentacle bulbs, while *Hydra* produces migratory nests of 4/8/16/32 cells from multipotent interstitial stem cells[145,263]. In contrast, the sea anemone *Nematostella vectensis* (an anthozoan) displays dispersed early-stage cnidocytes in the epidermis and mesenteries, even as it remains unclear whether they arise from the stem-like progenitor pool (also in[202,203]). This highlights the evolutionary plasticity of cnidocyte development, despite their conserved capsule genes. Thus, investigating whether orthologous transcription factors and neural effectors genes are used in cnidocytes and peptidergic neurons across cnidarians could shed light on neural cell evolution and diversification,

particularly considering their lineage-specific neural effectors expansion (e.g., voltage-gated potassium channels, GPCRs[98,264]).

Although cnidocytes and peptidergic neurons use distinct renewal modes, their early development is marked by transient *SoxC* expression, followed by downregulation during maturation. This parallels vertebrate neurogenesis, where *SoxC* factors (*SOX4/11/12*) play a critical role in neural progenitor differentiation[265], underscoring the evolutionary conservation of neurodevelopmental pathways across eumetazoans, despite lineage-specific diversification. Overall, we propose a model where *SoxC*, *SoxB(2)* and distinct bHLH and bZIP transcription factors are associated to neural class specification, while homeodomain and zinc finger genes specify subtype identities. Intriguingly, the homeodomain code delineating neural identities in *Nematostella* is unique to neural cells and absent in digestive secretory cells. This is not clearly associated to axial patterning, as body wall neurons do not appear clearly spatially segregated along the primary (oral-aboral) or secondary (directive) body axes, across currently characterized populations[101,180,181,194,201]. Furthermore, the neural-subtype-specific homeodomain proteins are distinct from the ANTP-class Hox genes specifying the secondary (directive) body axis[167,168,205], expressed in cnidarian-specific gastrodermal cells in adults. In bilaterians, homeodomain transcription factors are essential for primary body axis (antero-posterior) patterning (ANTP-class Hox genes), and act as terminal selectors of neural identity[266–270]. Thus, given the early diversification of homeodomain family genes before the cnidarian-bilaterian split, their neural subtype-specific expression in *Nematostella* suggests an ancestral role in neural subtype specification, separate from spatial patterning[46,266]. Supporting this hypothesis, the POU-family homeodomain gene *PouIV/Brn3* is critical for cnidocyte maturation[120,195], as its removal results in blocked immature cnidocytes that fail to initiate their secondary neural program. Systematic functional studies of homeodomain-class genes during adult neurogenesis would help shed light on this evolutionary link. Our identification of the populations, lineages, and transcriptional pro-

grams underlying adult neural cell generation in a sea anemone reveals how indeterminate growth organisms scale their nervous system. We uncover two distinct modes of renewal across the main neural classes, shedding light on the strategies underlying lifelong neurogenesis. The shared molecular features with bilaterians point to deeply conserved neural specification mechanisms, tracing back over 540 million years to the Precambrian. These findings provide new insights into the molecular basis of adult neurogenesis and the evolutionary diversification of neurons across animal lineages.

3.5 Methods

3.5.1 Transgenic line generation

For the *FoxL2>KikGR*, *SoxC>KikGR* and *Elav>KikGR* reporter line generation, plasmid constructs generated from a pNvT-MHC::mCh vector backbone (a gift from Ulrich Technau[169], Addgene plasmid #67943) with the KikGR fluorophore (from the pCAG::KikGR vector, a gift from Anna-Katerina Hadjantonakis[214], Addgene plasmid #3260833). KikGR reporter expression was driven by the published 2.4kb upstream region of *Elav* from[38] or the upstream regions of either *SoxC* (2.2kb) or *FoxL2* (2.4kb).

Stable transgenic animals were generated using I-SceI-mediated transgenesis[169], with the following modified injection mixture: 50ng/μL plasmid concentration, 1X I-SceI buffer (10mM TrisHCl, 10mM MgCl2, 1mM DTT, pH8.8), 0.5mM Patent Blue VF (Sigma, #198218), 0.2U/μL I-SceI (adjusted to 0.4U/μL after 30min at 37°C). All experiments used stably integrated *SoxC>KikGR*, *FoxL2>KikGR* or *Elav>KikGR* 5-8 week-old heterozygote animals (sexually immature adults) raised at 18°C in 12.5ppt artificial seawater in the dark and fed 4 times a week with freshly hatched *Artemia* nauplii.

3.5.2 Whole animal photoconversion and nerve net imaging and quantification

Individual *SoxC>KikGR*, *FoxL2>KikGR* or *Elav>KikGR* immature adults were relaxed in 12.5ppt artificial seawater and immobilized by a 1:1 dilution with 70mg/mL MgCl2, then transferred between two coverslips, with tape layers to adjust depth to ca. 0.5mm for photoconversion and live confocal imaging. Whole animal KikGR photoconversion was performed with a ZEISS Axio Imager 2 (ZEISS), with a X-Cite 120LED Boost (EXCELITAS) light source at 80% power, with manually tiled regions with 8 sec exposure (emission BP 445/50nm) at 20X magnification.

Photoconversion efficiency was confirmed by imaging and flow cytometry (Fig. 3.31) following animal dissociation into single cells. Photoconverted animals were kept in 12.5ppt artificial seawater, with either 1% DMSO or without. For confocal imaging of the body wall nerve net, individual *SoxC>KikGR*, *FoxL2>KikGR* or *Elav>KikGR* adults at 2-, 4-, or 7-days post KikGR photoconversion were immobilized (as above) and imaged at 40X on a Zeiss LSM800 inverted confocal microscope, with 8-10 optical sections per z-stack with 1μm between optical slices with the enhanced green fluorescent protein (eGFP) spectrum used for detecting non-photoconverted KikGR and mCherry spectrum used for the photoconverted KikGR protein, adjusted to remove overlap.

For quantification purposes, all images from the same reporter line at 2, 4 or 7 days post-photoconversion were acquired with the same settings. All cell segmentations and quantifications were performed on raw z-stack data. For cell quantification, a cellpose 2.0 neural net model[219,220] was manually trained on randomly sampled individual images slices from *Elav>KikGR*$^+$, *SoxC>KikGR*$^+$ and *FoxL2>KikGR*$^+$ live body wall confocal z-stacks, then used to segment all raw z-stacks at all timepoints post-photoconversion separately on either the red (photoconverted KikGR signal) or green (new KikGR signal) channel for all drivers with the following parameters: diameter: 0, chan2: 0, cellprob_threshold: -0.5, flow_threshold: 0.8, stitch_threshold: 0.25 after optimization. The two sets of cell masks

per image (red$^+$ set and green$^+$ set) were then merged into a single set for each z-stack for quantification.

3.5.3 Single cell dissociation and KikGR$^+$ cell sorting

For scRNA-seq sampling, individual *SoxC>KikGR*, *FoxL2>KikGR* or *Elav>KikGR* adults at 2-, 4-, or 7-days post KikGR photoconversion (with initial experiments containing 1% DMSO added to their medium during the timecourse, and not in latter experiments) were immobilized with MgCl_2 as above and sampled below their pharynxes (body without pharynx region or tentacles) to be consistent with our body wall imaging data, then dissociated into single cells according to[218], with 1-2 individuals per dissociation. In brief, sample medium was replaced with Calcium Magnesium free seawater (495 mM NaCl, 9.7 mM KCl, 27.6 mM NaHCO_3, 50 mM Tris–HCl pH 8) with collagenases (0.05mg/mL Liberase TM, Roche), and dissociated by gentle pipetting at room temperature. The cell suspension was filtered (Ø35μm) and stained with live cell dye Calcein Violet AM (ThermoFisher Scientific, 1:2000 dilution of 1μg/μL DMSO stock) or a combination of Calcein Violet and SytoxRed dead cell dye (Thermo Fisher scientific, 1:1000 dilution of a DMSO 5μM stock).

Individual live KikGR$^+$ cells were sorted using a Fluorescence-Assisted Cell Sorter (FACS Aria IIu and III, BD Biosciences) into 384-well plates containing 2μL lysis buffer and barcoded polyT capture oligonucleotides (MARS-Seq protocol[8,232,237]) then frozen at -80°C until processing. FACS gating strategy is shown in Fig. 3.31; in brief, to sort live KikGR$^+$ cells, cell-sized particles were selected using their FSC-A (logicle-scale) vs SSC-A (logicle-scale) signal (ca. $\geq$3μm in diameter, using sizing beads (SpheroTech) as an absolute size reference). Multiplets were excluded with FSC-W and FSC-H, then Calcein Violet AM-positive live singlets were selected (ex.: 405nm, em.: BP 450/40nm). Initial sorts included a SytoxRed signal: APC channel (ex.: 633nm, em. 660/20nm) and mildly depleted the Calcein Violet^{low}/SytoxRed^{mid} population. From live singlets, KikGR$^+$ cells were selected based on

their new KikGR signal (FITC-A, ex.:488nm, em. filters: LP 502nm and BP 530nm/30) and photoconverted KikGR (PE-A, ex.: 561nm, em. filters: BP 582/15nm) signal, with the FITC-A^{low}/PE-A^{low} autofluorescent cells excluded, with a negative wild-type (no KikGR) control. A shared KikGR^{+} gate was used for all *Elav>KikGR*$^{+}$ and *SoxC>KikGR*$^{+}$ cells, and for 17/41 of *FoxL2>KikGR*$^{+}$ libraries (Fig. 3.31).

For 16/41 FoxL2>KikGR^{+} libraries, we depleted its large epithelial cell fraction, as enabled by its specific photoconverted KikGR^{+}/new KikGR^{low} signal profile ("Epithelial-depleted KikGR^{+}" sorting). To enrich for green and orange cells, we also sampled only New KikGR$^{mid/high}$ cells for 4 FoxL2>KikGR^{+} libraries ("New KikGR^{+}" gate). All quantifications of reporter status (green/orange/red cell proportions) from FACS-sorted cells are from cells sorted with the shared KikGR^{+} gate throughout the paper, with ternary plots (R package ggtern[271]) or barplots (R package ggplot2) visualizations, filtering out clusters with less than 10 sampled cells. For each KikGR^{+} sorted cell, cell reporter status (i.e. green, orange or red) for all drivers (*Elav*, *FoxL2*, *SoxC*) was assigned using red, orange or green gates defined across all sorted cells (Fig. 3.32). Cell identifiers were corrected for two sorted 384-wells plates (4 libraries: RP0885_E/O, RP0886_E/O), as the flow cytometer mechanical arm moved by one column prior to the sort. For each library sorted with the KikGR^{+} gate and used for color reporter analysis, matched flow cytometry data was validated through its correlation between FSC-A signal (a cell volume proxy) and total UMI per cell.

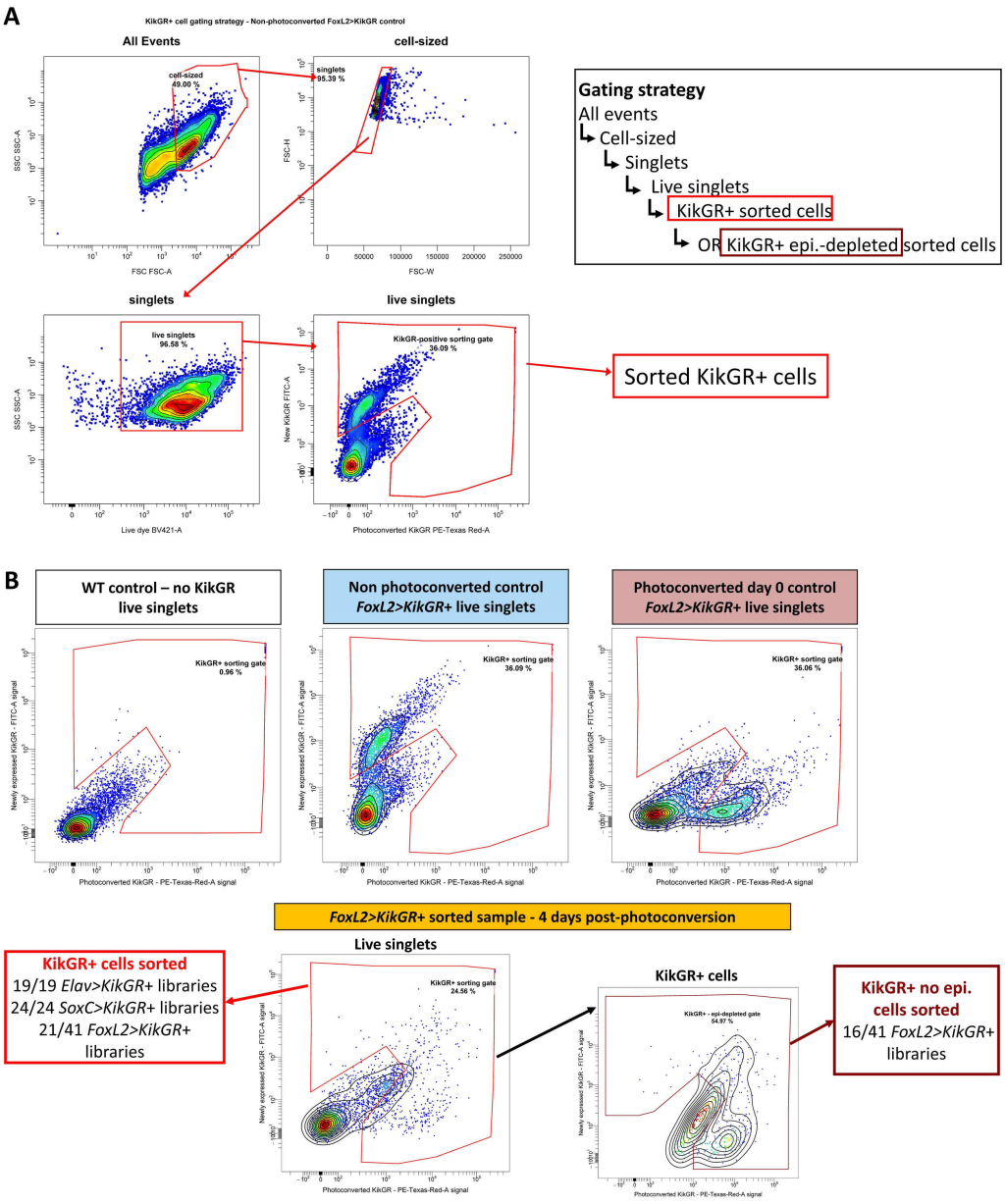

Figure 3.31: KikGR+ cell flow cytometer gating strategy

A Fluorescence-Activated Cell Sorter (FACS) strategy to sample KikGR$^+$ cells from adult body (excluding pharyngeal and tentacles regions) cell suspension for all drivers. Cells sized particles are selected, then multiplets are excluded based on their FSC-W/FSC-W signal and Calcein-positive cells are retained (live singlets fraction), then KikGR$^+$ cells are selected. **B** The KikGR$^+$ sorting gate is defined using WT (wild-type, no KikGR reporter), non-photoconverted (all 'green' cells) and fully-photoconverted (all 'red' cells) samples. Autofluorescent cells (low PE-Texas-Red-A/low FITC-A double-positive signal) are excluded from the KikGR$^+$ gate using the WT control sample. The same KikGR$^+$ sorting gate is used for all *Elav>KikGR*$^+$ and *SoxC>KikGR*$^+$ cells at all timepoints post- photoconversion. *FoxL2>KikGR*$^+$ cells are sorted with the KikGR$^+$ gate or the 'KikGR$^+$ no epi.' gate to deplete the epidermis epithelial cell fraction. All quantification of reporter color status (green/orange/red) for all drivers always use cells sorted with the KikGR$^+$ sorting gate.

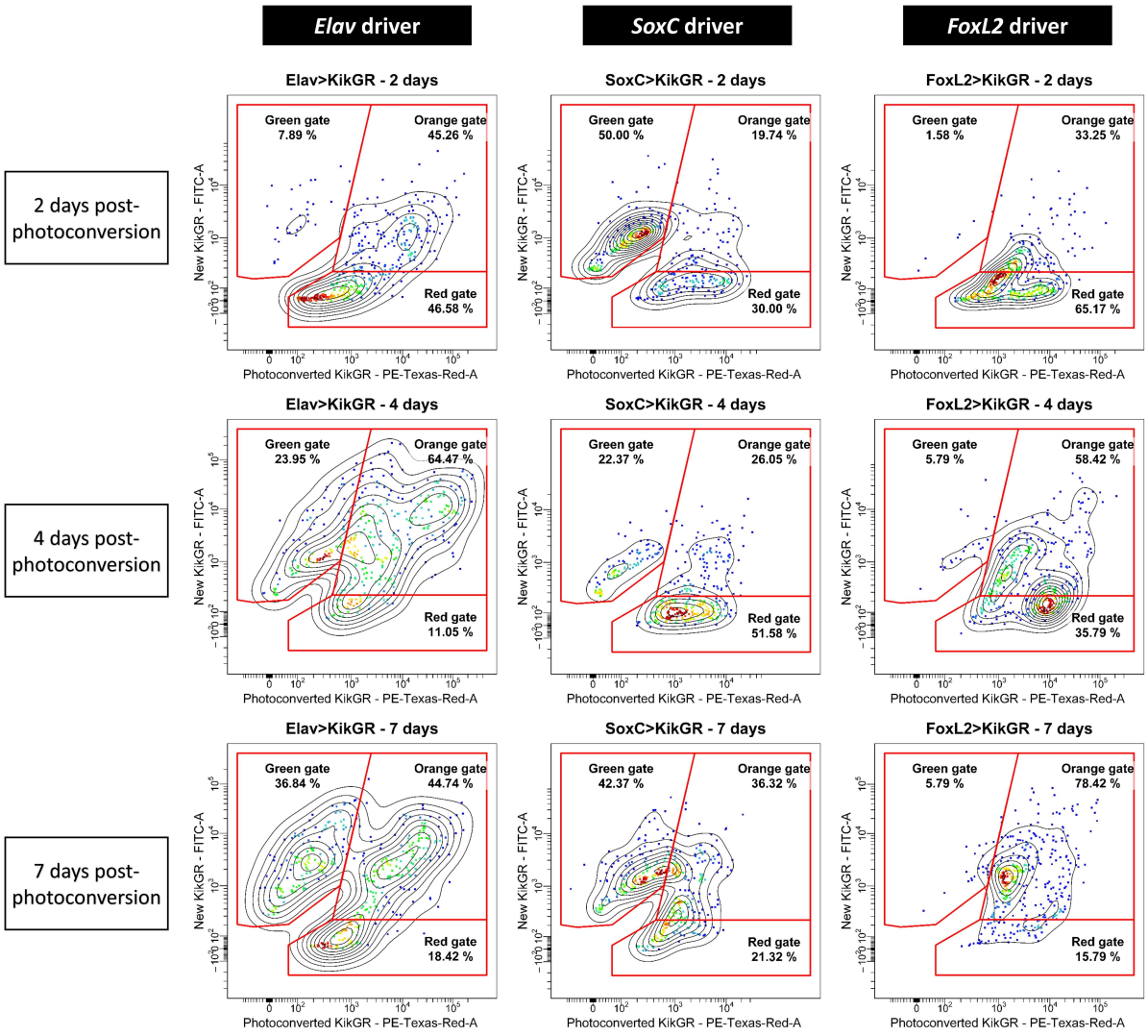

Figure 3.32: Reporter status attribution ('green', 'orange' or 'red' cell)
"Green", "orange" and "red" gates are defined for all KikGR$^+$ sampled cells (KikGR$^+$ sorting gate, see Fig. 3.31) at 2,4, or 7 days post-photoconversion for the *Elav*, *SoxC* and *FoxL2* drivers. Representative data from 380 sorted cells (2 half-plate libraries) shown for each driver and timepoint post- photoconversion.

3.5.4 Single cell RNA-Seq analyses

We generated 84 scRNA-Seq libraries according to published MARS-Seq protocols[8,232,237]. In brief, all plates were processed with the same protocol, mRNA was reverse transcribed into cDNA with a barcoded oligodT capture oligonucleotide in the lysis buffer containing the cell barcode and unique molecular identifier. Exonuclease treatment (ExoI, NEB) was used to remove unused oligodT nucleotides. Pools of cDNA were then amplified by *in vitro* transcription (T7 polymerase), then fragmented. Each RNA fragment was ligated to an oligonucleotide containing a library pooling barcode with a T4 ssDNA:RNA ligase before reverse transcription. Final libraries were obtained from PCR amplification with 17 cycles. Library fragment distributions and concentrations were assessed by Tapestation 4150 (Agilent) and Qubit3.0 (Invitrogen).

scRNA-Seq libraries were pooled equimolarly and sequenced in paired-end mode on a NextSeq550 (Illumina) using a high output 75 cycles kit (up to 32 libraries per run with 190 cells per half-plate library). Median sequencing depth per library was 13.4M reads (for 190 sorted cells). Fastq files were reformatted to the following read structure for compatibility with STARSolo: Read1: mRNA fragment; Read2: 4 nucleotides (nt) library barcode, 7 nt well barcode (11nt total for the cell barcode), 8 nt UMI barcode. Reads were mapped and demultiplexed using STARSolo, with the following parameters -soloType CB_UMI_Simple -soloUMIlen 8 -soloCBlen 11 -soloUMIstart 12 -soloCBstart 1 -soloCBmatchWLtype 1MM -soloUMIdedup Exact; with the cell barcode list for each sequencing run as the soloCB-whitelist. Pairedend fastq files per halfplate library are provided in the GEO archive (GSE288441), along with cell barcode lists for each pool of demultiplexing-compatible libraries.

Reads were mapped to the jaNemVect1.1 Nematostella vectensis genome assembly (from the Darwin Tree of Life (DToL) project – Wellcome Open Research[163]), with the KikGR sequence added[214], using the genes models from the RefSeq *Nematostella vectensis* Annota-

tion Release 101 from NCBI (GCF_932526225.1, NCBI) (in GEO archive GSE288441). Of the 84 libraries, from 27 independent biological dissociations (Fig. 3.33), a total of 19, 24, and 41 libraries correspond to the *Elav>KikGR*, *SoxC>KikGR*, and *FoxL2>KikGR* drivers, respectively, with an average of 2131 median RNA molecules recovered per library (see supplementary data files available online at[240] and Fig. 3.34). We removed the 164/16012 control empty wells, and the remaining 15848 cCells were filtered based on RNA molecules recovery (excluding cells with ≤300 or ≥13000 recovered RNA molecules, which removed 468/15848 (2.9%) low quality cells), and low quality clusters (lower UMIs, no specific marker gene, a large fraction of stressrelated genes and ribosomal proteins) were filtered out, as well as clusters with cells from one biological dissociation, only from $DMSO^+$ samples or from a single sequencing run for a total of 3030/15848 (19%) filtered out cells. After filtering, 12352/15848 (78%) cells were retained, comprising 78% of sorted cells. Filtered and unfiltered UMI tables are available in the GEO archive (GSE288441).

We then used Seurat v5 (5.2.1) with SCTransform v2 (sctransform 0.4.1) normalization for clustering, using the top 4000 variable genes (excluding the KikGR transgene), with the top 65 PCA components used and a 2.4 clustering resolution. Clusters were annotated based on literature-curated markers (from[50,101,120,151,173,178,181,188,202]. The *Elav* subclustering (Fig. 3.3) was processed similarly, with only *Elav>KikGR*$^+$ cells included. *Elav>KikGR*$^+$ cells, *SoxC>KikGR*$^+$, and *FoxL2>KikGR*$^+$ cells were also individually clustered as described above, with a resolution of 2, 2.5 and 2, respectively.

Due to the multiple genome assemblies and gene models used in the literature, gene annotations were transferred to the GCF_932526225.1 gene models (jaNemVect1.1 genome assembly, gene identifiers starting with LOC) using either BLASTP of protein sequences, primerBLAST of in situ hybridization probes from the original publications or best reciprocal BLASTP hits; as listed in Supplementary File 1, with the GCF_932526225.1 gene annotation used by default for all other genes. Transcription factors were predicted with

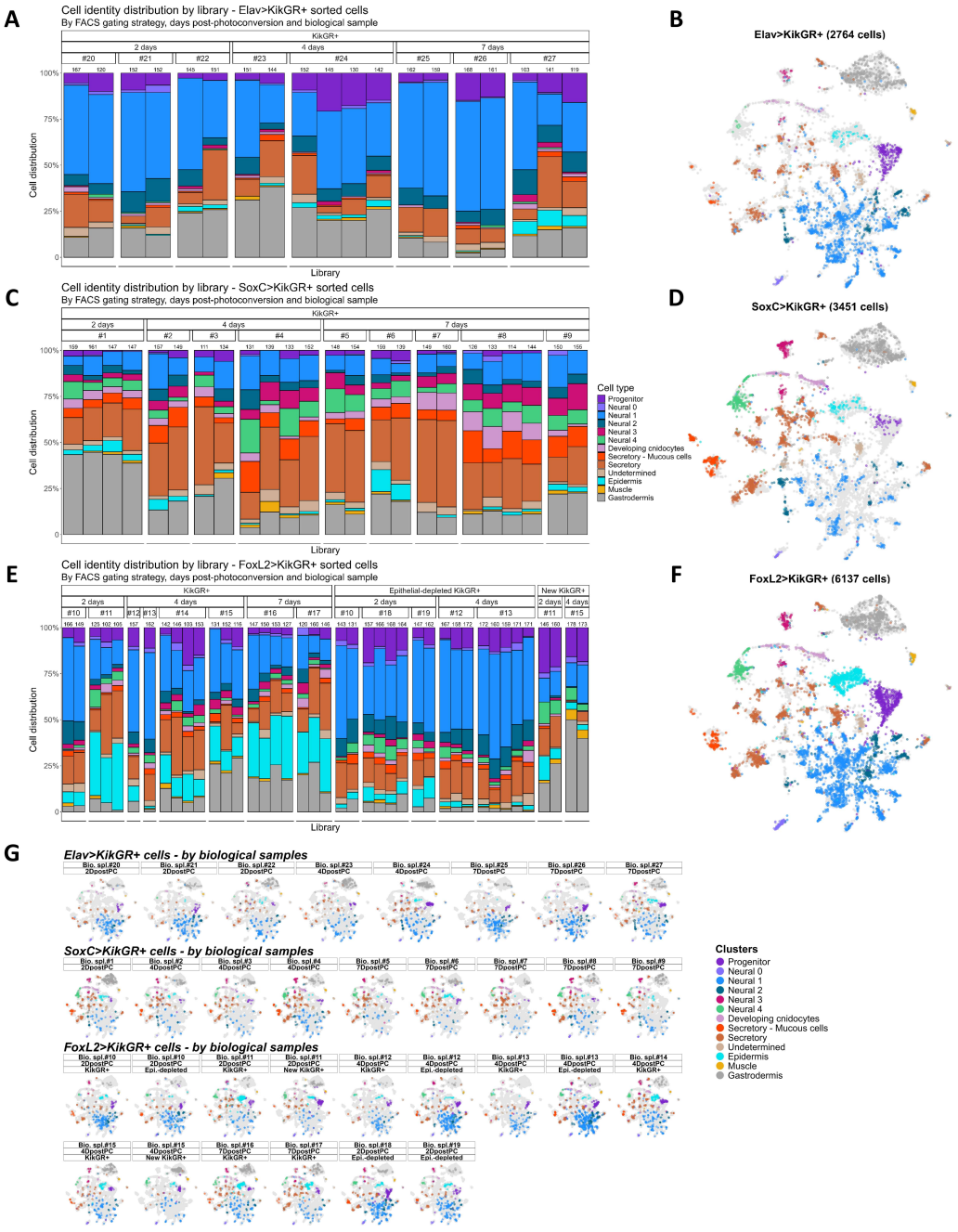

Figure 3.33: Cell identity distribution by reporter and scRNA-Seq library

A,C,E Cell identity distribution by scRNA-Seq library across all 84 libraries and 27 independent biological dissociations by reporter line sampled, FACS sorting strategy, day post reporter photoconversion and biological sample (**A**: *Elav>KikGR*$^{+}$ cells, **C** *SoxC>KikGR*$^{+}$ cells, **E** *FoxL2>KikGR*$^{+}$ cells). The number of cells recovered after filtering from each library is indicated above each barplot. **B**,**D**,**F** UMAP projection of scRNA-Seq sampled cells from each driver colored as in A/C/E by broad cell identity. **G** UMAP projection of each of the 27 independent biological dissociations (same identifier as in A/C/E) across all reporter lines, days post-photoconversion and FACS gating strategies.

AnimalTFDB83, with bZIP transcription factors named from their families from[272], and bHLH transcription factors annotated from[39,52]. Cell cycle module genes (22 G2/M and 29 S-phase genes) were identified in *Nematostella* by mapping Seurat's mouse cell cycle markers using eggNOG mapper annotations. The distribution of cell identities across libraries was consistent across biological samples for each driver and fluorescence-activated cell sorting (FACS) gating strategy Fig. 3.33. Gene expression across clusters is visualized using DotPlot from SCTransform data for the global clustering and from the RNA data for the line-specific reclustering, with expression values normalized to the maximum expression observed across clusters, and dot size representing the percentage of cells expressing the gene in each cluster.

To investigate differential neural effector gene usage between peptidergic neuron clusters (18 clusters) and late cnidocytes (1 cluster), we annotated genes coding for Class A GPCRs, mGluR-like receptors, and various ion channels based on PFAM domains using eggNOG mapper v2. The PFAM domains used included 7tm_1 and 7tm_3 for GPCRs (respectively Class A, and mGluR-like), as well as ASC, Anoctamin, NT_gated_channel, iGluR, PKD_channel, Lig_chan-Glu_bd, and Neur_chan_LBD for ion channels. De-orphanized neuropeptide GPCR receptors were further annotated from[98]. To assess the differential usage of neural-specific genes, we excluded lowly expressed genes (defined as those with fewer than 50 UMI across neural clusters) and genes whose highest expression occurred in non-neural clusters. Given the presence of multiple peptidergic clusters but only a single late cnidocyte cluster, we systematically compared the late cnidocyte cluster to the peptidergic cluster with the highest expression of each effector gene. The expression matrix was aggregated using Seurat's AggregateExpression function, with a scale factor of 1000. We calculated the log2 fold change in expression between the late cnidocytes and the peptidergic cluster with maximal effector expression, plotting this against the average expression between the two clusters for 499 genes across six categories.

For the cnidocytes developmental trajectory inference, cnidocytes (clusters 6 and 14)

were subsetted from the global scRNA-seq clustering and processed in Seurat with standard normalization to generate a UMAP embedding. Pseudotime analysis was performed using slingshot (2.12.0)[273] on the UMAP coordinates. Gene expression along pseudotime was then estimated using a generalized additive model with the gam function from the mgcv R (1.9-1) package. For the progenitor-specific reclustering, progenitors were subsetted from the global scRNA-seq clustering and processed in Seurat with standard normalization, using 30 PCs and a resolution of 1.5, and annotated based on the expression of progenitor-enriched genes and cell-type-specific marker genes.

Weighted Gene Co-expression Network Analysis (as implemented in the WGCNA R package) was used to identify 48 transcriptional modules in the global clustering[252]. Gene expression data from the Seurat object were filtered to retain genes with at least 50 UMIs, and aggregated expression was computed per cluster with log normalization and a scale factor of 1000. A weighted gene co-expression network was then constructed using a soft-threshold power of 6. The adjacency matrix was transformed into a topological overlap matrix (TOM), and hierarchical clustering was performed. Modules were identified via dynamic tree cutting (deepSplit = 3, minClusterSize = 20), then merged based on eigengene correlations, using a dissimilarity threshold of 0.25. Module eigengenes were computed and gene-module membership was determined by correlating gene expression with module eigengenes across clusters. A low (0.35) and high (0.5) threshold were used to visualize soft module membership for transcription factors. Gene module expression across clusters Fig. 3.29 was computed using Seurat's AddModuleScore function with each module composed of genes with module membership $\geq$0.5 (stringent threshold), and each gene only associated to its top scoring module. These thresholded gene modules were then used to visualize the network of module top co-expression relationships, derived from the average pairwise TOM values, by filtering out module-module average scores below 0.002.

3.5.5 Reanalysis of RNA-Seq data

Upregulated and downregulated genes from the *PouIV*-KO and *Prdm14d>GFP*$^{+}$ RNA-Seq datasets, as reported in the original publications[120,181], were converted to GCF_932526225.1 gene identifiers using best reciprocal BLASTp hits. Gene set enrichment analysis was then performed using the GSEA R package (fgsea 1.30.0), with cluster gene markers from the FindAllMarkers Seurat function on the global clustering (with an adjusted p-value threshold of 0.01) used for the 'pathways' lists (26-1265 genes per cluster-associated set).

3.6 Code availability

Data analysis code (R version 4.4.1) and processed annotated Seurat sc RNA-Seq objects are available on zenodo (DOI: 10.5281/zenodo.14977132).

3.7 Data availability

The *SoxC>KikGR*, *Elav>KikGR* and *FoxL2>KikGR* transgenic *Nematostella vectensis* animals and associated plasmids are available upon request to the corresponding author. Sequencing data is available in GEO (accession: GSE288441), with paired-end fastq files for each sequenced library, alongside the reference genome and associated gtf used for mapping (jaNemVect1.1 *Nematostella vectensis* genome assembly with the KikGR sequence added), cell metadata information and raw and filtered global count tables. Source data are provided with this paper as Source Data Files, as well as supplementary videos of the nerve net imaging.

3.8 Acknowledgments

We thank Sébastien Bastide for data analysis discussions and comments on the manuscript. We thank Caroline Kaiser, Ilaria Rebay, Paschalis Kratsios and Urs Schmidt-Ott, and mem-

bers of the Marlow lab for discussions and comments on the manuscript. We thank Colette Plessier for the Nematostella silhouettes. We thank Baptiste Saudemont for MARS-Seq training, Sophie Novault and Sandrine Schmutz for FACS training and expert advice. We thank the Flow Cytometry Core Facility in Institut Pasteur (Paris, France) for BD FACS AriaIII access. We thank David Leclerc, Mike Olson and Bert Ladd of the University of Chicago Cytometry and Antibody Technology Facility for BD FACS AriaII training and advice (Facility RRID: SCR_017760). We thank Marie-Pierre Mailhé for assistance with wild-type animal care. We thank the Developmental and Stem Cell Biology Department of Institut Pasteur (Paris, France) for Zeiss LSM800 confocal microscope access. We thank Matt Kaufman for image processing advice. We thank the University of Chicago's Research Computing Center for HPC access for image processing. We thank the Ecole Normale Supérieure de Lyon and the Ministère de l'Enseignement Supérieur et de la Recherche (France) for funding (to F.P.). This study was funded in part by NIH NIGMS grant 5R35GM155159 to H.M.

3.9 Author Contributions

H.M. conceived the study. H.M. and F.P. designed experiments and interpreted the data. F.P. performed experiments, developed and implemented data analyses, generated figures, and wrote the initial draft. F.P. and H.M. wrote and edited the manuscript. H.M. supervised the project and acquired funding.

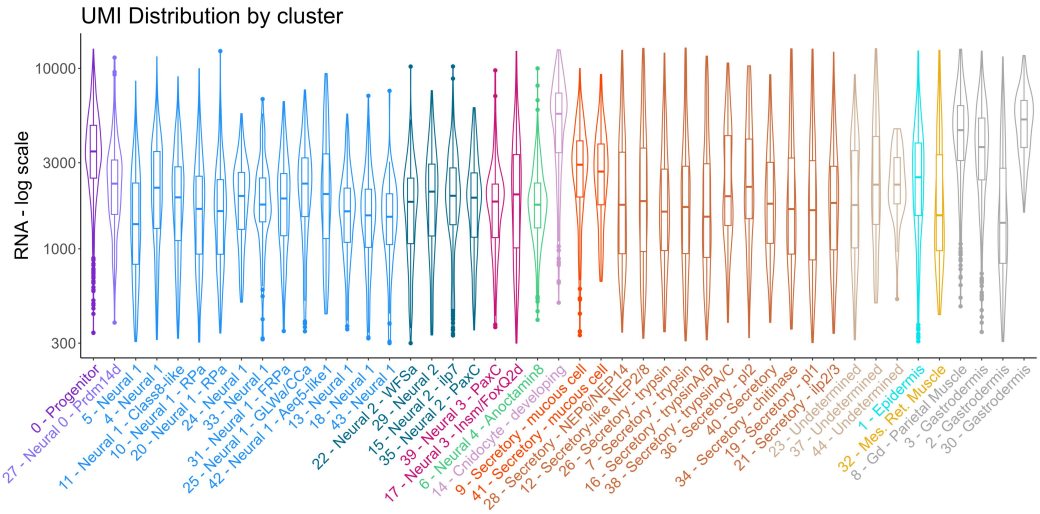

Figure 3.34: Total mRNA detected (total UMIs) by cell cluster

Distribution of total mRNA (total UMI, log scale) detected by cell cluster across all KikGR+ cells. Note the higher relative total UMIs detected in progenitors (purple), developing cnidocytes (pink), and some gastrodermis clusters (grey).

CHAPTER 4

DISCUSSION

4.1 Chapter overview

This dissertation work investigates the cellular lineages and gene programs underlying neural fate acquisition during lifelong homeostatic growth in the sea anemone *Nematostella vectensis* (Cnidaria, Anthozoa), characterizing how distinct neural cell types are generated, including neural lineage restrictions and their associated gene programs, helping inform our understanding of neural developmental processes in animals.

In section 4.2, we discuss mechanisms of adult neurogenesis, and we interpret findings from chapter 2 and chapter 3 as supporting a proportional and distributed, rather than niche-restricted, additive mode of lifelong neurogenesis, across neurons and cnidocytes, and compare this organization to other cnidarian and metazoan tissue renewal mechanisms. In section 4.3, we focus on the progenitor population(s) enabling this neurogenic process, and discuss whether the *Piwi1*$^{+}$ progenitor population constitutes a multipotent stem cell system, highlighting similarities with other multipotent stem cell states across animals, and overall supporting a model of hierarchical organization of lineage decisions in adult stem cell systems. In section 4.4, we discuss the broader significance of two distinct neural gene programs active in peptidergic neurons and late cnidocytes, and their relationships to other cnidarian neural programs, and the implications for the evolution of neural cell types at the base of Cnidaria.

4.2 Distributed adult neurogenesis

Cellular renewal, in varying grades, is ubiquitous across animals. From the lifelong production of gametes and select organ or tissue renewal to complete cellular turnover and indeterminate growth, these processes are critical for organismal function and overall fitness. In this section, we discuss the cellular processes that enable nervous system expansion and

generate distinct neural subtypes in the sea anemone *Nematostella vectensis*, an indeterminate growth context, integrating experimental data with recent findings and comparing with other cell renewal systems.

4.2.1 Is there a stable neural net tissue architecture?

The distributed nervous system of *Nematostella* appears spatially organized within distinct tissue layers and structures. However, how stable is this overall tissue organization in this lifelong growth context, and whether individual mature neurons can be stably positioned within the nerve net or are actively migrating or passively displaced due to tissue movement is unknown. Thus, experimental data of the localization of newly-generated neurons relative to existing ones, and their molecular identities and lineages relationships (chapter 3) can have different interpretations in a system with global tissue displacement (as seen in *Hydra* and the tentacle bulb of the hydrozoan jellyfish *Clytia hemisphaerica*) or with local tissue architecture stability. Thus, before discussing the generation of new neurons and cnidocytes, we review in this section experimental evidence of stable body wall neurons, and propose links between neuronal body position and neurite anatomical localization, with implications for neural circuit functions.

Interestingly, longitudinal in vivo imaging shown in chapter 2 identified the same morphologically detectable neurons in the same polyp body wall nerve net area over a 9 days period in vivo, which is longer that the timeline of neurogenesis (in chapter 3). To our knowledge, this is the first time that neurons could be longitudinally tracked in the nerve net in a cnidarian, and identified in the same position, thanks to the approach developed here. Furthermore, these tracked neurons displayed similar neurite branching patterns, for both epidermal-embedded and subepidermal neurons, across both oral (in section 2.3) and aboral body wall nerve net regions (not shown). This supports a stable organization of at least a subset of the nerve net. The timeframe imaged here (9 days) cannot exclude the

possibility of slower tissue displacement occurring, but this period is longer than the experimentally demonstrated neuron differentiation timescale in chapter 3, and as the same neurons can be detected across both oral and aboral body wall nerve net regions this overall supports some nerve net stability for mature neuron cell bodies and their neurites. Whether all mature neurons are stably localized is unknown, but longitudinal live imaging approaches could help track these cells across subtypes.

This experimental data supports that morphologically mature neurons with neurites can be stably integrated into the nervous system. Such overall tissue stability (at least in the short term) is also supported by recent experimental data in intact animals where body wall patches of *Kaede*$^{+}$ cells (a photoconvertible fluorophore) were photoconverted and remained in the same shape and relative position over 4 days, as assessed in younger animals (primary polyps)[274]. This stability under homeostatic renewal condition is further supported by the tissue-displacement and reorganization that are, in contrast, clearly detected upon injury. During oral[222] or physal (aboral) amputation[274], photoconverted patches of tissues show rearrangements and deformation, and oral patches can be found integrated into tentacular regions[222]. Such injury-specific reorganization is also supported by the presence of red *SoxC>KikGR*$^{+}$ mature cnidocytes in regenerated tentacles shown in section 2.4 - Fig. 2.6.

Furthermore, while neural circuits have not yet been characterized in *Nematostella*, its neurons are anatomically organized. In the body wall, there are known neurite tracts condensations (including neuronal cell bodies) at the base of each mesentery next to the parietal muscles, and sparse neuronal cells in the adjacent retractor muscle region, as well as neural cells embedded in the epidermis monolayer and in the subepidermal inter-mesenteric areas (see chapters 1 and 2). There are observations of preferential neurite orientation and localization related to the position of neuron bodies within nerve net regions. Indeed, imaging of a neuronal subtype with a *NvLWamide-like>mCherry* reporter line[52,178,201] showed that this neuronal subtype was integrated in multiple localization with stereotypical neurite pro-

jections numbers and anatomical orientation associated to their cell bodies localization, with unipolar or bipolar neurons with primarily oral-aboral neurite orientation located in neurite plexuses, and tri- and quadri-polar neurons with cell bodies in the epidermal compartment.

Even as their morphology has not been formally described or quantified across life stages except for the LWa subtype, the morphology of body wall neurons are usually described as composed of mostly sub-epidermal cells with a "ladder-like" organization of neuron bodies and neurites, with longitudinal neurites in the plexuses, and orthogonal circumferential neurites in the sub-epidermal layer in between plexuses (see schematic in chapter 1, from neural drivers data from[38,50,52,120,178,181,199] reviewed in[47]), in addition to sparser epidermal neurons with long neurites not aligned with neurite tracts. Consistently, throughout all *in vivo* imaging data from all three drivers here, sub-epidermal (and not epidermal ones) body wall neurons display this preferential orthogonal neurite positioning (see section 2.3, chapter 3 with associated supplementary videos available online at[240]). Furthermore, extending these descriptions, two anatomical neuronal subsets are detectable among non-epidermal body wall neurons. Indeed, across all imaging data for all drivers, we systematically observe that neurons whose bodies are basiepithelial but located between neurite tracts show a preferential circumferential neurite orientation orthogonal to the neurite plexuses, however, their neurites run above neurite plexuses in the epidermal/sub-epidermal mesoglea boundary layer and not through them (see chapter 2, chapter 3 and supplementary videos online associated to[240]). Meanwhile, neurons whose cell body are located within the neurite plexuses indeed display a preferential longitudinal neurite orientation, primarily within the neural plexus, with some neurites present in the inter-mesenteric area.

Whether these differences in neurite localization, which appear linked to the position of the cell bodies, are associated to distinct molecularly defined neuronal subtypes remain to be determined. Intriguingly, in young 2-week-old primary polyps, *NvLWamide-like*$^{+}$ neurons appear to primarily encompass either epidermal-embedded neurons or neural plexus neurons,

with the latter, as described above for most neural plexus neurons, showing longitudinal projections[177,178,201]. At that early polyp stage, this subtype represents less than a few dozens cells across the entire animal[178], and it is currently undetermined whether *NvLWamide-like*$^{+}$ neurons localized outside of the neurite plexuses might arise at a later stage, or if all *NvLWamide*$^{+}$ neurons are restricted to epidermal and neurite plexus localization, which would be indicative of molecular subtype regionalization. Future work investigating the anatomical localization of the molecularly described cell classes, for example using the cell-subtype-specific markers listed in Fig. 3.3 through hybridization chain reaction (HCR) *in situ* mRNA detection in polyps, in combination with transgenic reporters for live or fixed-sample imaging would help systematically characterize neuronal subtypes and their anatomical location as well as their neurite branching patterns. See below for further discussion on possible spatial regionalization of neuronal subtypes relative to nerve net subsets and body axes.

These observation support a stereotypical body wall nerve net organization, with cellular stability for at least a subset of its neurons and where their neurites branching and localization (which would determine their possible synaptic partners and neural circuits) are linked to the position of their neuronal bodies. Thus, understanding where and how new neurons integrate into these nerve net subsets in this continuous neurogenic process would provide insight into the cellular mechanisms underlying neural tissue renewal and expansion.

4.2.2 Lifelong growth and tissue renewal dynamics

Lifelong renewal of specialized cells in animals often relies on spatially-segregated lineage-restricted stem cells generating distinct differentiated cell types. Classical examples include the mouse intestinal epithelium and skin, where newly generated cells are progressively displaced away from the stem cell niche, and older differentiated cells are extruded off, preserving epithelial barrier integrity and maintaining tissue architecture. In these epithelial systems, local cell-cell contacts are preserved, with a spatial differentiation gradient across

the epithelial layer underlying cellular turnover. Some animals, often those with indeterminate growth, like planarians, acoels and cnidarians, among others, can renew all their specialized cell types throughout their lifetime, including their neural cells[241,275]. This trait is present across several lineages, distributed across the metazoan tree, relying on varying stem cell systems. In the cnidarian *Hydra*, which displays a simple bi-layered body plan with a distributed nervous system, a renewal strategy analogous to the epithelial stem cell systems mentioned above operates across the entire body, for all cell types, including neurons. Cells are continuously displaced along the primary body axis toward both extremities, and older cells are lost at the head and foot regions. This global tissue dynamics encompasses neural and secretory cells, which undergo position-dependent transcriptional changes as they are displaced[276,277].

In contrast to the cellular turnover of the entire cell complement in mostly epithelial renewing systems, lifelong renewal of neuronal tissues is often more limited across metazoans. As neural function relies on precise spatial relationships and synaptic connectivity between neurons, with neurites connecting to their targets at a distance, large-scale tissue displacement could present a fundamental challenge to maintain neural circuits integrity. Neural cells are often long-lived cells organized within stable neural tissues, with little to no cellular replacement throughout life, as seen in mammals. This can vary across lineages, with some vertebrates like the zebrafish *Danio rerio* exhibiting extensive adult neurogenesis, throughout brain regions, especially in its telencephalon[56]. Neural stem cells (radial glia cells) are organized in spatially defined progenitor pools that generate neurons that are integrated in nearby circuits. This addition is continuous, without disrupting tissue architecture and nervous system function.

Adult *Nematostella* exhibit lifelong, nutrition-dependent growth, with exponential size increase when food is abundant and shrinkage under starvation[166]. This nutrient-dependent body size and cell number scaling is accompanied by an increase in neurons and expansion

of its nerve net (see chapter 3 and [201]). In contrast to *Hydra*'s bi-layered body plan, the presence of mesenteries in sea anemones, heterogeneous internal tissue structures, make the hypothesis of global conveyor-belt–like tissue dynamics across the entire animal as seen in *Hydra* less likely. Especially as the distributed nervous system displays stereotypical nerve net condensations (neurite plexuses) present in double tracts at the base of each its mesenteries, containing both neuronal cell bodies and neurites bundles (see chapter 2).

How this distributed, but organized, nervous system expands during growth, when the animal can double its body length in a few days, is poorly understood. We outline several possible cellular mechanisms below and review the current experimental evidence in light of our data.

Overall, tissue growth can rely on distinct strategies: (*i*) distributed generation, where new cells are integrated throughout the organism (as seen in planarians, albeit with possible more localized progenitors in some areas), (*ii*) growth-zone–restricted addition, where progeny cells accumulate in regions nearby (like in the zebrafish telencephalon), or (*iii*) global conveyor-belt–like displacement with tissue-wide movement (as seen in *Hydra*, which also has local migration). These scenarios are not mutually exclusive, and stem or progenitor cells may be present locally in spatially restricted niches or originate from more distant locations and migrate actively or through passive tissue displacement. We discuss below experimental evidence to evaluate these competing models in *Nematostella* neurogenesis, examining whether the global nerve net architecture is stable, or part of a global tissue displacement dynamics, and where newly-generated neurons then integrate. These would help shed light on the mechanisms underlying adult neurogenesis.

4.2.3 Distributed nerve net scaling

Examining the localization of newly-generated green *Elav>KikGR*$^{+}$ neurons (chapters 2 and 3) highlighted their presence everywhere across the animal, and not restricted to a

specific region, including in both the neural plexus and inter-mesenteric body wall regions. Additionally, the similar proportion of green cells in the oral and aboral body wall areas further supports a model of a distributed nerve net expansion, without a spatially restricted neurogenic growth zone. We further find these new neurons to be incorporated throughout the anatomical subsets of the body wall nerve net, with green *Elav>KikGR*$^{+}$ cell bodies detected across both the neurite plexuses and inter-mesenteric areas (see section 2.5).

Overall, these data support a model of the body wall nerve net as a locally stable structure with a regionalized organization associated to neurite anatomical location, with the continuous addition of new neurons throughout, without a spatially-restricted neurogenic growth zone. Thus, experimental data do not support either a global conveyor-belt tissue displacement model or a restricted-growth-zone-addition model. As neurons are molecularly heterogeneous (see below), the current experimental data cannot formally exclude that some of these mechanisms might be at play in a restricted subset of neurons, and thus not detectable globally. And the presence of new neurons throughout the nerve net does not indicate whether these arise from local or migrating progenitors, or a combination of the two.

Furthermore, our work highlights that this neurogenic process concomitantly produces neurons with distinct molecular identities, quantitatively generated proportionally to existing cell types. This is consistent with previous data for the *NvLWamide-like*$^{+}$ subset, where neuron numbers showed a linear increase with polyp body length[201].

Furthermore, across all neuronal subtypes, our quantitative approach provides an estimate of the neurogenic rate under homeostatic growth conditions of around one new neuron generated per two pre-existing ones in under a week *in vivo* (see chapter 3). This very high neurogenic rate is on the same order of magnitude as estimates in planarians under homeostatic conditions, with the around 25% of neural cell turnover in a week evoked in[278], although the primary data it is based on is unclear. This very high rate in *Nematostella*

might appear surprising, as it suggests an exponentially additive process in fed conditions, but it is fully line with the global exponential growth rate in body size by area and total cell number quantified in[166], with a total polyp size and total cell numbers doubling time at under 3 days in *ad libitum* feeding conditions.

This neurogenic rate reflects the net output of neuron production and loss. Neuron production is the total generation of neuronal cells from undifferentiated progenitors, as well as possible conversion from non-neural to neural fates and, if considering individual neuron subtype production rate, interconversion between neuron identities. Neuron "loss" is the combination of cell death (like apoptosis) or other processes (like cell extrusion or transdifferentiation to a non-neural fate). The net difference between neuron production and loss can lead to overall nervous system net equilibrium, if the generation and destruction terms are similar, nerve net growth, if the generation term is higher, or shrinkage, if the decay is higher.

In *Nematostella*, there is a single reported case of neuron apoptosis of a specific neuronal subtype ($RPamide^{+}$) associated to the apical organ, a larval structure that disintegrates at settlement[183]. Outside of this one-time life-stage-transition-specific remodeling event, there is currently no evidence for apoptosis of neurons in homeostatic growth conditions. In young primary polyps, TUNEL (Terminal deoxynucleotidyl transferase dUTP nick end labeling) assays, detecting double-stranded DNA breaks generated during apoptosis, have not focused specifically on neural structures but recent data does not seem to highlight a specific signal there in uninjured conditions[279].

If there were ongoing apoptosis of existing neurons in the body wall nerve net, then these disintegrating neurons that are fragmenting would be present -and thus detectable- with our single-neuron-resolution imaging. Unless this is a rare event (which would represent a small to negligible contribution to overall nerve net homeostasis given the very high production rate), even a single digit percent apoptosis rate across neurons would amount to several cells

in each z-stack. Anecdotally, throughout hundreds of *in situ* live imaging data, we only ever observed less than 5 total occurrences of cells with a clearly neuronal morphologies and where neurites appeared somewhat discontinuous, possibly evoking cytoplasmic fragmentation, that may suggest some cells could be in the process of disintegrating. For reference, there are ca. 100-400+ cell bodies detected in each confocal imaging z-stack, with the total number of cell bodies detected across our untreated body wall imaging data being over 20,000. While we cannot exclude the possibility of cell migration/extrusion/remodeling events leading to cells ending up outside of any imaged regions (including in wider field imaging), continuous apoptosis at a rate within one order of magnitude of the neuron production rate under homeostatic growth conditions thus appears unlikely.

These observations and the published data overall support a net additive neurogenic process, with little evidence for widespread neuronal cell death under fed conditions. It is possible that this neurogenic process would shift in older animals, as despite its indeterminate growth, and asexual reproduction through budding of its aboral extremity (physal pinching), older animals (i.e. 6-month-old to 9 year+ in the lab, the maximum lifespan is unknown) do not seem to grow as fast. As neurons are not uniformly distributed in the animal, but organized spatially within subsets, and growth is anisotropic across the animal, with more lengthwise expansion (quantified in[166]), how this additive process scales different parts of the nerve net, across both feeding-induced growth and starvation-induced shrinkage, in unknown.

4.2.4 Cellular lineage and molecular identity of progeny neurons

Newly-generated neurons are integrated throughout the body wall nerve net (chapter 2 and chapter 3) and their generation appear to scale to the pre-existing identities. How these distinct identities are specified remain to be investigated, with the individual transcription factors identified for each subtype in chapter 3 providing targeted entry points for future studies. Importantly, if the overall tissue organization is fixed under homeostatic conditions,

and mature neurons themselves stay within their nerve net location, which the data currently supports, then a relationship between neuron localization and its molecular profile (and thus its subsequent neurite branching pattern and therefore connectome), could point towards a position-dependent cell-extrinsic model of cell-identity acquisition, similar to what is seen during spatial patterning of the mouse nerve cord for instance.

This raises the question of whether the transcriptionally distinct neuronal clusters identified by scRNA-seq reflect intermediate stages along a differentiation trajectory or mutually exclusive end states. Across drivers, the experimental approach used here supports that peptidergic neuron clusters detected in the body wall are not intermediate but distinct end states, and that peptidergic body wall neurons (neural classes 0/1/2, except for one, discussed below) are all generated from the same cellular lineage and pool of progenitors (further discussed in section section 4.3). Thus, global position-dependent transcriptional identities changes appear unlikely in this system, in contrast with what is seen in *Hydra* with its continuous global tissue displacement dynamics and position-associated neuronal identity changes[276,280,281]

The exception to this shared cellular lineage is cluster 22 in chapter 3, a neuronal subtype that expresses low levels of the *WFSa* neuropeptide precursor, but is otherwise uncharacterized. Indeed, its distinct differentiation trajectory is highlighted by the following: (*i*) this neuron subtype does not fit the green+orange to red *SoxC>KikGR*$^{+}$ cells proportional ratio detected across the other peptidergic neurons, as there is an excess of red progeny cells; (*ii*) its specific marker genes are not detected among cells that are part of the *Prdm14d*^{low} lineage, which encompasses all other peptidergic neurons, as seen from the analysis of our data combined with published RNA-Seq data from a *Prdm14d>GFP* driver[181]; and (*iii*) their relative abundance is not proportional across drivers, which is the case of all other peptidergic neurons (Neural 0/1/2). From the data, we do know that this cell type has transiently expressed the *SoxC>KikGR* reporter very early on its cellular history, earlier than for other

peptidergic neurons, pointing towards a more distant tip population, with an unknown intermediate maturation stage. It would be interesting to characterize this subtype further, using the specific effector genes (extensive list available in the supplementary data files associated to the published paper[240]). This would be especially informative as, despite the differences above clearly separating it from the other Neural 0/1/2 peptidergic neurons, this subset also expresses the same shared peptidergic gene module, which is otherwise associated with a shared cellular lineage.

There are other neural-like populations (e.g. labeled as Neural class 3 in chapter 3), which includes *FoxQ2d*$^+$ sensory cells[180], that are also distinct from the peptidergic neuron group. It is currently unknown whether these cells, with their basiepithelial sensory processes and body wall epidermis-exclusive localization, are neurons and what their role is. Similar experimental evidence show their cellular lineage is also distinct from the peptidergic neuronal lineage, as for neural cluster 22 seen above. Thus, it would be interesting to study if these may be part of an epidermis-restricted subclass, possibly generated from distinct sources.

Whether there is regionalization of the neuron subtypes within the body wall nerve net relative to either body axes or to the neuron position in a distinct anatomical location is currently unknown. There is evidence for some molecular regionalization between body wall, pharyngeal and tentacular neurons[8,53,177,178], with for instance some neuropeptides mostly specifically expressed in tentacular neurons, like *VHRa*$^+$ cells[100].

Within the body wall, currently described expression patterns for neural-specific genes do not appear to show clear oral-aboral or directive axis preferential regionalization. This is currently limited by the resolution of the data, which is often whole-mount *in situ* hybridization detection of mRNA localization in entire polyps at the primary polyp stage that may not provide enough cellular resolution across the spatial axes considered here. These include *NvLWamide-like*, *serum amyloid A*, *Aeq-like1*, *Aeq-like2*, *POUIV*, *RFa*, and others[8,50,52,101,177,201]. Further investigations into the spatial regionalization of neuronal

subtypes, especially in light of the recent development of spatial transcriptomics approaches for polyps would help shed light on whether there is regionalization within the body wall nerve net here[274]. This preliminarily suggests there is a radially symmetrical distribution of neuronal subtypes across the body wall nerve net, without clear oral-aboral polarity, and we further discuss the implications of this scenario relative to the progenitor cell source(s) in section 4.3 below).

4.2.5 Conserved cell-type but variable cell renewal: the case of cnidocytes

In contrast to the molecular, morphological and anatomical heterogeneity found across peptidergic neurons, cnidocytes in *Nematostella* display very limited diversity and follow a distinct differentiation trajectory. In the body wall, mature cnidocytes are embedded within the epidermal monolayer among epithelial cells, with their capsules oriented perpendicular to the epithelial sheet layer, as described in early life stages[186,188,193], and display short thin bifurcating cellular projections (neurites) forming a basiepithelial network, notably distinct from that of the neurons. Previous electron microscopy studies, like in the tentacles of the hydrozoan jellyfish *Coryne tubulosa* had noted thinner processes for nematocytes than for hair cells and other neurons[282]. The function and connectivity of these projections is unknown, as is how they develop.

Focusing on body wall cnidocytes, we examine whether they also show evidence for a spatially distributed renewal, or if, as in some other cnidarians like the jellyfish stage of *Clytia hemisphaerica*[145], *Nematostella* displays spatially organized maturation zones. Cnidocytes mature through a two-step process, with a morphologically, and molecularly completely distinct intermediate developing stage to build their capsule followed by a later neural-like maturation phase (chapter 3). This clear separation of developmental stages there enables the detection and spatial localization of early differentiating cnidocytes throughout the animal. To understand how cnidocyte renewal occurs in the body wall, we focused

on the localization of the developing cnidocytes. Due to the cytoplasmic localization of the KikGR fluorophore used in our transgenic lines, both developing cnidocytes (with their assembling bent capsules) and mature cnidocytes (with their straight rod-like capsules) are clearly differentially detectable *in situ* in live adult polyps (see for example Figs. 3.21 and 3.22 in chapter 3).

In *Nematostella*, we find developing cnidocytes present as isolated cells throughout the body wall area, displaying a rounder, irregular cell shape, and they occupy a basiepithelial position relative to the epithelial monolayer (see supplementary videos associated to[240]). As mature cnidocytes are epithelial, their change of localization during maturation further suggest that they undergo a mesenchymal to epithelial transition. Anecdotally, there is a molecular signature associated to both mature (but not developing) cnidocytes, epidermis epithelial cells, and some mucous cells (a subset of mucous cells is known to localize in the epidermis epithelial layer), separate from their cell-type-specific profile. We hypothesize this might be associated to their localization in the epidermis layer at the late (but not developing) stage; warranting further investigations. Thus, the specific presence of developing cnidocytes interspersed across the animal, across the body wall (see chapter 3), mesenteries and in the tentacles throughout the tentacular proximo-distal axis (unpublished observations), supports a distributed cnidogenic process, similar to what is seen with neurons, even as these cells are embedded in the epithelial tissue as mature cells.

This is consistent with several specific scenarios for their stem/progenitor source population, either a local generation from cnidocyte-competent cells or progenitors nearby, or migration from other sources (like the proposed *Piwi1*$^{+}$ cells in the mesenteries). Indirect evidence further supports a more local process: as part of the generation of a cnidocyte reporter line (with a memOrange2 fluorophore fused to a capsule gene), Sunagar *et al.* showed images of F0 mosaic transgenic polyps, which display body-wide oral-aboral longitudinal stripes of cnidocytes throughout their body, in primary polyps from the tip of the tentacles from the

aboral end and in contiguous stripes in tentacles in older 2.5 month-old polyps[194]. As these are mosaic animals, it suggests that the cellular sources of the newly generated cnidocytes are located in the same oral-aboral narrow stripe region where the cnidocytes end up in, and that there is no circumferential migration pattern from their cell origin source (it does not exclude migration from a mesenteric patch in the same area). Whether there are resident cnidocyte-specific stem cells throughout the animal, or these arise from a more broadly multipotent source or some epidermis epithelial cells might become cnidocyte-fated is further discussed in section 4.3 below.

Cnidocytes are a rare example of a cell-type that can clearly be traced along large evolutionary distances, and considered homologous across cnidarians. However, despite similarities in effectors genes, and in some regulators, data in *Nematostella* (in chapter 3 and in[188] and others) highlight that, despite this "homologous" cell-type, its cellular renewal can be quite variable from what has been described in other cnidarians, namely the distantly related hydrozoans *Hydra* (polyp) and *Clytia hemisphaerica* (jellyfish stage). In the hydrozoan *Hydra*, the interstitial stem cells (i-cells) localized in the ectodermal layer in the body column give rise to neurons, cnidocytes, secretory cells and gametes. Cnidocyte-committed i-cells divide into nests of 4/8/16 cells, connected by cytoplasmic bridges (first identified in 1959[283]), before forming their intracellular capsule structure[276]. These intracellular bridges were recently shown to be due to the co-option of a germline protein (*Kif23*) into developing cnidocytes[284]. These maturing cnidocytes nests actively migrate in the epithelial tissues towards the extremities of the animal, and can be embedded as part of epithelial battery cells. All cnidocytes are renewed in $\tilde{9}$ days. However, the repeated synchronized division of connected nematocyte-committed i-cells, followed by active migration as nests, and then initiation of their capsule program does not appear to be broadly shared across cnidarian lineages, with early reports in the 1960s indicating similar cellular nests were not present in a sea anemone[285]. Even in the more closely related hydrozoan cnidarian *Clytia hemisphaerica*,

at the jellyfish life stage, there a cellular conveyor belt system at the base of the tentacle bulb spatially segregating dividing stem cells, and developing cnidocytes in an ordered proximo-distal spatial differentiation gradient[145], unlike hydra' system. Thus, while cnidocytes share a common origin and effector genes (discussed further in section 4.4 below) despite the long evolutionary time between these extant cnidarians, their cellular mode of renewal does not appear to be fixed. This illustrates that developmental processes are plastic, with the same outcome, i.e. the specification of functional cnidocytes, arising through distinct strategies, even for a clearly identifiable and unambiguous "homologous" cell type.

Additionally, there is experimental evidence in both the hydrozoan *Hydra*, the colonial hydrozoan *Hydractinia echinata* and *Nematostella vectensis*[51,191,210,286,287], using a combination of pharmacological inhibition with DAPT (a γ-secretase inhibitor), as well as over-expression of the *Notch* intracellular domain (NICD) or *delta* ligand, showing that Notch signaling is involved in cnidocyte maturation across these species, with active Notch signaling acting (likely in *SoxB(2)*$^{+}$ progenitors[51]) as a negative regulator of neuron and cnidocyte fate specification. This might indicate that if the cellular localization of progenitors, migration and division properties might be variable across cnidarians, fundamental aspect of their differentiation, including regulators are nonetheless conserved. Critically, in all the cnidarian systems mentioned above, *Notch* signaling inhibition (through DAPT soaking) led to the accumulation of immature cnidocytes, with misshapen capsule, indicating that the initiation of the cnidocyte fate was not abolished, and initial capsule formation event occur, but that the maturation process is perturbed. The more recent genomic inventory of Notch pathway components across metazoans, including many cnidarians, further provides perspective on the possible requirement of this signaling pathway for cnidocyte capsule development. Notch pathway components are heavily conserved across extant cnidarians, except for myxozoans, a large group of microscopic parasites which have relatively recently been identified as part of cnidarians, due to the presence of a polar capsule; identified as a modified cnidocyte capsule,

which appear to lack many genes (like a delta/jagged ortholog, and the unclear presence of the *Notch* receptor itself)[288]. It remains to be determined whether the association between Notch signaling and cnidocyte capsule maturation is actually conserved across Cnidaria, which would support an ancestral role (with possible loss in myxozoans, although this has not been investigated, as they have many gene losses and adaptation linked to their reduced parasitic nature).

4.3 Anthozoan stem cell system(s) for lifelong neurogenesis

4.3.1 Identification of a *Piwi1*$^+$ progenitor pool

In this indeterminate growth system, distributed additive neurogenesis scales the nervous system, raising the question of the lifelong cellular source(s) of these newly-generated neurons. The main cell candidate pool to sustain cell renewal including neurons in *Nematostella* has been the *Piwi1*$^+$ putative progenitor cells described molecularly in[8]. These were annotated as such from their absence of known cell-type-specific markers and presence of genes associated with stem cells and highly proliferative cells across Metazoa (see section 4.3.5). We focus in this section on the cellular source(s) of neuronal cells, evaluating experimental data and the recent literature regarding (1) their potency and lineage-restriction, (2) their localization, within spatial niches or broadly distributed, and (3) their cycling and self-renewal dynamics.

By essence, static molecular snapshots of undifferentiated cells as provided by scRNA-Seq atlases cannot on their own resolve whether these cells differentiate into specialized cell types or determine their developmental potency. Several complementary reporter approaches, including ours, have been used to experimentally characterize these cells and their progeny[202,203]. By combining *in situ* detection of germline and stem cell markers (*Piwi1* and *Vasa2*) in adult *Nematostella vectensis*, with fluorescent transgenic reporter lines, Miramón-Puértolas, Pascual-Carreras & Steinmetz showed the existence of reporter-labeled somatic

cells that lack detectable *Vasa2* or *Piwi1* protein[202]. Several genetic reporters (e.g. a *Piwi1* reporter, an endogenous *Piwi1-mOrange* fusion protein, and a *Piwi1-P2A-GFP* reporter lines[202], and distinct *Piwi1* and *Nanos2* reporter lines[203]) further support, through fluorescent reporter inheritance, that somatic derivatives can arise from $Piwi1^{+}/Vasa2^{+}$ cells. Accordingly, the identification in our study of a pool of $Piwi1^{+}$ progenitor cells giving rise to multiple somatic derivatives is consistent with both scRNA-Seq predictions and reporter-based lineage tracing studies. Together, these data establish $Piwi1^{+}$ cells as a progenitor pool, but do not resolve what specific progeny they give rise to; what molecular and lineage heterogeneities may exist within this pool, and their associated spatial organization, and cell renewal dynamics.

4.3.2 Lineage restriction within the $Piwi1^{+}$ pool

Taken together, multiple lines of evidence (chapter 3), from reporter-based lineage tracing, scRNA-Seq characterization, in conjunction with analysis of published data indicate that the $Piwi1^{+}$ progenitor pool is not uniformly multipotent, but rather displays evidence of partial lineage restriction, with subsets contributing preferentially to specific neuronal and non-neuronal lineages. However, these lineage biases are not reflected in transcriptionally discrete $Piwi1^{high}$ populations at the current resolution. This includes the observation that despite differential progeny recovery, the distribution of $Piwi1^{high}$ cells populations between the *Elav>KikGR* and *FoxL2>KikGR* drivers is virtually identical [1] (see Fig. 3.14).

The presence in adults of $Piwi1^{high}/Elav{>}KikGR^{low}$ with an undifferentiated molecular profile, and a cycling transcriptional signature may appear surprising, especially as $Piwi1^{high}$ progenitors are proposed to represent a self-renewing "stem" cell pool. Indeed, ELAV-family proteins are classically described as post-mitotic neuronal markers across bilaterians, with conserved roles in the establishment and maintenance of neuronal programs. In the fruit

1. Except for the putative cnidocyte precursors, not present in $Elav{>}KikGR{+}^{+}/Piwi^{low}$ precursors

fly *Drosophila melanogaster*, *elav* is a well-established post-mitotic pan-neural marker. But it is also transiently expressed in proliferating glioblasts and in embryonic glial cells, in addition to neurons[289], indicating that its expression is not strictly restricted to terminally differentiated neurons.

In *Nematostella*, the regulatory element used to drive the *Elav>KikGR* reporter has previously been described as broadly neuronal and largely post-mitotic during early life stages (gastrula and planula). Experimental evidence indicate that these $Piwi1^{high}/Elav^{low}$ adult cells are not a continuously expressing populations, but rather that *Elav* expression is low (from *Elav* mRNA, *KikGR* mRNA from*Elav>KikGR* driver and KikGR fluorescence levels), and associated to differentiation initiation (reporter photoconversion chase), with a subsequent high and sustained expression in neuronal cells. Consistent with this, a recent nitroreductase-based cell ablation approach using the same *Elav* regulatory element[38] showed the elimination of *Elav*-expressing neurons[290]. Interestingly, $FoxQ2d^{+}$ sensory cells (Neural 3 class, not part of the same $Elav^{+}$ lineage) are still present, which is fully in line with other experimental data showing these not being part of the same $Elav^{+}$ lineage (overview in Fig. 3.19). However, this *Elav*-promoter-driven cell ablation experiment[290] cannot rule out that the detected $FoxQ2d^{+}$ cells post-ablation are old cells, and thus whether the newly developing cells are not affected (i.e. indeed arising from a distinct cellular lineage) would be more conclusive. Altogether, this represents a very promising system to experimentally refine lineages and cell potency. Of note, new *Elav*-positive neurons reappear in about 4 days, which is the neurogenic timeline we described in chapter 3.

Together with our *FoxL2>KikGR* and *Elav>KikGR* data and associated analysis of *Prdm14d>GFP* RNA-seq dataset[181], these experimental data indicate lineage restrictions within the $Piwi1^{+}$ pool despite the lack of transcriptionally discrete $Piwi1^{high}$ progenitor populations only detected in *FoxL2>KikGR+* cells, which show broader derivatives than the *Elav>KikGR* line. From the data, there is at least one peptidergic-exclusive lineage, as well

as evidence for a digestive secretory lineage as well, and we find distinct *Piwi*^{low} precursor states which express the same markers genes, one with pan-peptidergic markers, one with pan-digestive secretory markers, and one with pan-cnidocyte markers, supporting these as intermediate committed subpopulations.

Surprisingly, the relative peptidergic neuron subtypes proportions are the same across all drivers (with one exception, see section 4.2.4), despite drivers having diverging expression dynamics (i.e. mostly ON/OFF, OFF/ON/OFF and OFF/ON in the peptidergic neurons during their differentiation for the *FoxL2>KikGR*, *SoxC>KikGR*, and *Elav>KikGR* drivers). In addition to the lack of a clear transcriptional cycling signature across the different neuronal subtype populations, along with other experimental data (chapter 3), these data support that it is the same peptidergic cellular lineage that is sampled across all drivers, and that the distinct end states are post-mitotic terminally differentiated. For these peptidergic neurons, this argues against the existence of dedicated transit-amplifying cells, otherwise a hallmark of neuronal differentiation in bilaterian systems.

Thus, this more direct specification mode for peptidergic neuron subtypes may limit the diversity of molecular identities, but also sustain the high growth rates and fast differentiation timelines at hands. Interestingly, across distantly related animals, like in planarian flatworms[259,291,292], the hydrozoan cnidarian *Hydra*[249,280], the acoel *Hofstenia miamia*[293,294], and the annelid *Pristina leidyi*[295]), the organism-wide homeostatic renewal, including of their neural cells, is associated with a more limited diversity of neuronal subtypes generated. Across these systems, broadly non-neural exclusive multipotent cells renew cells with more neural-restricted fate. Current experimental evidence points towards a similar model in *Nematostella*, with a broadly multipotent *Piwi1*^{high} pool where lineage decisions occur and that renews more restricted *Piwi1*^{low} precursors, which display lineage-associated transcriptional signatures.

4.3.3 Niche-organized or distributed stem cell system(s)

Although *Piwi1*^{high} progenitors appear transcriptionally similar at the current resolution, the presence of lineage restriction within this pool and the distributed localization of new neurons and developing cnidocytes raises questions regarding the presence of one or multiple (stem) cell pools, their localization and self-renewal dynamics.

The question of the localization of the *Piwi1*$^{+}$ progenitor cells in *Nematostella* is surprisingly non-trivial. A knock-in fusion protein reporter line shows very widespread low-grade fluorescent reporter expression (1 in 2 cells in the juvenile, 1 in 5 in older adults), with high-reporter-expressing cells representing 4 per 10,000 cells in sexually immature adults and 4 per 1,000 cells in older adults[202]. As *Piwi1* antibody staining and mRNA detection show expressing cells within the mesenteries[202], the lowly *Piwi1-mOrange*$^{+}$ cells are interpreted as progeny cell (in addition with other experimental data). Detection of *Piwi1* mRNA by *in situ* hybridization[202] or hybridization chain reaction (as we showed in supplementary material for[240]) all indicate that this gene appears to be lowly expressed, and thus its expression pattern is hard to resolve; with increased signal detected within the mesenteries (mostly at the base of the cnidoglandular/ciliated tract area), alongside low-grade signal throughout the animal. Thus, *Piwi1*^{high} cells are likely rare, definitely present in the mesenteries, and possibly across the entire animal. Genes also expressed in *Piwi1*$^{+}$ progenitors, that could be used to determine the localization of these cells are overwhelmingly either expressed at very low levels, like *Piwi1*, hindering their unambiguous detection, or are also expressed across other cell types, limiting their usefulness as specific marker(s). Thus, a current model proposes that *Piwi1*^{high} cells reside deep within the mesenteries, at the base of the cnidoglandular tract, and then migrate into the body wall[202]. Consistent with this, mesenteries are required for regeneration from grafting experiments, suggesting that cells with broadly-multipotent to pluripotent properties are located there[225]. And transplanted genetically labeled cells from mesenteries (as well as other regions) can expand in the host

(from a pool of 10,000 to 100,000 cells)[255], and show enrichment of the same marker genes including *Piwi1*.

The characterization of stem cell population(s) in this system is still in its infancy, hence the "progenitor" label used throughout here, however, there is accumulating evidence from various experimental angles, that there exist cells that functionally fulfill criteria to be considered *bona fide* "stem cells". Briefly, animal stem cells are undifferentiated cells that (1) display defined potency (the ability to generate specific differentiated cell types) and (2) can self-renew (produce progeny that retain stem-cell identity). In this system, there is experimental evidence for long-term tissue contribution, clonality, self-renewal, capacity to regenerate cell lineages, with strong association with expression of a specific gene cassette including *Piwi1* (see below) even if their precise molecular and lineage heterogeneities are still under active characterization[202,225,255].

In both planarian and acoels, their collectively pluripotent stem cells are the only somatic proliferative cells, and, in planarians, neoblasts represent 1 in 5 cells. In *Nematostella*, a large share of *Piwi1-mOrange*^{high} cells are EdU positive[202,256], and *Piwi1*$^{+}$ cells display a clear transcriptional cell cycle signature (specific to the *Piwi1*^{high} subset) (see Fig. 3.9 and Fig. 3.14). More recently, cumulative EdU experiments under different feeding/starvation regimen show that when the nutritional conditions are optimal, >=96% of *Piwi1-mOrange*^{high} cells can accumulate EdU signal within a week, showing they are actively proliferative. However, there is also evidence for proliferation outside of the *Piwi1*^{high} cell subset; and that proliferation is distributed across the entire animal. Thus, *Piwi1*^{high} proliferative cells located at the base of the ciliated tracts in the mesenteries are unlikely to provide the full picture of this highly renewing system.

Indeed, active cell division (from short-EdU-pulse incorporation) is distributed across the entire animal[225], and *Piwi1-mOrange*^{low} cells, interpreted as mostly progeny cells of the *Piwi1*^{high} cells, also incorporate EdU (see supplementary figures in[256]). There is additionally

a weaker transcriptional cycling signature in a subset of epidermis epithelial cells (see also Fig. 3.9). This raises the question of whether additional self-renewing compartments exist outside the canonical *Piwi1*^{high} pool, and especially within the epidermal epithelium.

Additionally, several reports indicate the presence of molecularly uncharacterized EdU-label retaining cells after longer (1 month+) chase of the EdU label, i.e. proliferation-capable quiescent cells, throughout the animal, including at the base of the mesenteries close to the parietal muscles[203,225], some of which show *nanos2>mOrange* expression and have a smaller, rounder morphology[203]. *Nanos2* is part of the same molecular signature present in the *Piwi1*$^{+}$ progenitor pool. Intriguingly, we also observe small, round, morphologically undifferentiated cells with *Elav>KikGR*^{low} signal and *FoxL2>KikGR*^{low} signal present at the base of the neurite tracts, next to the parietal muscles, in the same area as the slow cycling *nanos2*$^{+}$ cells. Whether these might be locally resident, slow-cycling cells involved in neuronal cell renewal remain to be determined.

Together, these observations support a model in which neurogenesis in *Nematostella* rely on a broadly distributed population of proliferative and quiescent progenitor-like cells (putative stem cells), without a single spatially delineated proliferative niche. And the evidence presented in chapter 3 show that the *Piwi1*$^{+}$ progenitor pool is broadly multipotent, and not neural-exclusive, with lineage decisions likely occurring alongside the cell cycle, without evidence for transient-amplifying-cells, and a peptidergic-committed lineage.

4.3.4 Are there other self-renewing cell pools?

In *Nematostella*, it is unclear if *Piwi1*^{high} progenitors are a multipotent population that also give rise to epithelial epidermis cells, as seen in the colonial hydrozoan cnidarian *Hydractinia*[257], or if these are a separate self-renewing cellular compartment, as seen in *Hydra*, where there are three stem cells populations, an endodermal and an ectodermal epithelial stem cell population, as well as the multipotent i-cell lineage. More broadly, this raises the

question of the existence of several lineage-restricted stem cells populations that co-exist or if there is one pluri- to broadly multipotent cell pool that may give rise to all somatic (and germline derivatives), including both epidermis epithelial cells and neurons.

Focusing on this question, we note the large presence of *FoxL2>KikGR*$^{+}$ epithelial epidermis cells, representing almost 1 in 5 *FoxL2>KikGR*$^{+}$ sorted cell in the body (Fig. 3.5), while it is not the case with the *Elav>KikGR* or *SoxC>KikGR* drivers. This could be indicative of a lineage relationship, with *Piwi1*$^{+}$*/FoxL2*$^{+}$ cells giving rise to epithelial epidermis cells; but other experimental data (imaging and flow cytometry) indicates that these epidermis epithelial cells actually express low levels of the *FoxL2>KikGR* driver, which precludes identification of possible lineage relationships between these populations. The epithelial epidermis cells are homogeneously *FoxL2>KikGR*$^{+}$ orange cells, thus this does not provide information to determine whether they might be a separate cellular compartment. It remains undetermined whether the *FoxL2* gene is endogenously expressed in this population, given its low expression level in all cells except for tentacular spirocytes (as assessed by both scRNA-Seq and HCR detection of endogenous mRNAs). However, this weak *FoxL2>KikGR* expression also mirrors the weak continuous *Piwi1-mOrange* (fusion protein) signal also detected broadly across epidermis epithelial cells in[202], suggesting this compartment might be expressing low levels of *Piwi1* as well, (also detected by scRNA-Seq in our data here (see Fig. 3.9). *FoxL2>KikGR* is thus both lowly continuously expressed in *Piwi1*$^{+}$ progenitors, and also across epithelial epidermis cells, like the *Piwi1-mOrange* fusion reporter.

That some epidermis epithelial cells might be self-renewing is further supported by the enrichment of transcriptional cycling signature in a subset of these cells. This is also consistent with several reports of widespread EdU incorporation (showing active DNA synthesis) in uncharacterized cells throughout the epidermis tissue layer[203,222,225,279]. Overall, there is evidence supporting that a subset of epidermis epithelial cells may be proliferative. However, whether this is an independent cellular pool from the *Piwi1*$^{+}$ progenitor pool remains to be

experimentally demonstrated. In this regard, the low expression of *Piwi1* and *FoxL2>KikGR* across the epidermis layer precludes the determination of whether some epidermal epithelial cells can be renewed from the *Piwi1*$^{+}$ cellular pool or if this is a completely independent self-renewing compartment. In addition, the presence of large contiguous epidermis epithelial fluorescent patches in several mosaic F0 juveniles polyps[8,248] suggests that there is no widespread intercalation of newly-generated cells or extensive cellular rearrangements, at least at early juveniles stages, which would favor this compartment being a distinct self-renewing pool.

The activity of the *FoxL2>KikGR* reporter in epidermal epithelial cells also raises the possibility of alternative (or additional) cellular source for the several cell subtypes that are not part of the *Piwi1*$^{+}$/*Elav*^{low} lineage, i.e. cells that are not recovered in *Elav>KikGR*$^{+}$ cells, but are detected as progeny in the *FoxL2>KikGR* lineage. These progeny cells include developing cnidocytes, *FoxQ2d*$^{+}$ epidermal sensory cells, and mucous secretory cells.

Due to technical specificity of the experimental approaches used here, current reporters cannot fully resolve whether all developing cnidocytes mature from *Piwi1*$^{+}$ progenitors. And there is conflicting evidence about the presence or absence of *Piwi1-mOrange*$^{+}$ cnidocytes in the literature, which might have helped shed light on this issue: *Piwi1-mOrange*$^{+}$ cnidocytes are either detected[203] or not detected[202], but the transgenic lines used are distinct (ectopic reporter driver versus endogenous knock-in), so it might represent a technical detectability issue.

We know that cnidocytes are largely absent from *Elav>KikGR*$^{+}$ derivatives, and arise from *FoxL2>KikGR*$^{+}$ cells. But as both epithelial epidermis and *Piwi1*$^{+}$ cells are continuously *FoxL2>KikGR*$^{+}$ expressing, the experimental data is chapter 3 is compatible with developing cnidocytes arising from either *Piwi1*$^{+}$ progenitor cells, *Piwi1*^{low} epidermis epithelial cells, or a combination of the two. And our detection of a *Piwi1*^{low} precursor subset which expresses canonical developing cnidocyte transcription factors (and not recovered from

the *Elav>KikGR*$^{+}$ cells) is also compatible with all scenarios (see Fig. 3.14). As developing cnidocytes are present in the epidermis layer throughout the body and tentacles, as well as in the cnido-glandular tract at the tip of the mesenteries, the exclusive generation of these from epithelial epidermis cells appears unlikely. However, two other reporter lines we established, labeling epidermis epithelial cells, would experimentally addresses some of these aspects, especially combined with DAPI labeling of the cnidocyte capsule (with high DAPI concentrations emitting a GFP-like fluorescence following near UV excitation specific to the cnidocyte capsule and used to label mature cnidocytes), along with EdU incorporation detection.

4.3.5 A conserved germline program associated to multipotent stem cells

The pool of *Piwi1*$^{+}$ progenitor cells identified here is characterized by the specific expression of a conserved gene cassette (see Fig. 3.9) that has been termed the germline multipotency program (GMP)[296]. This germline multipotency program include members of the piwi/Argonaute family, DEAD-box helicases like *vasa* (*Ddx4* in mammals), and PL10, as well as other genes like *nanos*, *pumilio*, and *Bruno*. GMP proteins are often *nuage*-localized, i.e. in perinuclear membrane-less granules found in germline cells, also called polar granules in *Drosophila*, P granules in the nematode *C. elegans*, or chromatoid bodies in planarians. These have RNA-binding abilities and are involved in post-transcriptional gene regulation. In bilaterians, piwi-family genes are critical for transposon element silencing in germline cells and genome integrity.

This germline multipotency program appears to not be exclusively restricted to germline cells, and across multiple metazoan lineages, including in non-bilaterians, GMP proteins are also present in multipotent somatic stem cell, in addition to germline expression. In the freshwater sponge *Ephydatia fluviatilis* (Porifera, Demospongiae), the *piwi*-family gene *EfPiwiA* is expressed in archeocytes, cells that self-renew and can generate all sponge cell types,

and downregulated upon differentiation in most cell types[297]. Neoblasts, the collectively pluripotent stem cells of planarians (*Schmidtea mediterranea*, Annelida, Platyhelminthes) specifically express SMEDWI-2, a *piwi* protein, functionally required for homeostasis and regeneration[298,299]. In the hydrozoan cnidarian *Hydra*, both *Piwi*-like proteins (*hywi* and *hyli*) are expressed in its multipotent interstitial stem cells (i-cells) and in nematoblast nests (developing cnidocytes). These proteins are further localized in the cytoplasm, and enriched in perinuclear granules[300]. *Hydra* i-cells also express the piwi-family protein *CnPL10*, the nanos genes *Cnnos1* and *Cnnos2*, and the vasa genes *Cnvas1* and *Cnvas2*. *Hywi* knock-down in all cells but interstitial stem cells leads to epithelial disaggregation and is eventually lethal[300], likely due to its downregulation in the epidermal epithelial stem cell population. In addition to their robust gonadal tissue expression, sparse *vasa*-positive cells have also been reported in somatic tissues of ctenophores[301], and in the sea anemone *Exaiptasia diaphana* (Cnidaria, Anthozoa), with *Vasa*$^+$ cells distributed in tentacles, pharynx, mesenteries, and the pedal disc[302]. The pedal disc is an epithelial structure at the base of the polyp used for substrate attachment and this species can reproduce asexually from pedal disc lacerates, fully regenerating a mature polyp with functional gonads from a small (somatic) tissue piece. It is currently unknown whether the sparse uncharacterized *Vasa*$^+$ cells found in this area could be multipotent GMP-expressing stem cells, and whether they play a role in regeneration.

The current lack of genetic inducible knock-down or knock-out approaches in cnidarians, including in *Nematostella*, limits functional investigations into the role of these genes in their multipotent cell populations. Indeed, in *Nematostella*, early embryonic knock-down of *Piwi2* leads to developmental arrest[303,304]. These observations support the view that GMP expression, in addition to germline cell expression, is also pervasively found in undifferentiated cells, often with broad (somatic) potential. In both planarians and hydra, where broadly multipotent GMP-expressing stem cells are well-characterized, these cells can be restricted to somatic progeny, but, under injury, also replenish the germline lineage. Planarian neoblasts

can regenerate germ cells when those or gonadal structures are lost[305–308]. While *Hydra*'s i-cells have been historically described as multipotent stem cells that give rise to both somatic (neurons, cnidocytes, and secretory cells) and germline cells, experimental evidence now supports that the somatic and germline lineages are in fact distinct under homeostatic conditions[309,310], and that, under injury or germline loss, the somatic i-cell multipotent population can regenerate germline stem cells as well[309,310]. Overall, across distant animal lineages, GMP expression is associated with a conserved, undifferentiated stem-cell state compatible with multiple somatic and germline trajectories.

The pervasive identification of canonical germ plasm proteins in both germline and multipotent somatic stem cells across metazoans has prompted several authors to revisit the classical dichotomy of somatic lineage versus germline segregation. In[311], Solana proposed that animals have primordial stem cells (PriSC) defined as "*a group of multipotent or pluripotent cells that can have mixed germ/soma potential and that are always intercalated between the zygote and the PGCs during embryonic development.*". This proposed cellular lineage continuity has been hypothesized to underlie the conserved presence of germ plasm components in multipotent (somatic) stem cells across animals.

If gastrulation and the associated segregation of embryonic germ layers that imposes lifelong restriction of cellular potential is a hallmark of animal development, then the presence, across very distant animal lineages, of GMP-expressing multipotent cells capable of generating multiple derivatives (that are otherwise generated from distinct germ layers during embryonic processes), raises questions about the evolution of cell type specification mechanisms in animals, and its constraints. In several lineages, like in vertebrates, tissue-resident stem cells are specific from the same germ layer as the cells they replace. However, animals with extensive cellular renewal of all their cell types (including acoels, planarians and some cnidarians) often rely on pools of broadly multi to pluripotent stem cells. In acoels, the embryonic origin of these collectively pluripotent stem cells have been proposed to arise from

a single blastomere[312]. While the embryonic origin(s) of cnidarian stem cells is not known, with work in the hydrozoan *Clytia* proposing de novo generation of i-cells at the oral pole of the planula, in parallel to ectodermal non-i-cell early neurogenesis and cnidogenesis[313,314]. Thus, how these multi to collectively pluripotent cells are specified during early life stage remain to be further investigated.

Together, these observations support a model in which neurogenesis and cnidocyte renewal in adult *Nematostella* rely on a broadly distributed population of proliferative and quiescent progenitor-like cells, rather than a single spatially delineated stem-cell niche. Lineage restrictions appear to operate within an otherwise molecularly similar GMP-expressing $Piwi1^{high}$ pool. More work is needed to specifically identify these stem cell "pools" and how they are functionally related, and what signals regulate the balance between differentiation and self-renewal.

4.4 Two neural gene programs at the base of Cnidaria

4.4.1 Ontogeny and neural fate specification in *Nematostella*

Cell differentiation is intrinsically linked to cellular lineage, with the progressive acquisition of specialized gene programs from competent cells. In this adult neurogenic context, we experimentally characterize the diverging cellular lineages generating the two main neural cell classes of *Nematostella*, peptidergic neurons and cnidocytes. Combining lineage tracing information with transcriptional module analysis then highlights that the distinct peptidergic neurons share both a developmental lineage and a pan-neural gene program, not shared with cnidocytes which arise from a distinct cellular lineage, thus linking adult cellular lineage restrictions and identity acquisition. These lineage-associated cell-class-specific transcriptional modules (i.e, pan-peptidergic vs cnidocyte-exclusive) encompass both effector genes as well as transcription factors.

Mapping the presence of all transcription factors across both cell-lineage-associated and

cell-subtype-specific gene modules for neuronal and non-neuronal cell types highlighted general associations with classes of transcription factors. Namely, bHLH and bZIP transcription factors are specifically associated to neural cell classes with one set across all peptidergic neurons and one for developing cnidocytes, while Homeodomain-class transcription factors showed restricted subtype-specific associations, and both class-associated and subtype-specific zinc fingers genes as well. We have already discussed some of these aspects in the Discussion in chapter 3 in section 3.4, which would warrant further functional investigations into the homeodomain transcription factors selectively associated to neuronal subtypes here (see Figs. 3.28 and 3.30, and supplementary files published online with the paper for gene lists[240]). Additionally, in further support of the functional relevance of the transcription factors identified in these modules, the top pan-peptidergic-specific genes encompass a GATA gene, islet, and an ETS transcription factors (see Fig. 3.28). ETS and GATA motifs had been found to be enriched in open chromatin regions of *Elav>mOrange*$^{+}$ cells by bulk ATAC-Seq[8]. Recently, scATAC-Seq of gastrula and adult *Nematostella*[315] also identified islet/GATA motifs in a large subset of neurons; further supporting these genes as integral components of the regulatory landscape specific to this cellular lineage.

Furthermore, every transcription factor that has been functionally implicated in embryonic neurogenesis (the only current life stage where transient morpholino or shRNA knockdown approaches are available) is specifically present in each corresponding gene module here (see section 3.3.6 above). As neurons present during early life stages show overlapping molecular identities with adult ones, except for larval-specific apical-organ-associated neurons,[8], this suggests the re-use of embryonic gene regulatory networks for neural identity specification. This is preliminarily supported by recent single cell ATAC-Seq data across gastrula and adult life stages, highlighting similarities in cis-regulatory elements across both stages, although most neuronal subtypes are not resolved due to the sampling[315]. Linking specific transcription factors to the effector genes they regulate within defined cellular lineages is key

to understanding how distinct neuronal subtypes are specified. Increased scATAC-seq sampling to get to neuronal subtype resolution (or the use of complementary cell-type-specific chromatin assays targeting active enhancers, such as H3K4me1/H3K27ac profiling), would help refine the landscape of the cis-regulatory elements (CRE) used for distinct effector genes in each subtype, as well as whether the pan-peptidergic genes are regulated using the same CREs.

However, these approaches rely on accurate descriptions of transcription factor binding sites. This is particularly limiting in *Nematostella*, where many transcription factors lack well-characterized DNA-binding motifs. This is especially problematic for zinc-finger transcription factors, and we find these expressed specifically across both cell-class- and cell-subtype-specific transcriptional modules in our data. In animals, C2H2 (Krüppel-type) zinc-finger proteins constitute the largest and most diverse class of transcription factors, with lineage-specific expansions[14,15]. Their binding motifs are not well-characterized, as the binding specificity is determined by the number and relative spacing of their (numerous) zinc finger domains, and can vary greatly.

Overall, integrating lineages, transcription factors presence, cis-regulatory elements accessibility and effector gene expression data would help define what effector sets might be combinatorially regulated by multiple transcription factors depending on the cell-type context, which is critical to understand cell-type specification. Additionally, these questions are especially interesting as while cnidocytes and peptidergic neurons mature through two distinct lineages and use largely distinct effectors, they do share a smaller convergently expressed neural module (which includes some transcription factors). Examining how this shared program is regulated, whether through these shared transcription factors or distinct ones would be informative.

4.4.2 The divergent late cnidocyte neural program in *Nematostella*

The neural program of cnidocytes is largely distinct from other neurons. But even as the neural genes used differ, both programs rely on common neural effector families (including members of voltage-gated ion channels, ligand-gated ion channels, metabotropic (GPCR) glutamate receptors families among many others). One class of neural effectors appears conspicuously missing from the cnidocyte neural program: those associated with neuropeptide signaling. This is significant as cnidarians nervous systems are thought to rely heavily on peptidergic neurotransmission, with the neuropeptide complement of *Nematostella* under active characterization[98,100,101,105].

Peptidergic-signaling associated genes include over a dozen neuropeptide precursors, and specific neuropeptide processing machinery enzymes (prohormone convertases, amidating enzymes, carboxypeptidase), as well as, on the receiving end, their recently deorphanized cognate receptors (GPCRs)[92,98,100,101]. Each *Elav*$^{+}$ neuronal subtype (Neural classes 0/1 in chapter 3) is uniquely defined by neuropeptide precursors, most of which are specific to one neuronal subtype only, hence the "peptidergic neurons" designation throughout as these neuropeptide-associated gene are absent from late cnidocytes. Neuropeptide precursors and neuropeptide processing enzymes are highly expressed genes, which exclude technical detectability issues in late cnidocytes. And because the deorphanized neuropeptide GPCR receptors are also not detected in late cnidocytes, but specific to peptidergic neurons (and a few non-neural types), this preliminarily supports that neuropeptide-GPCR signaling may be restricted within peptidergic neurons, and not used between peptidergic neurons and cnidocytes. This pattern appears quite specific, although we cannot exclude that the possibility of currently uncharacterized peptidergic genes existing and being in fact expressed in the late cnidocyte program. The overall absence of currently known peptidergic precursors and receptors in late cnidocytes, in contrast to their clear presence across other neurons, raises questions regarding their mode of communication and the possible conservation this

non-peptidergic cnidocyte neural program across other cnidarians (see section 4.4.3 below).

Cnidocyte discharge is a one-time event, extremely fast and tightly regulated, with the integration of chemical and mechanical stimuli[123]. In *Nematostella*, cnidocytes in the body wall have short branching neurites forming a network between the epidermal layer and the sup-epidermal compartment. These are visible here with the *SoxC>KikGR* and *FoxL2>KikGR* reporters due to the cytoplasmic nature of the KikGR fluorophore, and also in[8,130,316,317]. Their integration into neural circuits and synaptic partners are currently undetermined. Early ultrastructural studies from other sea anemones (*Exaptaisia*) and scyphozoan jellyfish indicate that at least some cnidocytes can be integrated into neural circuits, with synapses (uni and bi-directional) between neurons and nematocytes reported in some species[144,318,319]. Both dense-cored and clear vesicles, with their content uncharacterized, have been reported. Cnidocyte discharge has been linked to specific ion channels from several families (including voltage-gated calcium channels, and TRP and Piezo channels in hydrozoans), but how this process is regulated remain poorly characterized. This raises the intriguing question of how much of cnidocyte function arises from parallel functional evolution relative to other neurons. And more broadly, as cnidocytes are the defining synapomorphy of Cnidaria, whether these two distinct gene programs are also present across other cnidarians has implications for our understanding of early neural evolution and neuronal subtype diversification. Indeed, neuropeptide precursors have undergone likely independent diversification events at least across anthozoans and hydrozoans (see overview in chapter 1).

4.4.3 Transcriptional cell-type similarities across cnidarians

The secondary neural program of cnidocytes in *Nematostella* is largely distinct from that of other neurons. As cnidocytes are considered homologous cell types and appeared early in cnidarian evolution, if this secondary neural program were both conserved across the phylum and distinct from the other neuronal programs, this would support a scenario with

at least two neuronal programs present at the base of the phylum. As this first split within Cnidaria occurred early after their divergence from Bilateria (see Fig. 1.1 in chapter 1), this has implications for our understanding of early neural cell type evolution, and how these diversified across lineages since the last common eumetazoan ancestor.

In the Appendix (section A.1), we show cell-type transcriptome comparisons between distantly related cnidarians spanning the first split within Cnidaria. These comparisons highlight transcriptional similarities between developing cnidocytes across all species (as expected), as well as between their secondary neural programs. The species compared there are very distantly related to *Nematostella*, with one anthozoan octocoral *Xenia sp.*[142,143] (last common ancestor with *Nematostella* at least 438 Mya), and two even more distantly related hydrozoans *Clytia hemisphaerica*[320] and *Hydra vulgaris*[138,249], thus spanning the earliest split within the Cnidaria phylum (see Fig. 1.2 in the Introduction). Across such a large evolutionary distance, akin to the divergences between bilaterian phyla, these comparisons identify around 8 transcriptional programs with some degree of conservation, covering developing cnidocytes, late cnidocytes, neurons, mucous secretory cells, digestive secretory cells, gastrodermis cells, muscle cells, and progenitor cells, consistently with described cnidarian cell types across the literature. These comparisons imply that in addition to their developing cnidocyte program, the secondary neural program of cnidocytes is both *ancestral to Cnidaria*, and that it was *already neural-like but distinct from that of other neurons* at that time (i.e. >540 Mya). We discuss further in section 4.4.4 below the implications for evolutionary scenarios of neuronal cell types evolution at the base of Cnidaria and in eumetazoan ancestors.

It may seem surprising that an earlier cross-species comparison[142], which included mostly overlapping atlas data, did not identify this conservation of a secondary cnidocyte program. However, in that study, the secondary program of cnidocyte was annotated only as another neural transcriptional state (which it also is), with only the developing program of cnidocytes

recognized as a cnidocyte state[142]. We further propose in the Appendix section A.1 that one of the cell states from the octocoral *Xenia* (unidentified in the original publication and identified as a neuronal state in[142]) may correspond to late cnidocytes. Other studies across cnidarians also overlooked this, likely due to different aspects. Indeed, cell atlases in hydrozoans identified the transcriptionally distinct developing cnidocyte and mature cnidocytes states, as in *Hydra*[249], *Clytia hemisphaerica*[320], and more recently in *Hydractinia*[140,141]; but did not identify that mature cnidocyte state as a neural program as it appears distinct from that of the other neural cells. Across the other cnidarian subphylum, Anthozoans, cell atlases identified the developing cnidocyte state as corresponding to cnidocytes, in *Nematostella*[8], *Xenia sp.*[143], *Stylophora pistillata*[142], but the late cnidocyte program appears to have been identified as a neural state (which it is), without association to cnidocytes. In the case of *Nematostella*, another atlas described the biphasic maturation strategy of cnidocytes based purely on the scRNA-Seq 2D projection of the data, but did not identify that secondary program as neuronal[53]. And in the case of *Nematostella*, direct atlases comparisons were further complicated by the separate genome and gene annotations used (also distinct from the more recent assembly and gene models used in chapter 3). Entire cell type transcriptomes were not available prior to scRNA-Seq approaches, which also explain this unidentified relationship, despite cnidocytes being known as a phylum-restricted innovation, and 50+ years of literature featuring electrophysiological studies (mostly in hydrozoans), and electron microscopy studies indicating that cnidocytes are mechanosensory cells, with a sensory cilium structure, capable of generating action potentials, and showing electrical coupling with other cnidocytes at a distance, with thin neurite-like cellular outgrowth, sometimes with vesicles detected near putative synaptic sites. And the ambiguity around cnidocytes is still ongoing, with the latest atlas of neuronal cell states in *Hydra*[280] only describing the non-cnidocytes neural cells, and recent scATAC data in *Nematostella* identifying the developing cnidocyte state, with several unlinked neural states that are likely to include mature cnidocytes[315].

4.4.4 Evolution of two neural developmental programs at the base of Cnidaria

Overall these comparisons indicate that the biphasic maturation of cnidocytes appears conserved across distantly related cnidarians, with a distinct end neural program compared to other neurons. As cnidocytes and other neurons arise through distinct cellular lineages (experimentally demonstrated here in *Nematostella* in chapter 3, known for *Hydra*, assumed for *Clytia*[320], and suggested by the atlas UMAP representation in *Xenia* in section A.1), this further supports that neural differentiation and acquisition of distinct neuronal gene programs is constrained by cellular developmental history, which may thus be much more evolutionary conserved that the physical organization of stem/progenitor/progeny cells in tissues (see Discussion in section 4.2 above). How did these two neural programs evolve at the base of Cnidaria remain an open question.

Several regulators of both neuron development and cnidocyte development can be traced back to a eumetazoan ancestor, including members of the Sox B, Sox C, *atonal/neurogenin* and *achaete-scute* families (reviewed in chapter 1). This is interesting as the initial gene program of cnidocytes is primarily non-neuronal, with high expression of cnidarian-specific capsule-building genes that are then cleared off prior to expression of the secondary program.

Furthermore, transcription factors with neuronal specification functions in bilaterians have been associated with regulating the early (non-neural) phase of cnidocyte maturation as well. In *Nematostella*, targets of *POUIV/Brn3* have been identified from comparisons between constitutive KO and WT animals with bulk RNA-Seq. These include both genes of the capsule-building program and of the late cnidocyte program (data from[120], further analyzed in Fig. 3.24). This homeodomain transcription factor is known for neuronal specification roles in bilaterians, and is cell-autonomously required in *Nematostella* for cnidocyte maturation[120], with accumulation of blocked immature cnidocytes in mutants (and *POUIV* also regulates mechanosensory hair cell development[196]).

In addition, other regulators of cell differentiation which often play roles in neuronal spec-

ification have been implicated in the early maturation of cnidocytes. Work in *Nematostella* had identified that while *Lsd1* (Lysine-specific demethylase 1, an epigenetic modifier) is broadly expressed, its knock-out specifically induces cnidocyte maturation defects, while the initial specification of immature cnidocyte is conserved[195,198]. This phenotype is similar to a knock-out for another component of the CoREST complex, broadly involved in cell differentiation processes, usually mediating gene repression. To what extend these may participate in the repression and downregulation of cnidocyte capsule genes and switch to the neural program of late cnidocyte remain to be characterized.

The differentiation process of cnidocytes involves a separate cellular lineage and regulators that are associated to neurogenesis in Bilaterians (reviewed for *Hydra* in[48]), like *POUIV* or *Lsd1* in *Nematostella*, as well as many cnidarian-specific transcription factors, like the zinc finger *ZNF845/CnZNF1* (not an ortholog of bilaterian ZNF845), in *Nematostella*[136], and *Hydra*; *PaxA*[188], *NvSox2*[193] (not a SOX2 ortholog, nor a SOX B family gene), the cnidarian specific *cnidoFos* and *cnidoJUN*[194], among others. Initiation of this cnidocyte neurogenic specification thus relies on both upstream regulators shared with other neurons (like *SoxB(2)*, *Atonal-related_Nem10*, etc..), and cnidocyte-specific regulators, many of which have arisen as cnidarian-specific duplications.

Recruitment of "neural genes" to the developing cnidocyte program

Tracing the origin of the genes used in this developing cnidocyte program to shed light into its evolution highlights that these are often taxonomically-restricted, with origins at the base of Cnidaria for most currently characterized capsule-building genes[8,118]. But, in addition to these, and despite the broadly "non-neural" program of this early differentiation phase, there are several genes usually associated with neural programs in bilaterians specifically expressed in the developing cnidocyte gene program in *Nematostella*, and not in late cnidocytes, highlighting the composite origins of this initial phase.

Indeed, the presence of putative vesicular glutamate transporter in developing cnidocytes (and not in later mature stages) in *Nematostella* (see also in Fig. 3.23) may be functionally associated to the presence and loading of glutamate during cnidocyst formation, initially built from a Golgi vesicle. Glutamate gets polymerized as poly-γ-glutamate in the capsule and its high concentration in the mature capsule is thought to provide the high osmotic pressure associated to explosive release and harpoon function[125]. The enzyme for the synthesis of poly-γ-glutamate, the poly-γ-glutamate synthase, was shown in the hydrozoan *Clytia hemisphaerica* to have arisen through horizontal gene transfer from bacteria[321]. And a putative vesicular glutamate transporter is present in developing cnidocytes (nematoblast stage) in the hydrozoan *Hydra* and not expressed in mature nematocytes or neurons, as shown in the recent *Hydra* nervous system atlas[280]. However, while both nematocyte-specific and distinct neuron-specific vGluT-like genes are also described in *Clytia*, these were noted to lack an arginine residue required for its function and presence in canonical bilaterian vesicular glutamate transporters[320], thus their function remain unclear.

In addition, there are several unclassified prohormone convertases, identified as largely independent-expansions from the canonical bilaterian PCSK-like subfamilies[179], expressed in the developing cnidocyte program of *Nematostella*, and not in its late cnidocyte program (see Fig. 3.23). These enzymes are endopeptidases that are often used as part of proteolytic cascades to generate functional short peptides like neuropeptides, with distinct prohormone convertases paralogs expressed in peptidergic neurons (see Fig. 3.3). We speculate here that these could potentially play a role in the cleavage of capsule components, as all minicollagen sequences are structured with a N-terminal signal peptide, followed by a propeptide sequence and ending with a KR endopeptidase cleavage site, and the latter is being canonically recognized by PCSK prohormone convertases. Among exopeptidases, expression of a carboxypeptidase specifically in cnidocytes (now annotated as developing cnidocytes) has been noted in the first *Nematostella* atlas[8]. The functional roles of these proteases remain

to be determined.

Thus, the initial cnidocyte gene program encompasses both taxonomically restricted cnidarian-specific novel genes, but also genes from other sources, including from cnidarian-specific diversification of gene families otherwise used in neural programs. In that specific case, these "neural"-related genes do not show sustained expression in the secondary neural program of cnidocytes. Overall, these elements support a scenario where, at the base of Cnidaria, a combination of novel genes built using pre-existing protein domains, along with a horizontally transferred gene, and pre-existing glutamate transporter and paralogs of prohormone convertases were used together to build a new intracellular structure: the cnidocyst. This indicate a mosaic origin of this novel neural cell type, associated to a distinct developmental lineage, using with regulators include both cnidarian-specific genes, whose emergence can be traced to the base of Cnidaria through its conservation in extant cnidarian species.

Diversification of neural gene programs across cnidarians

Regarding the diversification of these neural programs, across cnidarians currently sampled, cnidocytes show little transcriptional diversity with 1 to 3 broadly similar developing cnidocyte programs (and corresponding late neural programs) despite historical descriptions of many subtypes based on capsule morphology. These developing cnidocyte programs are broadly similar across species, except for the spirocyte, which is restricted to some anthozoans and known to rely on anthozoan-restricted paralogs of capsule genes. In contrast, there is much greater transcriptional diversity across neurons (i.e. non cnidocytes neural cells), with between 8 (for the well-sampled hydrozoan *Hydra*[138,249,280]) to 25+ neuronal states (in *Nematostella*[8,53] and chapter 3). There is little one-to-one cell-type similarities between anthozoans and hydrozoans for neurons, echoing what was described in early comparisons[142]. Within hydrozoans, which are the most closely related, some one-to-one or few-to-few com-

parisons could be detected (in section A.1). Overall, this suggests extensive lineage-specific diversification within cnidarians, specifically for neurons.

In *Nematostella*, the molecularly distinct peptidergic neurons appear to differentiate through a direct process (see chapter 3), in contrast with the biphasic maturation of cnidocytes. This is reminiscent of the also more direct differentiation seen in *Hydra* neural specification, with neural cells specified directly from i-cells, going through a final cell cycle, and dividing once to produce two cells that then differentiate into neurons, giving rise to a limited number of transcriptionally distinct states[322].

In *Nematostella*, most, but not all, neuronal states share a restricted peptidergic-specific cellular lineage and this cellular lineage is further associated to a shared gene program, with is combined with individual subtype-specific genes to make up their transcriptional program as seen in chapter 3. Subtype-specific genes include neuropeptide precursors, which encompass multiple largely unrelated gene families across cnidarian lineages (see chapter 1 for overview). Thus, overall, the specification of distinct neuronal identities across cnidarian lineages might be associated to cnidarian-lineage-specific diversification events of gene families, which has been noted for several "neural effector" genes families. Examples include the mechanosensory-associated ion channel family transient-receptor-potential N protein (TRP-N), with independent diversification within anthozoans and hydrozoans[323], the KCNQ voltage-gated K^+ channel family[324], or class A GPCRs[96,98] among others. There are also mostly uncharacterized expansions of cys-loop ligand-gated ion channels in cnidarians, as inferred from comparisons between two anthozoan genomes and other metazoans[20].

4.5 Summary

Together, these findings provide a comprehensive foundation for understanding adult neural differentiation in the sea anemone *Nematostella vectensis* (Cnidaria, Anthozoa), and suggest a more complex evolutionary scenario for neuronal subtype evolution and diversification at

the base of Cnidaria.

In brief, we investigated here:

(1) How is adult neurogenesis spatially and temporally organized, and are there spatially restricted neurogenic niches?

Chapter 2 develops in vivo imaging, photoconversion, and genetic reporter-based lineage-tracing approaches to enable temporal tracking of neural cell birth at single-cell resolution. Using these tools, Chapter 3 shows that neurons are generated in a distributed manner across nerve net territories and along the primary body axis, thus underlying nerve net expansion. Chapter 4 interprets these findings as supporting a distributed, rather than niche-restricted, additive mode of lifelong neurogenesis, across all body wall neurons and cnidocytes, and compares this organization to other cnidarian and metazoan tissue renewal strategies. This further highlights the variability of cellular renewal mechanisms across large evolutionary distances, even when gene regulatory programs appear conserved.

(2) What are the cellular origins and lineage relationships of newly-generated neural cells?

Chapter 2 extends live nerve net imaging to molecular characterization by combining reporter-based lineage tracing with targeted single-cell RNA-seq with recorded reporter status. With sampling from all three reporters (*Elav>KikGR*, *SoxC>KikGR* and *FoxL2>KikGR*) generated in chapter 2, chapter 3 shows that the molecular identities of newly-generated neurons are specified proportionally to existing ones. Furthermore, this approach uncovers that the distinct drivers largely sample the same cellular renewal processes for peptidergic neurons (the largest neural cell class), and digestive secretory cells. In contrast, experimental data indicate that cnidocytes arise from a separate lineage from both neuronal and secretory cells. This further identifies a pool of progenitor cells with a GMP signature that differentiate into somatic progeny cells with a peptidergic neuron-restricted lineage, also highlighting the absence of a neural-exclusive bipotent peptidergic neuron/cnidocyte progenitor state. chap-

ter 4 discusses whether the *Piwi1*$^{+}$ progenitor population constitutes a multipotent stem cell system, as well as its cellular localization, molecular profile and evaluates the degree of lineage restriction within this pool and similarities with other stem cell systems. This supports a model of hierarchical cellular lineage renewal from a distributed broadly multipotent pool of stem/progenitor cells.

(3) How are neural cell identities specified across major neuronal cell classes (peptidergic neurons and cnidocytes)?

Chapter 2 and chapter 3 show that newly generated neurons are integrated into a stable nerve net architecture throughout anatomical regions, with neuronal identities scaled proportionally. Integrating reporter expression data, chapter 3 provides experimental evidence for the direct generation of distinct peptidergic neuron subtypes, highlighting that neuronal states occupy separate terminal branches rather than sequential stages of a temporal differentiation cascade. This work further identifies a core peptidergic gene regulatory program linked to a common peptidergic-exclusive cellular lineage, alongside subtype-specific gene modules for all peptidergic neurons. In contrast, cnidocytes arise from a separate cellular lineage and undergo a biphasic transcriptional maturation, with their neural program only secondarily initiated and largely distinct from that of peptidergic neurons. chapter 4 discusses how cellular lineage architecture shapes the acquisition of distinct identities across neural classes and further examines how transcription factor classes are differentially associated with distinct cellular lineages, and the corresponding usage of shared neural class programs versus subtype-specific regulatory modules across neural cell types.

(4) How do these neural developmental trajectories and transcriptional programs compare within cnidarians? What does adult neurogenesis in this anthozoan cnidarian reveal about the evolutionary origins of neural diversification?

Chapter 3 identifies two neural gene regulatory programs corresponding to peptidergic neurons and late cnidocytes, associated with two distinct cellular lineages and regulatory

modules, with limited molecular overlap. Cross-species comparisons in Appendix A.1 support conservation of this biphasic mode of cnidocyte maturation with a secondary neural program, distinct from other neural cells, across distantly related cnidarians. Thus, chapter 4 proposes that these findings support both an ancestral biphasic mode of cnidocyte specification and an early divergence of at least two neural programs (cnidocyte and non-cnidocyte) at the base of Cnidaria, and discusses implications for evolutionary scenarios of neural cell type diversification.

APPENDIX A

CNIDARIAN CELL-TYPE COMPARISONS

A.1 Cell-type comparisons across distantly related cnidarians

Whether the two divergent neural programs identified in *Nematostella* in chapter 3 (for peptidergic neurons and cnidocytes) are also conserved across distantly related extant cnidarians would help inform our understanding of early neural cell type evolution and diversification. We examine here whole-transcriptome cell-type-similarities across distantly related cnidarian species.

Within the phylum Cnidaria, the earliest lineage divergence separates medusozoans, including hydrozoans, staurozoans, scyphozoans and cubozoans, (and parasitic endocnidozoans), from anthozoans (corals, sea anemones, sea pens). This split dates back to at least 500 Mya, and occurred early after the divergence from the last Cnidarian-Bilaterian ancestor[3,24] (see Figs. 1.1 and 1.2 in chapter 1), similarly to the earliest split within bilateria, between protostomes and deuterostomes.

To examine the conservation of cell-type transcriptional programs that might date back to that early cnidarian ancestor, we used SAMap[9] to score transcriptional similarities between cell types from atlases data from four extant species spanning the two branches of the earliest cnidarian split, the sea anemone *Nematostella vectensis* (Anthozoa, Hexacorallia), the coral *Xenia sp.* (Anthozoa, Octocorallia), a hydra (Medusozoa, Hydrozoa) and the jellyfish *Clytia hemisphaerica* (Medusozoa, Hydrozoa) (Fig. A.1) (see below for data sources and details).

Specifically, we expect some transcriptional similarities across the cnidocyte capsule programs, as several capsule forming genes have been identified across at least anthozoans and hydrozoans, with some expansions and sequence divergences shared across these distantly related lineages; as well as across known cnidarian cell classes like gastrodermis cells, neurons as a whole, etc. And previous comparisons highlighted similarities across cell states, but did

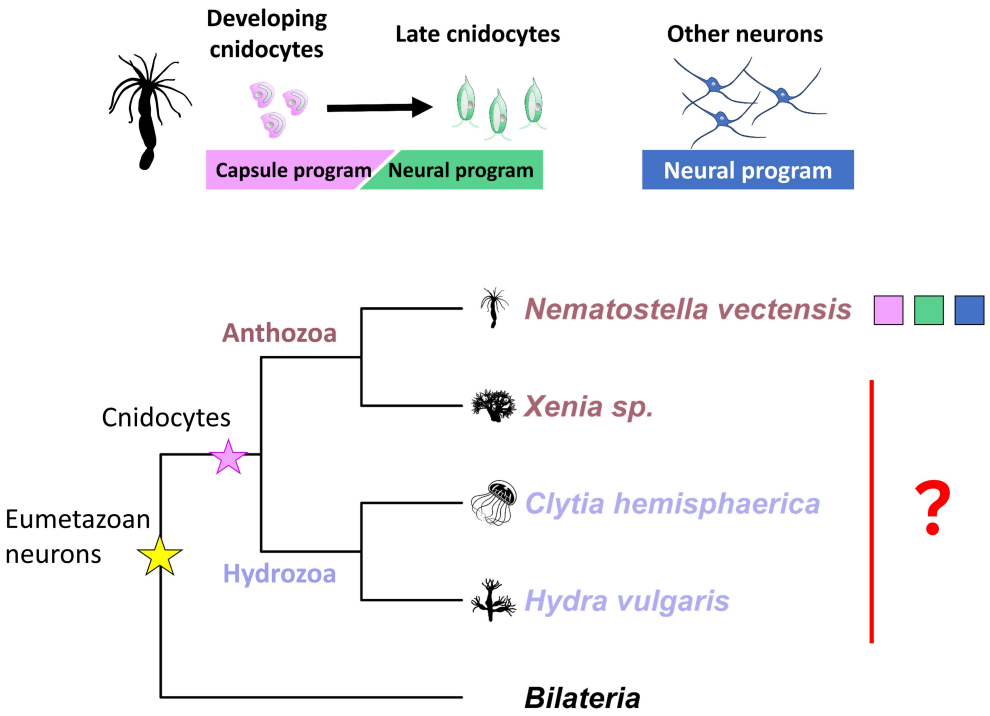

Figure A.1: Neural programs comparisons across distantly-related cnidarians

Distinct neural programs for neurons and cnidocytes in *Nematostella* (see chapter 3. Using extant cnidarian atlases to investigate the evolutionary conservation of these neural programs, spanning the earliest split within Cnidaria (Anthozoa, in maroon, versus Medusozoa, in light purple) to understand the evolutionary history of the secondary neural program of cnidocytes and neural programs in cnidarians. Silhouettes from phylopic.org, except for *Nematostella*: Colette Plessier from F.P. draft.

not specifically identify a developing and late cnidocyte cell state, (which was logical given knowledge at the time) and thus could not compare the latter specifically to other neural cell states[142]. See section 4.4 in chapter 4 for overview.

The scRNA-Seq atlases datasets (Figs. A.1 and A.2) are the following:

- *Nematostella vectensis* (class: Hexacorallia, order: Actiniaria, family: Edwardsiidae) scRNA-Seq data includes raw data originally published in [8], remapped to the jaNemVect1.1 genome and NCBI gene models, as well as *Elav*>*KikGR*$^+$, *FoxL2*>*KikGR*$^+$, SoxC>KikGR$^+$ body with mesenteries cells (described in chapter 3) and additional tentacle cells specific *Elav*>*KikGR*$^+$, *FoxL2*>*KikGR*$^+$, SoxC>KikGR$^+$ libraries (generated and pro-

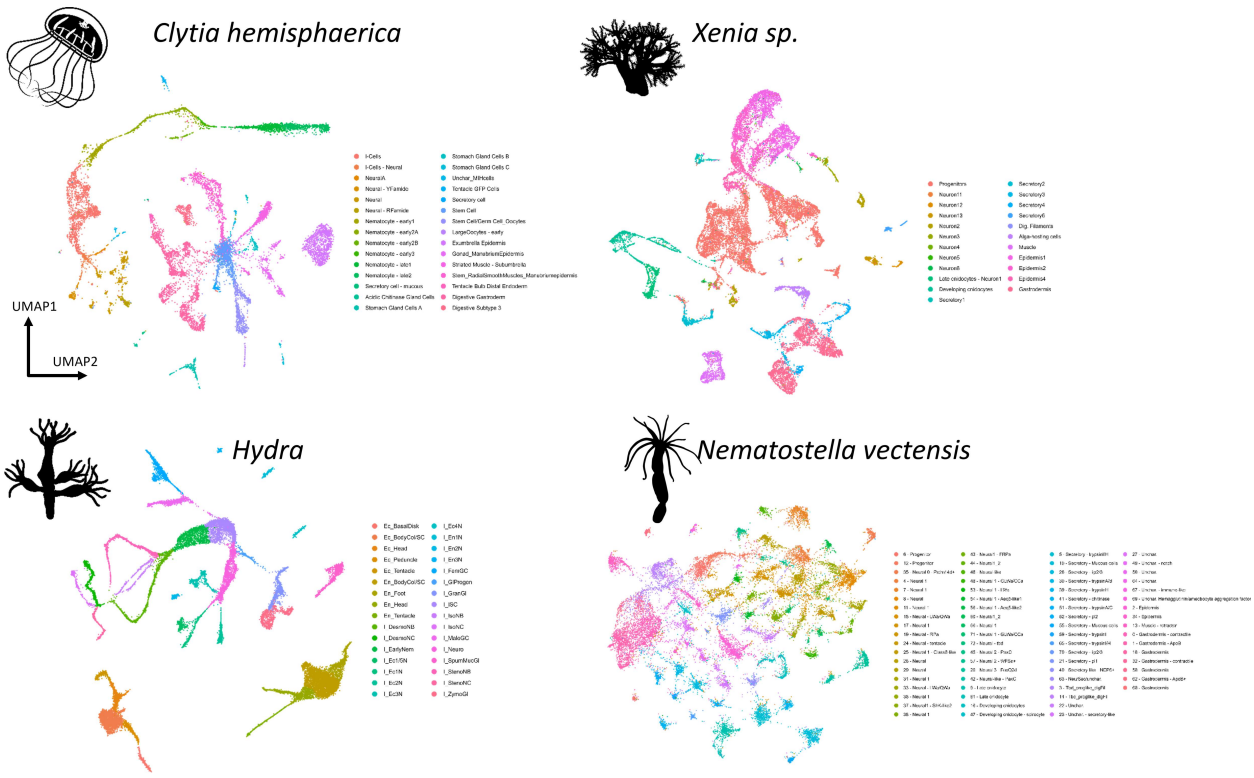

Figure A.2: Cnidarian scRNA-Seq atlases from distantly related species

Cnidarian single cell atlases from hydrozoans *Clytia hemisphaerica* (data from[320]) and *Hydra vulgaris* (data from[138,249]), anthozoan octocoral *Xenia sp* (Anthozoa, Octocorallia) (data from[143]), and anthozoan sea anemone *Nematostella vectensis* (Anthozoa, Hexacorallia), with new 2D UMAP representations (except for *Hydra*, projection from[138] used) with random cluster colors.

cessed by F.P. as described for the body libraries chapter 3), with manually curated clusters annotations.

- *Hydra vulgaris* (class: Hydrozoa, order: Anthoathecata, family: Hydridae) strain AEP scRNA-Seq data from the polyp stage was originally published in[249], then extended and remapped to a reassembled genome in [138], with these updated cell clusters annotations used here.

- *Clytia hemisphaerica* (class: Hydrozoa, order: Leptothecata, family: Campanulariidae) sc RNA-Seq data from the jellyfish life stage is the original data from[320], reclustered and re-annotated for harmonization between the experimental batches.

- *Xenia sp.* (class: Octocorallia, order: Malacalcyonacea, family: Xeniidae), scRNA-Seq data from the polyp stage, originally published in[143], then remapped and annotated in[142], and with additional manual annotations (see below).

Data from these cnidarian whole animal sc RNA-Seq atlases (UMAP 2D representations shown in Fig. A.2), with annotations primarily from the original publications ([8,138,142,143,249,320]) are used for cell-type comparisons. SAMap is used to score transcriptional similarities[9] (from SAM: Self-Assembling Manifold[325]), which does not relies only on one-to-one orthologous genes between species but on BLAST protein sequences similarities across the entire proteome of each species pair compared, thus integrating gene similarities, with an iterative graph-based approach to identify and refine mutual nearest neighbors across species. This provides a continuous [0-1] score (SAMap score hereafter), where 1 is the highest transcriptional similarity and 0 is the lowest (with arbitrary low cutoffs around 0.2 (or lower depending on studies)).

A.2 Eight pan-cnidarian cell class transcriptional programs

Pairwise species atlases comparisons highlighted high SAMap similarity scores for several cell states across these four species, further associated to known cnidarian cell types. Overall, eight cnidarian conserved cell-type-specific transcriptional programs are identifiable: progenitor/stem cells, developing cnidocytes, late cnidocytes, other neurons, mucous secretory cells, digestive secretory cells, gastrodermis, muscle cells (Figs. A.3 and A.4). Regarding the latter, *Hydra* is known to possess myoepithelial cells, not muscle cells, with a contractile unit embedded within epithelial cells, and this cell type shows some transcriptional similarity to both anthozoan muscle cells from *Nematostella* and *Xenia* and to *Clytia*'s muscle cells, as well as to epithelial epidermal cells. *Nematostella* epithelial epidermal cells show very limited similarities across species, while these are detected across other comparisons. Across all species, distinctions within secretory cells delineated a mucous secretory cell class distinct

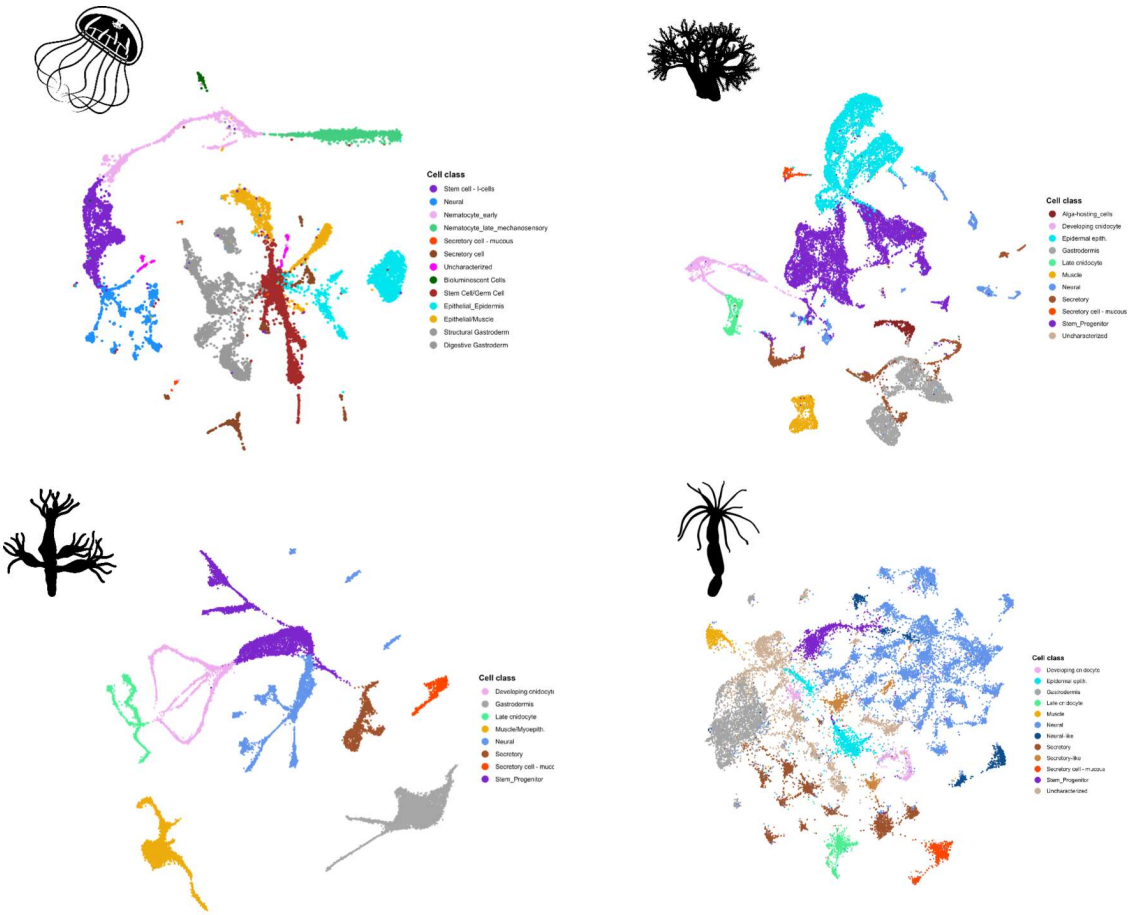

Figure A.3: Cnidarian atlases by broad cell classes

Cnidarian single cell atlases colored by broad cell classes from cell-type similaries. 2D UMAP representation of sc RNA-Seq data from hydrozoans *Clytia hemisphaerica* (top left) (data from[320]) and *Hydra vulgaris* (bottom left) (data from[138,249]), anthozoan octocoral *Xenia sp* (top right) (data from[143] remapped in[142] with proposed annotation for the late cnidocyte state in green), and sea anemone *Nematostella vectensis* (bottom right).

from the other secretory cells (primarily "digestive" secretory cells in *Nematostella*).

We combined original datasets curated annotations and these results to annotate broad cell classes (Figs. A.5 to A.7). Specifically, because the late cnidocyte program was not identified in the *Xenia* dataset[143] (or in its remapped and reannotated version in[142]), we propose here that one of the previously annotated neural state (Neuron 1 annotation in[142], unannotated in the original publication[143]) corresponds to the late cnidocyte cell state, hence the labeling and color here throughout figures (in Fig. A.3). This is supported by several

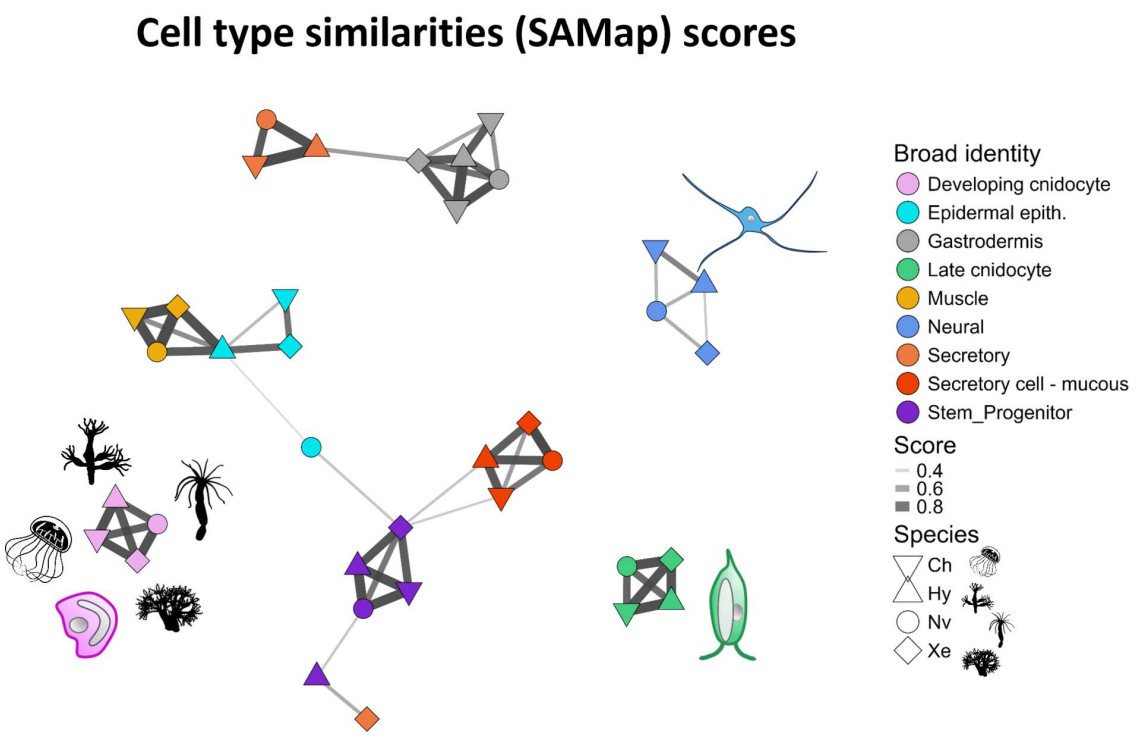

Figure A.4: Eight cnidarian shared cell classes

Network of pairwise cell-class-level SAMap scores across four cnidarian species highlighting broad transcriptional similarities for progenitor/stem cells (in purple), gastrodermis cells (in grey), muscle cells (in yellow), mucous secretory cells (in deep orange), digestive secretory cells (in orange), neural cells (in blue), developing cnidocytes (in pink) and late cnidocytes (in green). *Hydra* is known to have myoepithelial cells, and not separate muscle cells. Ch: *Clytia hemisphaerica*, Hy: *Hydra vulgaris*, Nv: *Nematostella vectensis*, Xe: *Xenia sp.*

elements: (1) the spatial heterogeneity of cnidocytes transcripts reported in the original publication[143], suggestive of developmental maturation with distinct transcriptional profiles, and identifying two cnidocyte maturation states; (2) the new UMAP projection here (see Figs. A.2 and A.3), suggestive of a continuous trajectory from developing cnidocytes to this state (original publication used a t-SNE projection), (3) developing cnidocyte states from other cnidarians match the labeled "cnidocyte" state, with the late cnidocyte state across other cnidarians matching that "Neuron 1" state, proposed to be *Xenia*'s late cnidocyte state.

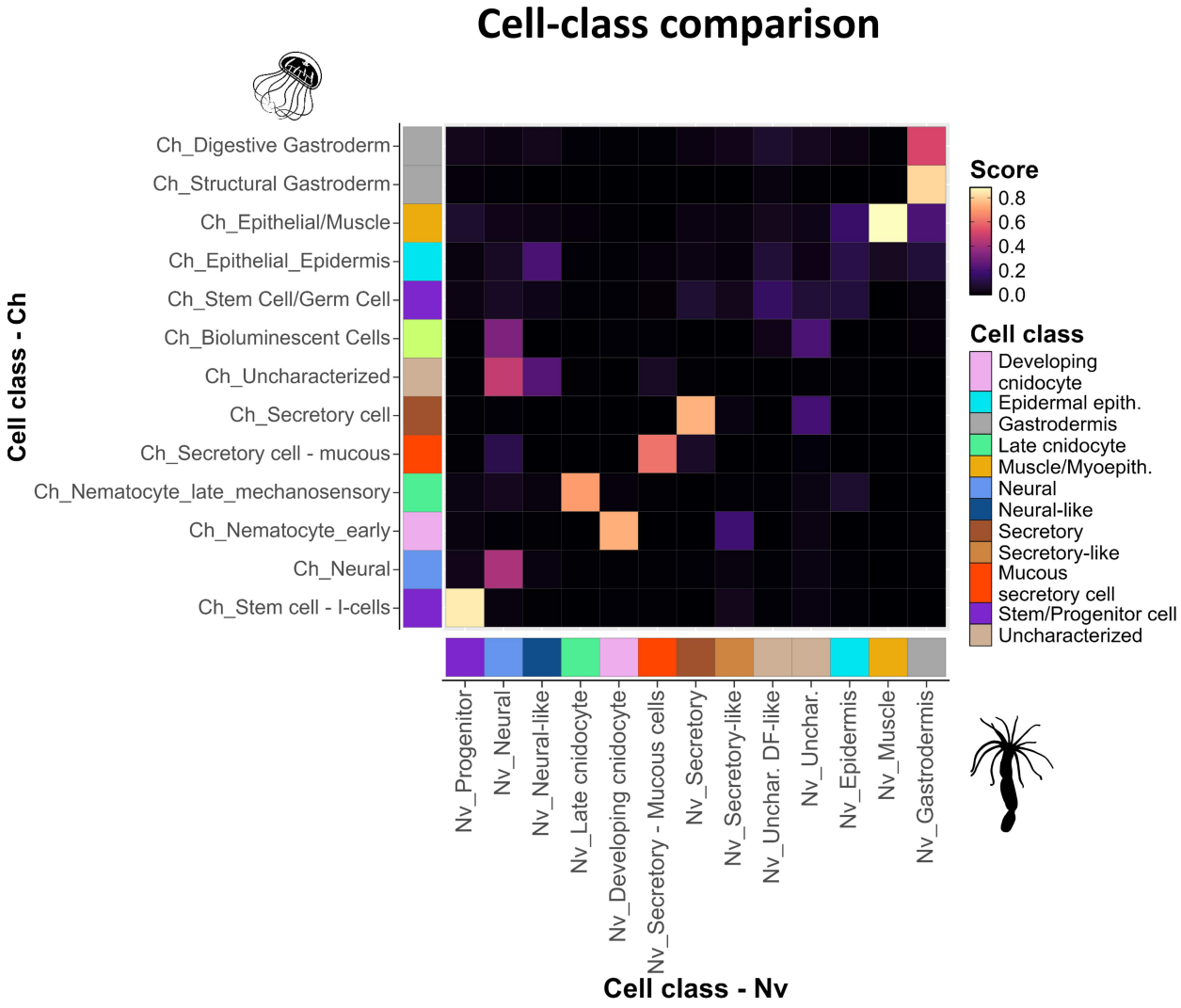

Figure A.5: *Nematostella* and *Clytia hemisphaerica* broad cell class similarities

Broad cell class similarities (SAMap scores) between the anthozoan sea anemone polyp *Nematostella vectensis* and the hydrozoan jellyfish *Clytia hemisphaerica* (data from[320], with curated annotations).

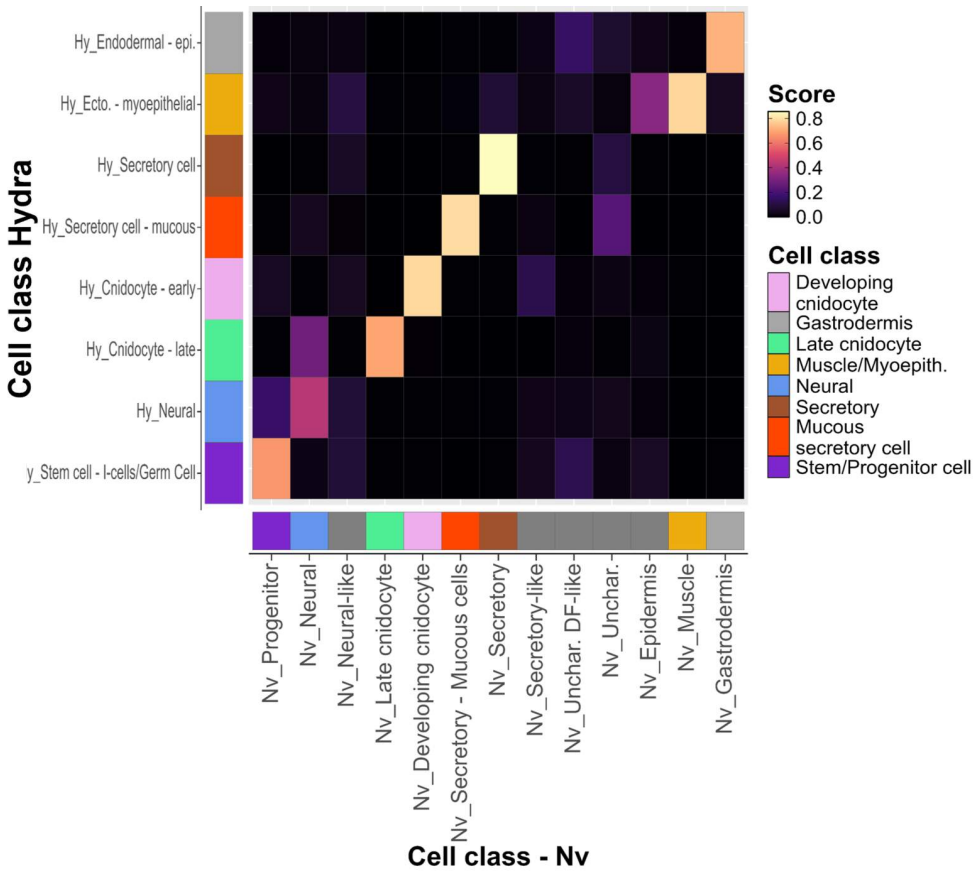

Figure A.6: *Nematostella* and *Hydra* broad cell class similarities

Broad cell class similarities (SAMap scores) between the anthozoan sea anemone polyp *Nematostella vectensis* and the hydrozoan *Hydra*, highlighting a shared late cnidocyte program, among others.

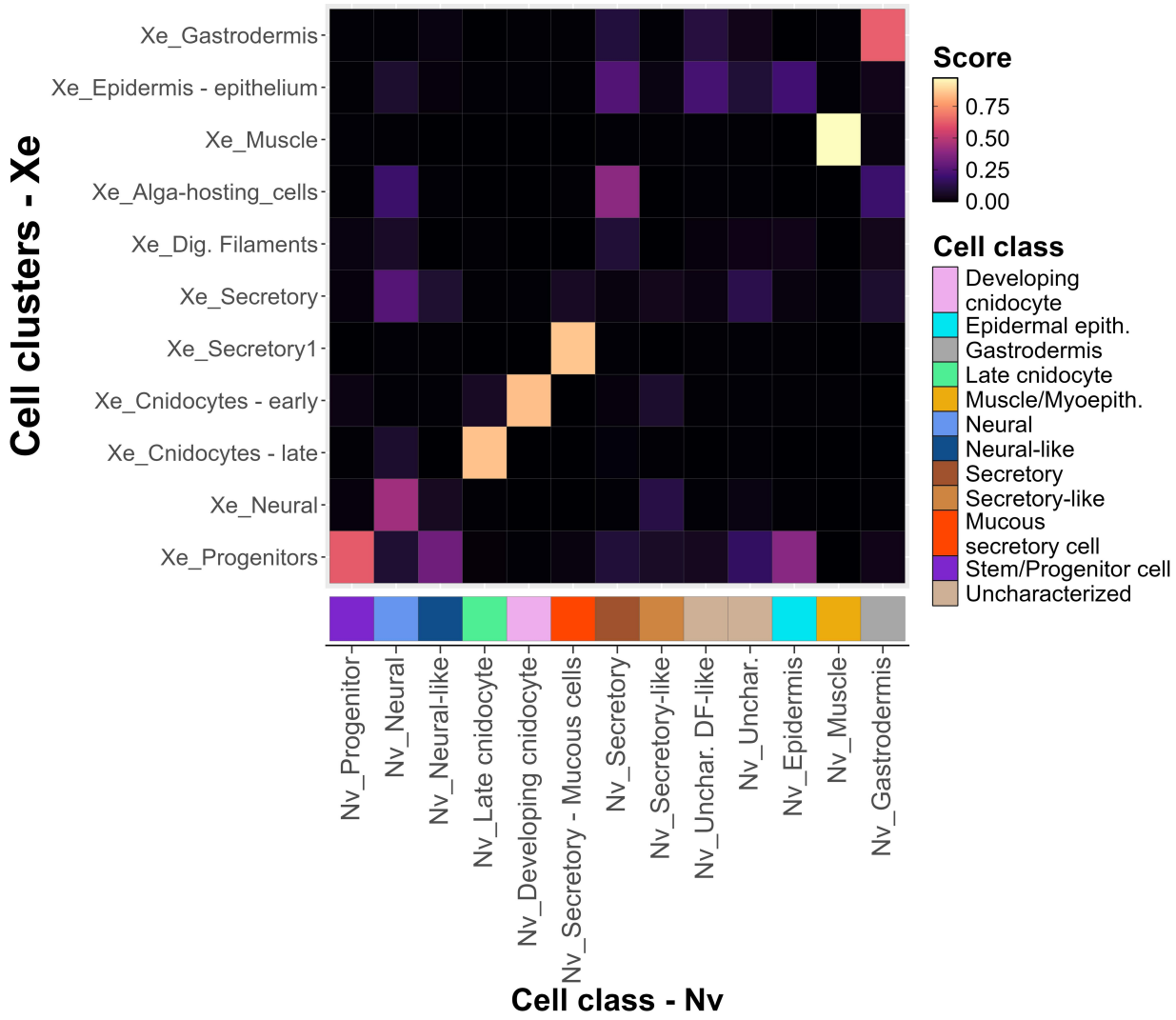

Figure A.7: *Nematostella* and *Xenia sp.* broad cell class similarities

Broad cell class similarities (SAMap scores) between the anthozoan sea anemone polyp *Nematostella vectensis* and the anthozoan octocoral *Xenia sp.* with original data from[143], annotations from[142] with the putative late cnidocyte program annotated here in green.

A.3 Cnidocyte, neuronal and secretory transcriptional similarities

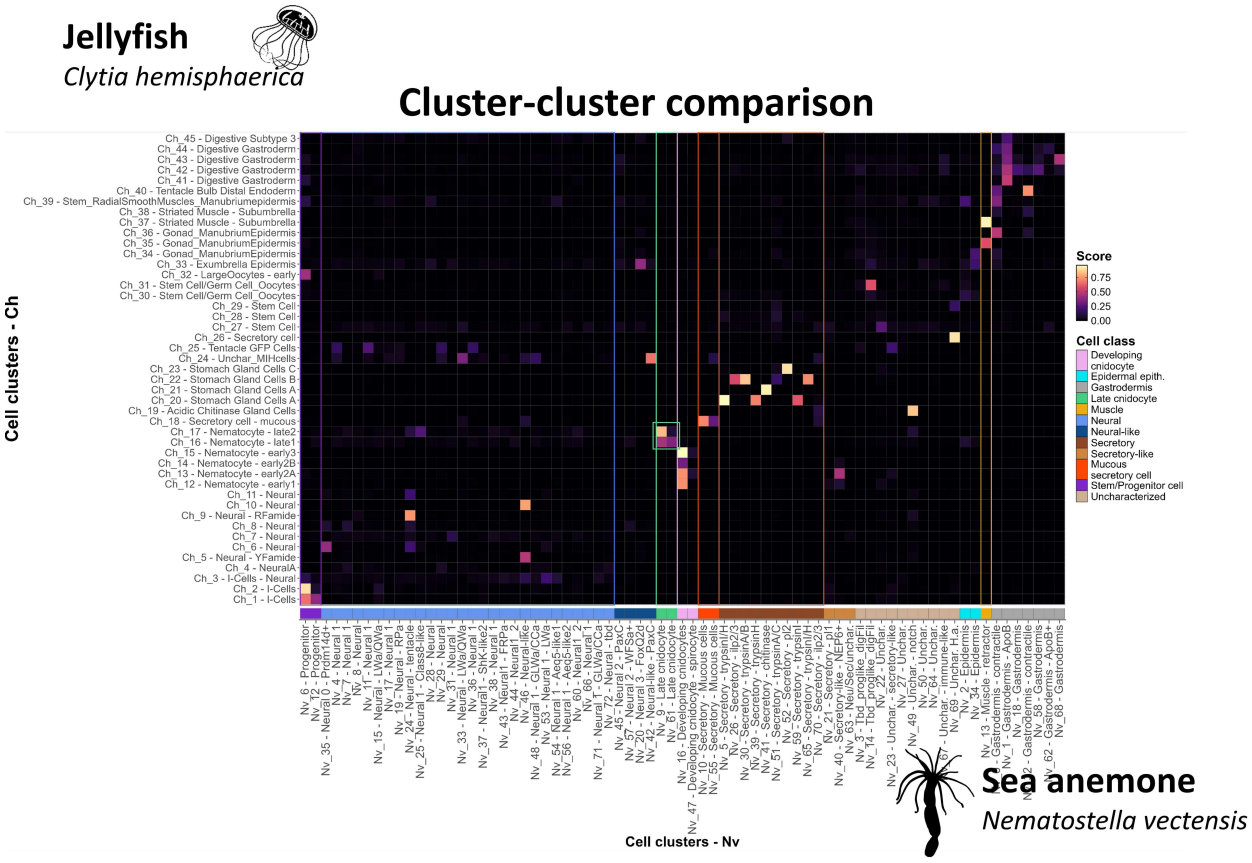

Figure A.8: *Nematostella vectensis* and *Clytia hemisphaerica* cell-type transcriptome similarities

Pairwise cell cluster similarities (SAMap scores) between the anthozoan sea anemone polyp *Nematostella vectensis* and the hydrozoan jellyfish *Clytia hemisphaerica* (data from[320], with curated annotations).

At cell cluster resolution, there are transcriptional similarities between developing cnidocytes states across all comparisons (see Figs. A.8 to A.12). There is little to no transcriptional similarity for one of the two developing cnidocyte state present in *Nematostella*, likely corresponding to spirocytes, an anthozoan-specific cnidocyte (but likely not present in octocorals[326]). This is consistent with spirocyte representing a more restricted innovation within Anthozoa. Indeed, *Xenia* and *Nematostella* are extant anthozoan representatives of the earliest split within this subphylum, separating the class Hexacorallia that *Nematostella*

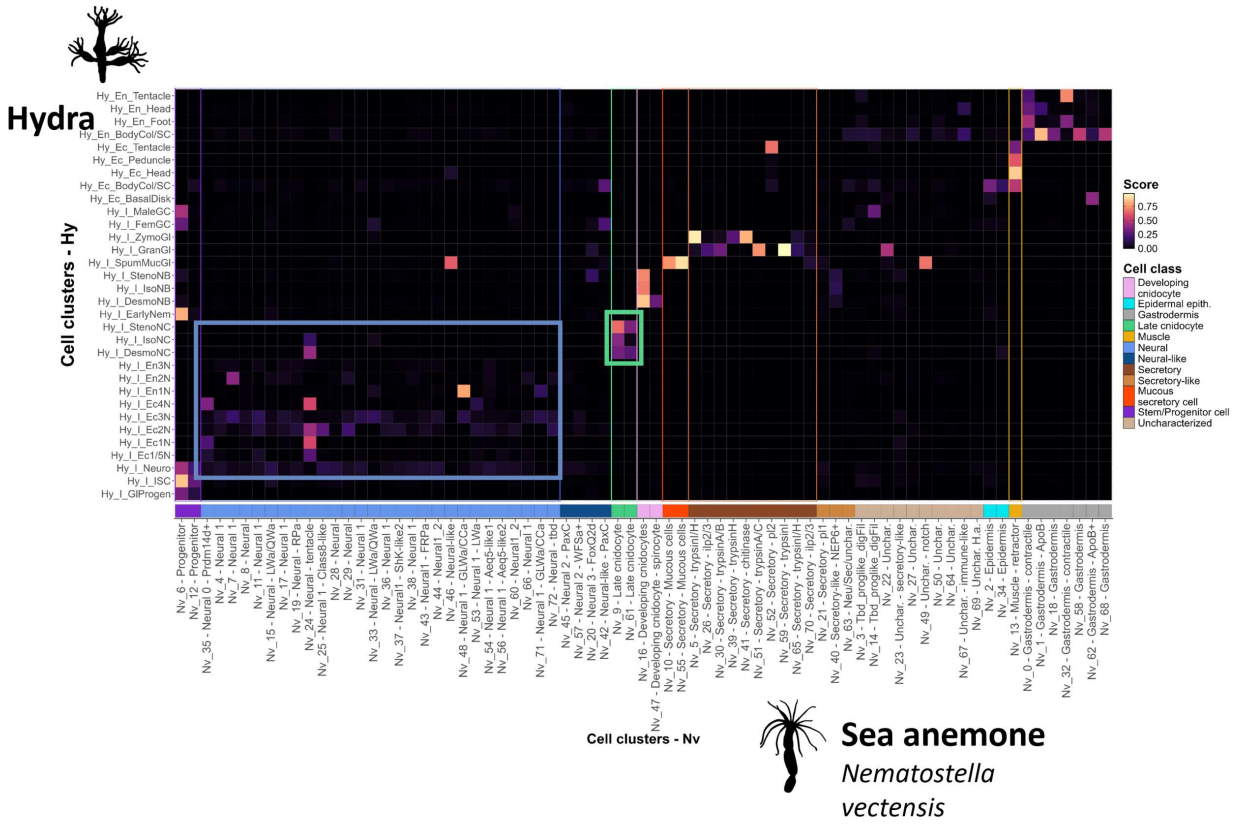

Figure A.9: *Nematostella vectensis* and *Hydra* cell-type transcriptome similarities
Pairwise cell cluster similarities (SAMap scores) between the anthozoan sea anemone polyp *Nematostella vectensis* and the hydrozoan *Hydra* with the updated annotations from[138,249].

belongs to Hexacorallia and Octocorallia that *Xenia* belongs to. This divergence is ancient, at least earlier than 438 Mya (hexacorallian diversification listed in figure Fig. 1.2 from[24] in chapter 1). Beyond broad cell classes, and the developing cnidocyte, late cnidocyte and mucous secretory cell states, there is relatively few subtype-specific cell type similarities. This may reflect the large evolutionary distances between these lineages. Thus, between these two distantly related anthozoans, transcriptional similarities are detected between the two *Nematostella* progenitor clusters, and specific *Xenia* progenitor clusters (note that *Xenia*'s "progenitor" cell states are putatively annotated by[142], with currently no experimental demonstration of their progenitor or stem cell properties).

The distinct developing cnidocyte states likely represent temporal intermediates for *Clytia*

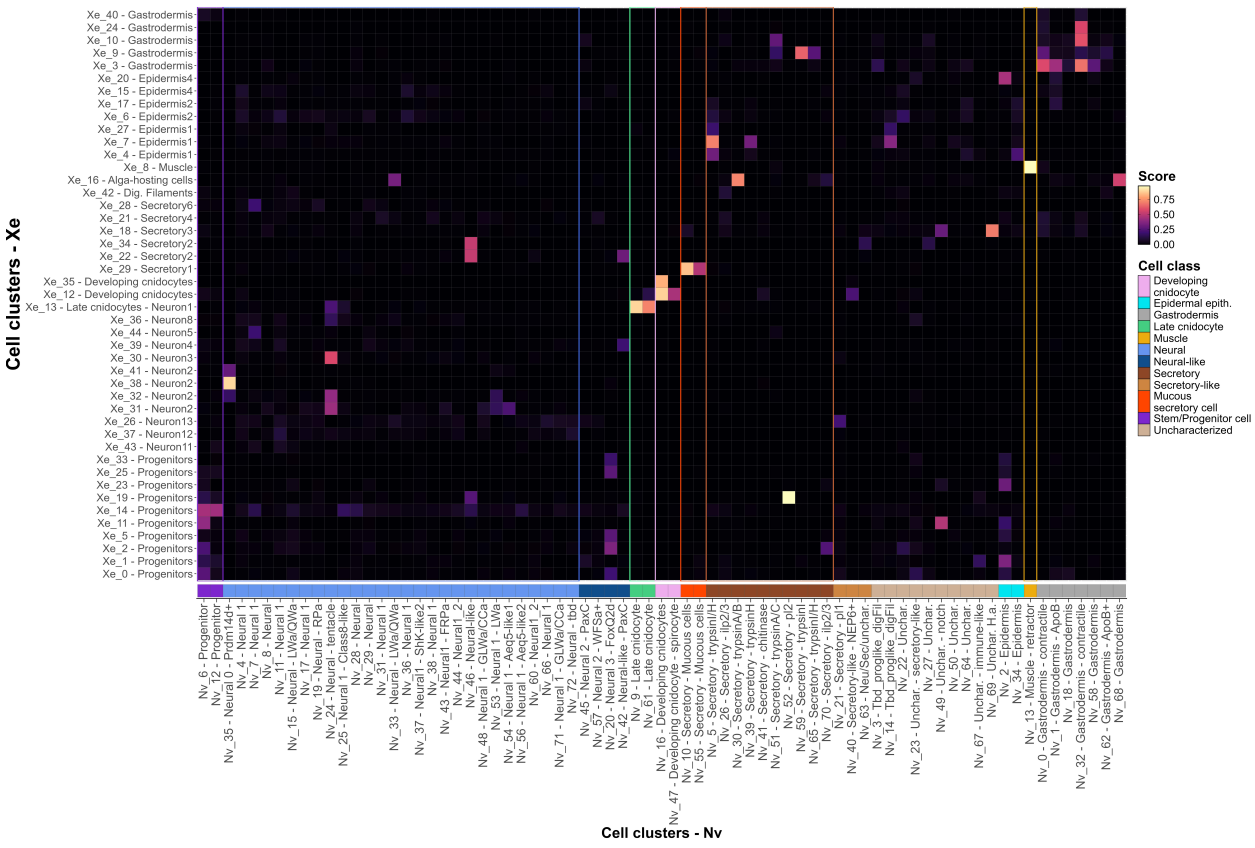

Figure A.10: *Nematostella* and *Xenia sp.* cell-type transcriptome similarities
Pairwise cell cluster similarities (SAMap scores) between the anthozoan sea anemone polyp *Nematostella vectensis* (columns) and the anthozoan octocoral *Xenia sp.* (rows, original data from[143], annotations from[142] with additional annotation for the putative late cnidocyte program).

and *Xenia*, which show broad similarities to developing nematocyte states in *Nematostella* and *Hydra* (across all of its three distinct nematocyte subtypes).

Regarding neuronal populations, there are some one-to-one similarities between *Hydra* and *Clytia*, and much less across *Nematostella* or *Xenia*, with the anthozoans showing an increased molecular diversity of neuronal cell states. Between *Hydra* and *Clytia*, their eight neuronal subtypes show some one-to-one similarities Fig. A.12. Overall, there are little one-to-one similarities between individual neuronal subtypes, except for a tentacular specific neural subtype in *Nematostella*, which shows similarities to several uncharacterized *Xenia* neuron subtypes, one (RFa^{+}) neuronal subtype in the hydrozoan *Clytia*, and to several

ectodermal neurons in *Hydra*.

In *Nematostella*, *Prdm14d*$^{+}$ neurons (identified in[181], and annotated as class 0 Neurons throughout chapter 3) are transcriptionally distinct from other neurons. Throughout all pairwise comparisons here, this neuronal subtype exhibit similarities to one or a few neuronal subtypes across species: to one uncharacterized neural cluster in *Xenia* (a subset of the Neuron 2 type annotated by[142]), as well as similarities between a tentacular neuronal cell subtype in *Nematostella* and several of *Xenia*'s neurons. Some similarities for neuronal subtypes had been noted in[142], more specifically among anthozoans, with different gene models and cell cluster annotations used for *Nematostella* cell types.

Across all pairwise comparisons (see below), the dichotomy between mucous secretory cells (the spumous cells of *Hydra*) and other secretory cells is conserved, often with very high 1-to-1 similarities scores for specific secretory clusters across anthozoans and hydrozoans, thus spanning a large evolutionary distance. This is especially intriguing as the trypsin-domain-containing proteins that are often specific to distinct digestive secretory cells in *Nematostella*[151] have undergone cnidarian-specific and anthozoan-specific diversification events[327]. This suggests that several secretory cell programs (one for mucous cells and at lease two across other cells) might have been present at the base of Cnidaria but remain to be investigated.

Overall there is support for broadly similar gene regulatory programs specifically within developing cnidocyte (one program), late cnidocytes (one program distinct from that of developing cnidocytes and distinct from other neuronal states), other neurons, which much weaker similarities across distantly related species, and at least three subsets of secretory cells: mucous cells on one side, and two subsets within other secretory cells, corresponding to the described zymogen and spumous subtypes in *Hydra*. These similarities are further discussed in relationship with the distinct developmental trajectories of neuronal populations in *Nematostella* in chapter 4 (in section 4.4).

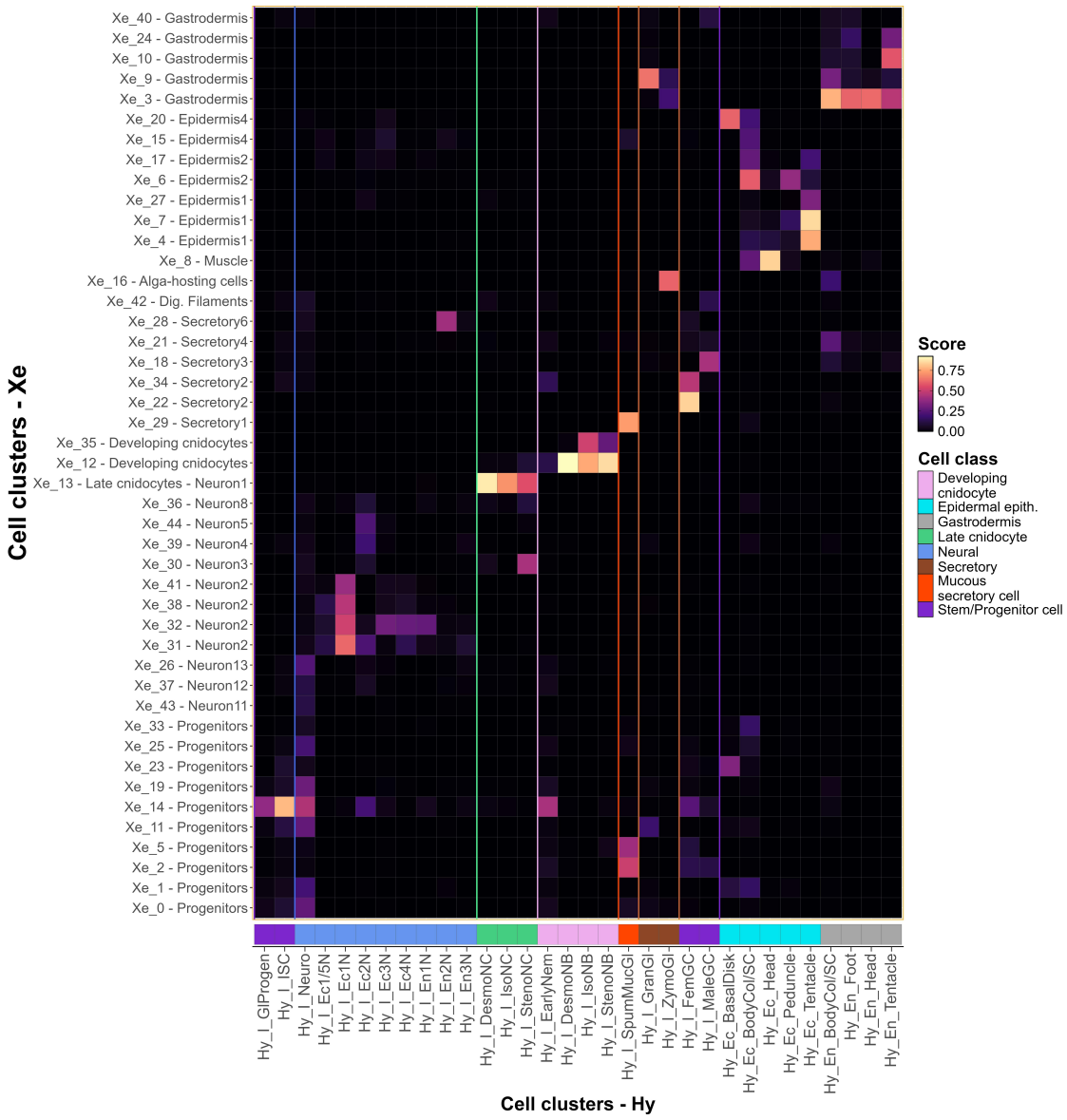

Figure A.11: *Hydra* and *Xenia* cell-type transcriptome similarities

Pairwise cell cluster similarities (SAMap scores) between the hydrozoans *Hydra* and *Xenia*, highlighting similarities across neuronal, late cnidocytes, developing cnidocytes, subsets of secretory cells. Annotations for Xenia cell types are putative and primarily from[142].

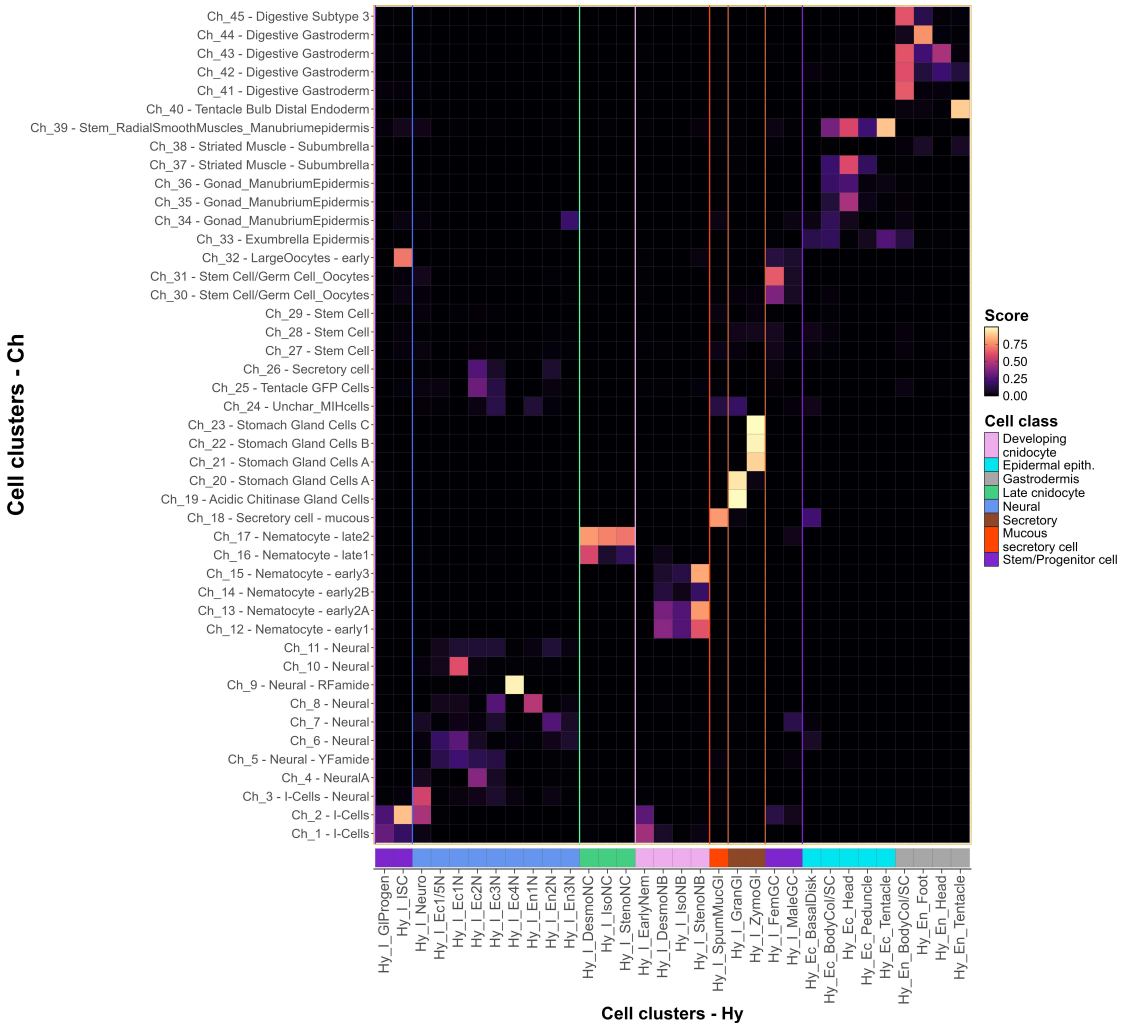

Figure A.12: *Hydra* and *Clytia* cell-type transcriptome similarities

Pairwise cell cluster similarities (SAMap scores) between the hydrozoans *Hydra* and *Clytia*, highlighting similarities across neuronal, late cnidocytes, developing cnidocytes, subsets of secretory cells (with distinct cell types in *Clytia* matching the zymogen, granular, and mucous cell subtypes from *Hydra*). NC: nematocytes (mature cnidocytes), NB: nematoblasts (developing cnidocytes).

A.4 Acknowledgements

We thank Sébastien Bastide for scientific discussions and insights and for providing help with object conversion.

REFERENCES

1. Moody, E. R. R. *et al.* The Nature of the Last Universal Common Ancestor and Its Impact on the Early Earth System. *Nat Ecol Evol* **8,** 1654–1666. doi:10.1038/s41559-024-02461-1 (2024).

2. Knoll, A. H. The Multiple Origins of Complex Multicellularity. *Annual Review of Earth and Planetary Sciences* **39,** 217–239. doi:10.1146/annurev.earth.031208.100209 (2011).

3. Carlisle, E., Yin, Z., Pisani, D. & Donoghue, P. C. J. Ediacaran Origin and Ediacaran-Cambrian Diversification of Metazoa. *Sci Adv* **10,** eadp7161. doi:10.1126/sciadv.adp7161. PMID: 39536100.

4. 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. doi:10.1038/nrn.2015.15 (1 2016).

5. Tanay, A. & Sebé-Pedrós, A. Evolutionary Cell Type Mapping with Single-Cell Genomics. *Trends in Genetics* **37,** 919–932. doi:10.1016/j.tig.2021.04.008 (2021).

6. Arendt, D., Bertucci, P. Y., Achim, K. & Musser, J. M. Evolution of Neuronal Types and Families. *Curr Opin Neurobiol* **56,** 144–152. doi:10.1016/j.conb.2019.01.022. PMID: 30826503 (2019).

7. Pomaville, M. B., Sattler, S. M. & Abitua, P. B. A New Dawn for the Study of Cell Type Evolution. *Development* **151,** dev200884. doi:10.1242/dev.200884. PMID: 38722217 (2024).

8. Sebé-Pedrós, A. *et al.* Cnidarian Cell Type Diversity and Regulation Revealed by Whole-Organism Single-Cell RNA-Seq. *Cell* **173,** 1520–1534.e20. doi:10.1016/j.cell.2018.05.019. PMID: 29856957 (2018).

9. Tarashansky, A. J. *et al.* Mapping Single-Cell Atlases throughout Metazoa Unravels Cell Type Evolution. *eLife* **10,** e66747. doi:10.7554/eLife.66747 (2021).

10. Sebé-Pedrós, A., Degnan, B. M. & Ruiz-Trillo, I. The Origin of Metazoa: A Unicellular Perspective. *Nat Rev Genet* **18,** 498–512. doi:10.1038/nrg.2017.21 (2017).

11. López-Escardó, D. *et al.* Reconstruction of Protein Domain Evolution Using Single-Cell Amplified Genomes of Uncultured Choanoflagellates Sheds Light on the Origin of Animals. *Philos Trans R Soc Lond B Biol Sci* **374,** 20190088. doi:10.1098/rstb.2019.0088. PMID: 31587642 (2019).

12. De Mendoza, A., Suga, H. & Ruiz-Trillo, I. Evolution of the MAGUK Protein Gene Family in Premetazoan Lineages. *BMC Evol Biol* **10,** 93. doi:10.1186/1471-2148-10-93 (2010).

13. Degnan, B. M., Vervoort, M., Larroux, C. & Richards, G. S. Early Evolution of Metazoan Transcription Factors. *Current Opinion in Genetics & Development. Genomes and Evolution* **19,** 591–599. doi:10.1016/j.gde.2009.09.008 (2009).

14. De Mendoza, A. *et al.* Transcription Factor Evolution in Eukaryotes and the Assembly of the Regulatory Toolkit in Multicellular Lineages. *Proc Natl Acad Sci U S A* **110,** E4858–E4866. doi:10.1073/pnas.1311818110. PMID: 24277850 (2013).

15. De Mendoza, A. & Sebé-Pedrós, A. Origin and Evolution of Eukaryotic Transcription Factors. *Current Opinion in Genetics & Development. Evolutionary Genetics* **58–59,** 25–32. doi:10.1016/j.gde.2019.07.010 (2019).

16. Burkhardt, P. *et al.* Evolutionary Insights into Premetazoan Functions of the Neuronal Protein Homer. *Mol Biol Evol* **31,** 2342–2355. doi:10.1093/molbev/msu178. PMID: 24899667 (2014).

17. Liebeskind, B. J., Hofmann, H. A., Hillis, D. M. & Zakon, H. H. Evolution of Animal Neural Systems. *Annual Review of Ecology, Evolution, and Systematics* **48,** 377–398. doi:10.1146/annurev-ecolsys-110316-023048 (Volume 48, 2017 2017).

18. Liebeskind, B. J., Hillis, D. M. & Zakon, H. H. Evolution of Sodium Channels Predates the Origin of Nervous Systems in Animals. *Proc Natl Acad Sci U S A* **108,** 9154–9159. doi:10.1073/pnas.1106363108. PMID: 21576472 (2011).

19. Gur Barzilai, M. *et al.* Convergent Evolution of Sodium Ion Selectivity in Metazoan Neuronal Signaling. *Cell Rep* **2,** 242–248. doi:10.1016/j.celrep.2012.06.016. PMID: 22854023 (2012).

20. Liebeskind, B. J., Hillis, D. M. & Zakon, H. H. Convergence of Ion Channel Genome Content in Early Animal Evolution. *Proceedings of the National Academy of Sciences* **112,** E846–E851. doi:10.1073/pnas.1501195112 (2015).

21. Colgren, J. & Burkhardt, P. The Premetazoan Ancestry of the Synaptic Toolkit and Appearance of First Neurons. *Essays Biochem* **66,** 781–795. doi:10.1042/EBC20220042. PMID: 36205407 (2022).

22. Moroz, L. L. & Kohn, A. B. Independent Origins of Neurons and Synapses: Insights from Ctenophores. *Philos Trans R Soc Lond B Biol Sci* **371,** 20150041. doi:10.1098/rstb.2015.0041. PMID: 26598724 (2016).

23. Moroz, L. L., Romanova, D. Y. & Kohn, A. B. Neural versus Alternative Integrative Systems: Molecular Insights into Origins of Neurotransmitters. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376,** 20190762. doi:10.1098/rstb.2019.0762 (2021).

24. Khalturin, K. *et al.* Medusozoan Genomes Inform the Evolution of the Jellyfish Body Plan. *Nat Ecol Evol* **3,** 811–822. doi:10.1038/s41559-019-0853-y (2019).

25. Ryan, T. J. & Grant, S. G. N. The Origin and Evolution of Synapses. *Nat Rev Neurosci* **10,** 701–712. doi:10.1038/nrn2717. PMID: 19738623 (2009).

26. Colgren, J. J. & Burkhardt, P. The Evolutionary Origins of Synaptic Proteins and Their Changing Roles in Different Organisms across Evolution. *Nat. Rev. Neurosci.,* 1–16. doi:10.1038/s41583-025-00983-6 (2025).

27. Burkhardt, P. & Jékely, G. Evolution of Synapses and Neurotransmitter Systems: The Divide-and-Conquer Model for Early Neural Cell-Type Evolution. *Current Opinion in Neurobiology. Evolution of Brains and Computation* **71,** 127–138. doi:10.1016/j.conb.2021.11.002 (2021).

28. Burkhardt, P. *et al.* Syncytial Nerve Net in a Ctenophore Adds Insights on the Evolution of Nervous Systems. *Science* **380,** 293–297. doi:10.1126/science.ade5645 (2023).

29. Sachkova, M., Modepalli, V. & Kittelmann, M. The Deep Evolutionary Roots of the Nervous System. *Annual Review of Neuroscience* **48,** 311–329. doi:10.1146/annurev-neuro-112723-040945 (Volume 48, 2025 2025).

30. Brette, R. Integrative Neuroscience of Paramecium, a "Swimming Neuron". *eNeuro* **8,** ENEURO.0018–21.2021. doi:10.1523/ENEURO.0018-21.2021. PMID: 33952615 (2021).

31. Senatore, A., Raiss, H. & Le, P. Physiology and Evolution of Voltage-Gated Calcium Channels in Early Diverging Animal Phyla: Cnidaria, Placozoa, Porifera and Ctenophora. *Front. Physiol.* **7.** doi:10.3389/fphys.2016.00481 (2016).

32. Moran, Y., Barzilai, M. G., Liebeskind, B. J. & Zakon, H. H. Evolution of Voltage-Gated Ion Channels at the Emergence of Metazoa. *J Exp Biol* **218,** 515–525. doi:10.1242/jeb.110270 (2015).

33. Ramos-Vicente, D. *et al.* Metazoan Evolution of Glutamate Receptors Reveals Unreported Phylogenetic Groups and Divergent Lineage-Specific Events. *eLife* **7,** e35774. doi:10.7554/eLife.35774 (2018).

34. Ramos-Vicente, D., Grant, S. G. & Bayés, À. Metazoan Evolution and Diversity of Glutamate Receptors and Their Auxiliary Subunits. *Neuropharmacology* **195,** 108640. doi:10.1016/j.neuropharm.2021.108640 (2021).

35. Pascale, A., Amadio, M. & Quattrone, A. Defining a Neuron: Neuronal ELAV Proteins. *Cell Mol Life Sci* **65,** 128–140. doi:10.1007/s00018-007-7017-y. PMID: 17928954 (2007).

36. Marlow, H. Q., Srivastava, M., Matus, D. Q., Rokhsar, D. & Martindale, M. Q. Anatomy and Development of the Nervous System of Nematostella Vectensis, an Anthozoan Cnidarian. *Developmental Neurobiology* **69,** 235–254. doi:10.1002/dneu.20698 (2009).

37. Kerner, P., Degnan, S. M., Marchand, L., Degnan, B. M. & Vervoort, M. Evolution of RNA-Binding Proteins in Animals: Insights from Genome-Wide Analysis in the Sponge Amphimedon Queenslandica. *Mol Biol Evol* **28,** 2289–2303. doi:10.1093/molbev/msr046 (2011).

38. Nakanishi, N., Renfer, E., Technau, U. & Rentzsch, F. Nervous Systems of the Sea Anemone Nematostella Vectensis Are Generated by Ectoderm and Endoderm and Shaped by Distinct Mechanisms. *Development* **139,** 347–357. doi:10.1242/dev.071902. PMID: 22159579 (2012).

39. Simionato, E. *et al.* Origin and Diversification of the Basic Helix-Loop-Helix Gene Family in Metazoans: Insights from Comparative Genomics. *BMC Evol Biol* **7,** 33. doi:10.1186/1471-2148-7-33. PMID: 17335570 (2007).

40. Baker, N. E. & Brown, N. L. All in the Family: Proneural bHLH Genes and Neuronal Diversity. *Development* **145.** doi:10.1242/dev.159426. PMID: 29720483 (2018).

41. Hartenstein, V., Takashima, S., Hartenstein, P., Asanad, S. & Asanad, K. bHLH Proneural Genes as Cell Fate Determinants of Entero-Endocrine Cells, an Evolutionarily Conserved Lineage Sharing a Common Root with Sensory Neurons. *Dev Biol* **431,** 36–47. doi:10.1016/j.ydbio.2017.07.013. PMID: 28751238 (2017).

42. Gao, Y. *et al.* The Emergence of Sox and POU Transcription Factors Predates the Origins of Animal Stem Cells. *Nat Commun* **15,** 9868. doi:10.1038/s41467-024-54152-x (2024).

43. Magie, C. R., Pang, K. & Martindale, M. Q. Genomic Inventory and Expression of Sox and Fox Genes in the Cnidarian Nematostella Vectensis. *Dev Genes Evol* **215,** 618–630. doi:10.1007/s00427-005-0022-y (2005).

44. Hartenstein, V. & Stollewerk, A. The Evolution of Early Neurogenesis. *Dev. Cell* **32,** 390–407. doi:10.1016/j.devcel.2015.02.004. PMID: 25710527 (2015).

45. Wegner, M. SOX after SOX: SOXession Regulates Neurogenesis. *Genes Dev* **25,** 2423–2428. doi:10.1101/gad.181487.111. PMID: 22156204 (2011).

46. Larroux, C. *et al.* Genesis and Expansion of Metazoan Transcription Factor Gene Classes. *Molecular Biology and Evolution* **25,** 980–996. doi:10.1093/molbev/msn047 (2008).

47. Rentzsch, F., Layden, M. & Manuel, M. The Cellular and Molecular Basis of Cnidarian Neurogenesis. *Wiley Interdiscip Rev Dev Biol* **6.** doi:10.1002/wdev.257. PMID: 27882698 (2017).

48. Galliot, B. *et al.* Origins of Neurogenesis, a Cnidarian View. *Developmental Biology. Special Section: Evolution of Developmental Regulatory Systems* **332,** 2–24. doi:10.1016/j.ydbio.2009.05.563 (2009).

49. Galliot, B. & Quiquand, M. A Two-Step Process in the Emergence of Neurogenesis. *European Journal of Neuroscience* **34,** 847–862. doi:10.1111/j.1460-9568.2011.07829.x (2011).

50. Richards, G. S. & Rentzsch, F. Transgenic Analysis of a SoxB Gene Reveals Neural Progenitor Cells in the Cnidarian Nematostella Vectensis. *Development* **141,** 4681–4689. doi:10.1242/dev.112029. PMID: 25395455 (2014).

51. Richards, G. S. & Rentzsch, F. Regulation of Nematostella Neural Progenitors by SoxB, Notch and bHLH Genes. *Development* **142,** 3332–3342. doi:10.1242/dev.123745. PMID: 26443634 (2015).

52. Layden, M. J., Boekhout, M. & Martindale, M. Q. Nematostella Vectensis Achaete-Scute Homolog NvashA Regulates Embryonic Ectodermal Neurogenesis and Represents an Ancient Component of the Metazoan Neural Specification Pathway. *Development* **139,** 1013–1022. doi:10.1242/dev.073221. PMID: 22318631 (2012).

53. Steger, J. *et al.* Single-Cell Transcriptomics Identifies Conserved Regulators of Neuroglandular Lineages. *Cell Reports* **40,** 111370. doi:10.1016/j.celrep.2022.111370 (2022).

54. Flici, H. *et al.* An Evolutionarily Conserved SoxB-Hdac2 Crosstalk Regulates Neurogenesis in a Cnidarian. *Cell Rep* **18,** 1395–1409. doi:10.1016/j.celrep.2017.01.019. PMID: 28178518 (2017).

55. Chrysostomou, E. *et al.* A Cellular and Molecular Analysis of SoxB-driven Neurogenesis in a Cnidarian. *eLife* **11,** e78793. doi:10.7554/eLife.78793. PMID: 35608899.

56. Chouly, M. & Bally-Cuif, L. Generating neurons in the embryonic and adult brain: compared principles and mechanisms. *Comptes Rendus. Biologies* **347,** 199–221. doi:10.5802/crbiol.167 (2024).

57. Gold, D. A. & Jacobs, D. K. Stem Cell Dynamics in Cnidaria: Are There Unifying Principles? *Dev Genes Evol* **223,** 53–66. doi:10.1007/s00427-012-0429-1 (2013).

58. Scimone, M. L., Kravarik, K. M., Lapan, S. W. & Reddien, P. W. Neoblast Specialization in Regeneration of the Planarian Schmidtea Mediterranea. *Stem Cell Reports* **3,** 339–352. doi:10.1016/j.stemcr.2014.06.001. PMID: 25254346 (2014).

59. Molinaro, A. M. & Pearson, B. J. In Silico Lineage Tracing through Single Cell Transcriptomics Identifies a Neural Stem Cell Population in Planarians. *Genome Biol* **17,** 87. doi:10.1186/s13059-016-0937-9. PMID: 27150006 (2016).

60. Westfall, J. A., Kinnamon, J. C. & Sims, D. E. Neuro-Epitheliomuscular Cell and Neuro-Neuronal Gap Junctions inHydra. *J Neurocytol* **9,** 725–732. doi:10.1007/BF01205015 (1980).

61. Pierobon, P. Coordinated Modulation of Cellular Signaling through Ligand-Gated Ion Channels in Hydra Vulgaris (Cnidaria, Hydrozoa). *Int. J. Dev. Biol.* **56,** 551–565. doi:10.1387/ijdb.113464pp. PMID: 22689363 (6-7-8 2012).

62. Kass-Simon, G. & Pierobon, P. Cnidarian Chemical Neurotransmission, an Updated Overview. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **146,** 9–25. doi:10.1016/j.cbpa.2006.09.008 (2007).

63. Moroz, L. L., Nikitin, M. A., Poličar, P. G., Kohn, A. B. & Romanova, D. Y. Evolution of Glutamatergic Signaling and Synapses. *Neuropharmacology,* 108740. doi:10.1016/j.neuropharm.2021.108740 (2021).

64. Wessler, I., Kilbinger, H., Bittinger, F., Unger, R. & Kirkpatrick, C. J. The Non-Neuronal Cholinergic System in Humans: Expression, Function and Pathophysiology. *Life Sciences. Proceedings of the 1st International Symposium on Non-neuronal Acetylcholine. Non-neuronal Acetylcholine: Signaling Pathways and Major Biological Effects in Humans* **72,** 2055–2061. doi:10.1016/S0024-3205(03)00083-3 (2003).

65. Horiuchi, Y. *et al.* Evolutional Study on Acetylcholine Expression. *Life Sciences* **72,** 1745–1756. doi:10.1016/S0024-3205(02)02478-5 (2003).

66. Chapman, J. A. *et al.* The Dynamic Genome of Hydra. *Nature* **464,** 592–596. doi:10.1038/nature08830. PMID: 20228792 (2010).

67. Goulty, M., Botton-Amiot, G., Rosato, E., Sprecher, S. G. & Feuda, R. The Monoaminergic System Is a Bilaterian Innovation. *Nat Commun* **14,** 3284. doi:10.1038/s41467-023-39030-2. PMID: 37280201 (2023).

68. Faltine-Gonzalez, D. Z. & Layden, M. J. Characterization of nAChRs in Nematostella Vectensis Supports Neuronal and Non-Neuronal Roles in the Cnidarian–Bilaterian Common Ancestor. *EvoDevo* **10,** 27. doi:10.1186/s13227-019-0136-3. PMID: 31700598 (2019).

69. Anctil, M. Chemical Transmission in the Sea Anemone *Nematostella Vectensis*: A Genomic Perspective. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* **4,** 268–289. doi:10.1016/j.cbd.2009.07.001 (2009).

70. Cole, A. G. *et al.* Muscle Cell Type Diversification Facilitated by Extensive Gene Duplications. *bioRxiv,* 2020.07.19.210658. doi:10.1101/2020.07.19.210658 (2020).

71. Van Giesen, L., Kilian, P. B., Allard, C. A. & Bellono, N. W. Molecular Basis of Chemotactile Sensation in Octopus. *Cell* **183,** 594–604.e14. doi:10.1016/j.cell.2020.09.008. PMID: 33125889 (2020).

72. Allard, C. A. H. *et al.* Structural Basis of Sensory Receptor Evolution in Octopus. *Nature* **616,** 373–377. doi:10.1038/s41586-023-05822-1 (2023).

73. Weir, K., Dupre, C., van Giesen, L., Lee, A. S.-Y. & Bellono, N. W. A Molecular Filter for the Cnidarian Stinging Response. *eLife* **9,** e57578. doi:10.7554/eLife.57578 (2020).

74. Denker, E., Chatonnet, A. & Rabet, N. Acetylcholinesterase Activity in *Clytia Hemisphaerica* (Cnidaria). *Chemico-Biological Interactions. Proceedings of the IX International Meeting on Cholinesterases* **175,** 125–128. doi:10.1016/j.cbi.2008.03.004 (2008).

75. Oren, M., Brikner, I., Appelbaum, L. & Levy, O. Fast Neurotransmission Related Genes Are Expressed in Non Nervous Endoderm in the Sea Anemone Nematostella Vectensis. *PLoS One* **9,** e93832. doi:10.1371/journal.pone.0093832. PMID: 24705400 (2014).

76. Hajj-Ali, I. & Anctil, M. Characterization of a Serotonin Receptor in the Cnidarian Renilla Koellikeri: A Radiobinding Analysis. *Neurochem Int* **31,** 83–93. doi:10.1016/s0197-0186(96)00134-9. PMID: 9185168 (1997).

77. Carlberg, M. & Anctil, M. Biogenic Amines in Coelenterates. *Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology* **106,** 1–9. doi:10.1016/0742-8413(93)90250-O (1993).

78. Peres, R. *et al.* Developmental and Light-Entrained Expression of Melatonin and Its Relationship to the Circadian Clock in the Sea Anemone Nematostella Vectensis. *EvoDevo* **5,** 26. doi:10.1186/2041-9139-5-26. PMID: 25243057 (2014).

79. Roopin, M. & Levy, O. Melatonin Distribution Reveals Clues to Its Biological Significance in Basal Metazoans. *PLOS ONE* **7,** e52266. doi:10.1371/journal.pone.0052266 (2012).

80. Roopin, M. & Levy, O. Temporal and Histological Evaluation of Melatonin Patterns in a 'basal' Metazoan. *J Pineal Res* **53,** 259–269. doi:10.1111/j.1600-079X.2012.00994.x. PMID: 22506978 (2012).

81. Yañez-Guerra, L. A. *et al. Functional and Phylogenetic Analysis of Placozoan GPCRs Reveal the Prebilaterian Origin of Monoaminergic Signalling* https://www.biorxiv.org/content/10.1101/2025.04.18.649542v1 (2026). Pre-published.

82. Kanaya, H. J. *et al.* A Sleep-like State in Hydra Unravels Conserved Sleep Mechanisms during the Evolutionary Development of the Central Nervous System. *Sci Adv* **6,** eabb9415. doi:10.1126/sciadv.abb9415. PMID: 33028524 (2020).

83. Aguillon, R. *et al.* DNA Damage Modulates Sleep Drive in Basal Cnidarians with Divergent Chronotypes. *Nat Commun* **17,** 3. doi:10.1038/s41467-025-67400-5. PMID: 41495058 (2026).

84. Pierobon, P. An Interesting Molecule: -Aminobutyric Acid. What Can We Learn from Hydra Polyps? *Brain Sci* **11,** 437. doi:10.3390/brainsci11040437. PMID: 33805330 (2021).

85. Jékely, G. *et al.* The Long and the Short of It – a Perspective on Peptidergic Regulation of Circuits and Behaviour. *Journal of Experimental Biology* **221.** doi:10.1242/jeb.166710. PMID: 29439060 (2018).

86. Jékely, G. The Chemical Brain Hypothesis for the Origin of Nervous Systems. *Philos Trans R Soc Lond B Biol Sci* **376,** 20190761. doi:10.1098/rstb.2019.0761. PMID: 33550946.

87. Luxmi, R., Kumar, D., Mains, R. E., King, S. M. & Eipper, B. A. Cilia-Based Peptidergic Signaling. *PLoS Biol* **17,** e3000566. doi:10.1371/journal.pbio.3000566. PMID: 31809498 (2019).

88. Kumar, D. *et al.* A Bioactive Peptide Amidating Enzyme Is Required for Ciliogenesis. *Elife* **6,** e25728. doi:10.7554/eLife.25728. PMID: 28513435 (2017).

89. Varoqueaux, F. *et al.* High Cell Diversity and Complex Peptidergic Signaling Underlie Placozoan Behavior. *Current Biology* **28,** 3495–3501.e2. doi:10.1016/j.cub.2018.08.067 (2018).

90. Najle, S. R. *et al.* Stepwise Emergence of the Neuronal Gene Expression Program in Early Animal Evolution. *Cell* **186,** 4676–4693.e29. doi:10.1016/j.cell.2023.08.027. PMID: 37729907 (2023).

91. Sachkova, M. Y. *et al.* The Unique Neuronal Structure and Neuropeptide Repertoire in the Ctenophore Mnemiopsis Leidyi Shed Light on the Evolution of Animal Nervous Systems. *bioRxiv,* 2021.03.31.437758. doi:10.1101/2021.03.31.437758 (2021).

92. Hayakawa, E. *et al.* Mass Spectrometry of Short Peptides Reveals Common Features of Metazoan Peptidergic Neurons. *Nat Ecol Evol* **6,** 1438–1448. doi:10.1038/s41559-022-01835-7 (2022).

93. Yañez-Guerra, L. A., Thiel, D. & Jékely, G. Premetazoan Origin of Neuropeptide Signaling. *Mol Biol Evol* **39,** msac051. doi:10.1093/molbev/msac051 (2022).

94. Mirabeau, O. & Joly, J.-S. Molecular Evolution of Peptidergic Signaling Systems in Bilaterians. *PNAS* **110,** E2028–E2037. doi:10.1073/pnas.1219956110. PMID: 23671109 (2013).

95. Jékely, G. Global View of the Evolution and Diversity of Metazoan Neuropeptide Signaling. *Proc Natl Acad Sci U S A* **110,** 8702–8707. doi:10.1073/pnas.1221833110. PMID: 23637342 (2013).

96. Elphick, M. R., Mirabeau, O. & Larhammar, D. Evolution of Neuropeptide Signalling Systems. *J Exp Biol* **221,** jeb151092. doi:10.1242/jeb.151092. PMID: 29440283 (Pt 3 2018).

97. Thiel, D., Franz-Wachtel, M., Aguilera, F. & Hejnol, A. Xenacoelomorph Neuropeptidomes Reveal a Major Expansion of Neuropeptide Systems during Early Bilaterian Evolution. *Molecular Biology and Evolution* **35,** 2528–2543. doi:10.1093/molbev/msy160 (2018).

98. Thiel, D. *et al.* Large-Scale Deorphanization of Nematostella Vectensis Neuropeptide GPCRs Supports the Independent Expansion of Bilaterian and Cnidarian Peptidergic Systems. *eLife* **12.** doi:10.7554/eLife.90674.1 (2023).

99. Takahashi, T. *et al.* Systematic Isolation of Peptide Signal Molecules Regulating Development in Hydra: LWamide and PW Families. *Proceedings of the National Academy of Sciences* **94,** 1241–1246. doi:10.1073/pnas.94.4.1241 (1997).

100. Hayakawa, E. *et al.* A Combined Strategy of Neuropeptide Prediction and Tandem Mass Spectrometry Identifies Evolutionarily Conserved Ancient Neuropeptides in the Sea Anemone Nematostella Vectensis. *PLoS One* **14,** e0215185. doi:10.1371/journal.pone.0215185. PMID: 31545805 (2019).

101. Sachkova, M. Y. *et al.* Toxin-like Neuropeptides in the Sea Anemone Nematostella Unravel Recruitment from the Nervous System to Venom. *Proc Natl Acad Sci U S A* **117,** 27481–27492. doi:10.1073/pnas.2011120117. PMID: 33060291 (2020).

102. Grimmelikhuijzen, C. J. P. & Hauser, F. Mini-Review: The Evolution of Neuropeptide Signaling. *Regul Pept* **177 Suppl,** S6–9. doi:10.1016/j.regpep.2012.05.001. PMID: 22726357 (2012).

103. Takahashi, T. & Takeda, N. Insight into the Molecular and Functional Diversity of Cnidarian Neuropeptides. *International Journal of Molecular Sciences* **16,** 2610–2625. doi:10.3390/ijms16022610 (2 2015).

104. Takahashi, T. Comparative Aspects of Structure and Function of Cnidarian Neuropeptides. *Front Endocrinol (Lausanne)* **11,** 339. doi:10.3389/fendo.2020.00339. PMID: 32528418 (2020).

105. Koch, T. L. & Grimmelikhuijzen, C. J. P. A Comparative Genomics Study of Neuropeptide Genes in the Cnidarian Subclasses Hexacorallia and Ceriantharia. *BMC Genomics* **21,** 666. doi:10.1186/s12864-020-06945-9. PMID: 32993486 (2020).

106. Yamamoto, W. & Yuste, R. Peptide-Driven Control of Somersaulting in *Hydra Vulgaris*. *Current Biology* **33,** 1893–1905.e4. doi:10.1016/j.cub.2023.03.047 (2023).

107. Prabhu, P. & Reddy, P. C. An in Silico Pipeline Identifies New Neuropeptides and Reveals a Non-Amidated Regulator of Muscle Contraction in the Freshwater Cnidarian Hydra. *Open Biol* **15,** 250073. doi:10.1098/rsob.250073 (2025).

108. Weissbourd, B. *et al.* A Genetically Tractable Jellyfish Model for Systems and Evolutionary Neuroscience. *Cell* **184,** 5854–5868.e20. doi:10.1016/j.cell.2021.10.021 (2021).

109. Quiroga Artigas, G. *et al.* A G Protein–Coupled Receptor Mediates Neuropeptide-Induced Oocyte Maturation in the Jellyfish Clytia. *PLoS Biol* **18,** e3000614. doi:10.1371/journal.pbio.3000614. PMID: 32126082 (2020).

110. Houliston, E., Momose, T. & Manuel, M. *Clytia Hemisphaerica*: A Jellyfish Cousin Joins the Laboratory. *Trends in Genetics* **26,** 159–167. doi:10.1016/j.tig.2010.01.008 (2010).

111. Nakanishi, N. & Martindale, M. Q. CRISPR Knockouts Reveal an Endogenous Role for Ancient Neuropeptides in Regulating Developmental Timing in a Sea Anemone. *eLife* **7,** e39742. doi:10.7554/eLife.39742. PMID: 30223943.

112. Golubovic, A. *et al.* A Peptide-gated Ion Channel from the Freshwater Polyp *Hydra. Journal of Biological Chemistry* **282,** 35098–35103. doi:10.1074/jbc.M706849200 (2007).

113. Assmann, M., Kuhn, A., Dürrnagel, S., Holstein, T. W. & Gründer, S. The Comprehensive Analysis of DEG/ENaC Subunits in Hydra Reveals a Large Variety of Peptide-Gated Channels, Potentially Involved in Neuromuscular Transmission. *BMC Biol* **12,** 84. doi:10.1186/s12915-014-0084-2. PMID: 25312679 (2014).

114. Gründer, S. & Assmann, M. Peptide-Gated Ion Channels and the Simple Nervous System of Hydra. *Journal of Experimental Biology* **218,** 551–561. doi:10.1242/jeb.111666 (2015).

115. Gründer, S., Ramírez, A. O. & Jékely, G. Neuropeptides and Degenerin/Epithelial Na+ Channels: A Relationship from Mammals to Cnidarians. *J Physiol* **601,** 1583–1595. doi:10.1113/JP282309. PMID: 36479972 (2023).

116. Aguilar-Camacho, J. M. *et al.* Functional Analysis in a Model Sea Anemone Reveals Phylogenetic Complexity and a Role in Cnidocyte Discharge of DEG/ENaC Ion Channels. *Commun Biol* **6,** 17. doi:10.1038/s42003-022-04399-1 (2023).

117. Williams, E. A. Function and Distribution of the Wamide Neuropeptide Superfamily in Metazoans. *Front Endocrinol (Lausanne)* **11,** 344. doi:10.3389/fendo.2020.00344. PMID: 32547494 (2020).

118. Babonis, L. S. That Stinging Sensation: Modularity and the Origin of the Stinging Cell. *Integr Comp Biol* **65,** 661–675. doi:10.1093/icb/icaf070 (2025).

119. Rentzsch, F., Juliano, C. & Galliot, B. Modern Genomic Tools Reveal the Structural and Cellular Diversity of Cnidarian Nervous Systems. *Current Opinion in Neurobiology. Neuronal Identity* **56,** 87–96. doi:10.1016/j.conb.2018.12.004 (2019).

120. Tournière, O. *et al.* NvPOU4/Brain3 Functions as a Terminal Selector Gene in the Nervous System of the Cnidarian Nematostella Vectensis. *Cell Rep* **30,** 4473–4489.e5. doi:10.1016/j.celrep.2020.03.031. PMID: 32234481 (2020).

121. Holstein, T. & Tardent, P. An Ultrahigh-Speed Analysis of Exocytosis: Nematocyst Discharge. *Science* **223,** 830–833. doi:10.1126/science.6695186 (1984).

122. Anderson, P. A. V. & Mckay, M. C. The Electrophysiology of Cnidocytes. *J Exp Biol* **133,** 215–230. doi:10.1242/jeb.133.1.215 (1987).

123. Anderson, P. A. V. & Bouchard, C. The Regulation of Cnidocyte Discharge. *Toxicon. Cnidarian Toxins and Venoms* **54,** 1046–1053. doi:10.1016/j.toxicon.2009.02.023 (2009).

124. Weber, J. Poly(Gamma-Glutamic Acid)s Are the Major Constituents of Nematocysts in Hydra (Hydrozoa, Cnidaria). *Journal of Biological Chemistry* **265,** 9664–9669. doi:10.1016/S0021-9258(19)38721-6 (1990).

125. Szczepanek, S., Cikala, M. & David, C. N. Poly--Glutamate Synthesis during Formation of Nematocyst Capsules in Hydra. *J Cell Sci* **115,** 745–751. doi:10.1242/jcs.115.4.745 (2002).

126. Özbek, S., Balasubramanian, P. G. & Holstein, T. W. Cnidocyst Structure and the Biomechanics of Discharge. *Toxicon. Cnidarian Toxins and Venoms* **54,** 1038–1045. doi:10.1016/j.toxicon.2009.03.006 (2009).

127. Pantin, C. F. A. The Excitation of Nematocysts. *J Exp Biol* **19,** 294–310. doi:10.1242/jeb.19.3.294 (1942).

128. Gitter, A. H., Oliver, D. & Thurm, U. Calcium- and Voltage-Dependence of Nematocyst Discharge in Hydra Vulgaris. *J Comp Physiol A* **175,** 115–122. doi:10.1007/BF00217442 (1994).

129. Watson, G. M. & Hessinger, D. A. Evidence for Calcium Channels Involved in Regulating Nematocyst Discharge. *Comparative Biochemistry and Physiology Part A: Physiology* **107,** 473–481. doi:10.1016/0300-9629(94)90028-0 (1994).

130. Karabulut, A. *et al.* The Architecture and Operating Mechanism of a Cnidarian Stinging Organelle. *Nat Commun* **13,** 3494. doi:10.1038/s41467-022-31090-0 (2022).

131. Özbek, S. *et al.* The Glycoprotein NOWA and Minicollagens Are Part of a Disulfidelinked Polymer That Forms the Cnidarian Nematocyst Wall. *J. Biol. Chem.* **279,** 52016–52023. doi:10.1074/jbc.M407613200. PMID: 15466471 (2004).

132. Zenkert, C., Takahashi, T., Diesner, M.-O. & Özbek, S. Morphological and Molecular Analysis of the Nematostella Vectensis Cnidom. *PLoS One* **6,** e22725. doi:10.1371/journal.pone.0022725. PMID: 21829492 (2011).

133. Beckmann, A. & Özbek, S. The Nematocyst: A Molecular Map of the Cnidarian Stinging Organelle. *Int. J. Dev. Biol.* **56,** 577–582. doi:10.1387/ijdb.113472ab. PMID: 22689365 (6-7-8 2012).

134. Weill, R. *Contribution à l'étude des cnidaires et de leurs nématocystes* 2 vols. *Travaux de la Station zoologique de Wimereux, publiés sous la direction de M. Caullery* t. 10-11 (Laboratoire d'évolution des êtres organisés, Paris, 1934).

135. England, K. W. Nematocysts of Sea Anemones (Actiniaria, Ceriantharia and Corallimorpharia: Cnidaria): Nomenclature. *Hydrobiologia* **216,** 691–697. doi:10.1007/BF00026532 (1991).

136. Babonis, L. S., Enjolras, C., Ryan, J. F. & Martindale, M. Q. A Novel Regulatory Gene Promotes Novel Cell Fate by Suppressing Ancestral Fate in the Sea Anemone Nematostella Vectensis. *Proc Natl Acad Sci U S A* **119,** e2113701119. doi:10.1073/pnas.2113701119. PMID: 35500123 (2022).

137. Hwang, J. S. *et al.* Nematogalectin, a Nematocyst Protein with GlyXY and Galectin Domains, Demonstrates Nematocyte-Specific Alternative Splicing in Hydra. *Proceedings of the National Academy of Sciences* **107,** 18539–18544. doi:10.1073/pnas.1003256107 (2010).

138. Cazet, J. F. *et al.* A Chromosome-Scale Epigenetic Map of the Hydra Genome Reveals Conserved Regulators of Cell State. *Genome Res* **33,** 283–298. doi:10.1101/gr.277040.122. PMID: 36639202 (2023).

139. Siebert, S., Anton-Erxleben, F. & Bosch, T. C. G. Cell Type Complexity in the Basal Metazoan Hydra Is Maintained by Both Stem Cell Based Mechanisms and Transdifferentiation. *Dev Biol* **313,** 13–24. doi:10.1016/j.ydbio.2007.09.007. PMID: 18029279 (2008).

140. Song, J., Jong, D. de, Waletich, J., Baxevanis, A. D. & Schnitzler, C. E. *An Updated and Spatially Validated Somatic Single-Cell Atlas of Hydractinia Symbiolongicarpus* https://www.biorxiv.org/content/10.1101/2025.06.03.657738v1 (2025). Pre-published.

141. Salamanca-Díaz, D. A. *et al.* The Hydractinia Cell Atlas Reveals Cellular and Molecular Principles of Cnidarian Coloniality. *Nat Commun* **16,** 2121. doi:10.1038/s41467-025-57168-z (2025).

142. Levy, S. *et al.* A Stony Coral Cell Atlas Illuminates the Molecular and Cellular Basis of Coral Symbiosis, Calcification, and Immunity. *Cell* **184,** 2973–2987.e18. doi:10.1016/j.cell.2021.04.005. PMID: 33945788 (2021).

143. Hu, M., Zheng, X., Fan, C.-M. & Zheng, Y. Lineage Dynamics of the Endosymbiotic Cell Type in the Soft Coral Xenia. *Nature* **582,** 534–538. doi:10.1038/s41586-020-2385-7 (7813 2020).

144. Westfall, J. A. Ultrastructure of Synapses in the First-Evolved Nervous Systems. *J Neurocytol* **25,** 735–746. doi:10.1007/BF02284838 (1996).

145. Denker, E., Manuel, M., Leclère, L., Le Guyader, H. & Rabet, N. Ordered Progression of Nematogenesis from Stem Cells through Differentiation Stages in the Tentacle Bulb of *Clytia Hemisphaerica* (Hydrozoa, Cnidaria). *Developmental Biology* **315,** 99–113. doi:10.1016/j.ydbio.2007.12.023 (2008).

146. Gold, D. A. *et al.* Mechanisms of Cnidocyte Development in the Moon Jellyfish Aurelia. *Evol. Dev.* **21,** 72–81. doi:10.1111/ede.12278. PMID: 30623570 (2019).

147. Holstein, T. W. The Significance of Ethel Browne's Research on *Hydra* for the Organizer Concept. *Cells & Development* **178,** 203907. doi:10.1016/j.cdev.2024.203907 (2024).

148. Browne, E. N. The Production of New Hydranths in Hydra by the Insertion of Small Grafts. *Journal of Experimental Zoology* **7,** 1–23. doi:10.1002/jez.1400070102 (1909).

149. Vogg, M. C., Galliot, B. & Tsiairis, C. D. Model Systems for Regeneration: Hydra. *Development* **146,** dev177212. doi:10.1242/dev.177212 (2019).

150. Steinmetz, P. R. H. *et al.* Independent Evolution of Striated Muscles in Cnidarians and Bilaterians. *Nature* **487,** 231–234. doi:10.1038/nature11180. PMID: 22763458 (2012).

151. Steinmetz, P. R., Aman, A., Kraus, J. E. & Technau, U. Gut-like Ectodermal Tissue in a Sea Anemone Challenges Germ Layer Homology. *Nat Ecol Evol* **1,** 1535–1542. doi:10.1038/s41559-017-0285-5. PMID: 29185520 (2017).

152. Haillot, E. *et al.* Segregation of Endoderm and Mesoderm Germ Layer Identities in the Diploblast Nematostella Vectensis. *Nat Commun* **16,** 7979. doi:10.1038/s41467-025-63287-4. PMID: 40858588 (2025).

153. Lebedeva, T. *et al.* -Catenin-Driven Endomesoderm Specification Is a Bilateria-specific Novelty. *Nat Commun* **16,** 2476. doi:10.1038/s41467-025-57109-w. PMID: 40075083 (2025).

154. Koizumi, O. *et al.* The Nerve Ring in Cnidarians: Its Presence and Structure in Hydrozoan Medusae. *Zoology. Animal Evolution: Early Emerging Animals Matter* **118,** 79–88. doi:10.1016/j.zool.2014.10.001 (2015).

155. Birch, S., Picciani, N., Oakley, T. & Plachetzki, D. in *Distributed Vision: From Simple Sensors to Sophisticated Combination Eyes* 21–47 (Springer International Publishing, Cham, 2023). doi:10.1007/978-3-031-23216-9_2.

156. Hand, C. & Uhlinger, K. R. The Culture, Sexual and Asexual Reproduction, and Growth of the Sea Anemone Nematostella Vectensis. *Biological Bulletin* **182,** 169–176. doi:10.2307/1542110. JSTOR: 1542110 (1992).

157. Fritzenwanker, J. H. & Technau, U. Induction of Gametogenesis in the Basal Cnidarian Nematostella Vectensis (Anthozoa). *Dev Genes Evol* **212,** 99–103. doi:10.1007/s00427-002-0214-7 (2002).

158. Röttinger, E. Nematostella Vectensis, an Emerging Model for Deciphering the Molecular and Cellular Mechanisms Underlying Whole-Body Regeneration. *Cells* **10,** 2692. doi:10.3390/cells10102692. PMID: 34685672 (2021).

159. Layden, M. J., Rentzsch, F. & Röttinger, E. The Rise of the Starlet Sea Anemone Nematostella Vectensis as a Model System to Investigate Development and Regeneration. *Wiley Interdiscip Rev Dev Biol* **5,** 408–428. doi:10.1002/wdev.222. PMID: 26894563 (2016).

160. Putnam, N. H. *et al.* Sea Anemone Genome Reveals Ancestral Eumetazoan Gene Repertoire and Genomic Organization. *Science* **317,** 86–94. doi:10.1126/science.1139158. PMID: 17615350 (2007).

161. Srivastava, M. *et al.* The Trichoplax Genome and the Nature of Placozoans. *Nature* **454,** 955–960. doi:10.1038/nature07191 (2008).

162. Srivastava, M. *et al.* The Amphimedon Queenslandica Genome and the Evolution of Animal Complexity. *Nature* **466,** 720–726. doi:10.1038/nature09201. PMID: 20686567 (2010).

163. *The Genome Sequence of the Starlet Sea Anemone,* Nematostella ... Wellcome Open Research | Open Access Publishing Platform. https://wellcomeopenresearch.org/articles/8-79 (2024).

164. Marlow, H. in *Evolutionary Ecology of Marine Invertebrate Larvae* 16–33 (Oxford University Press, 2017). doi:10.1093/oso/9780198786962.003.0002.

165. Ikmi, A. *et al.* Feeding-Dependent Tentacle Development in the Sea Anemone Nematostella Vectensis. *Nat Commun* **11.** doi:10.1038/s41467-020-18133-0. PMID: 32879319 (2020).

166. Garschall, K. *et al.* The Cellular Basis of Feeding-Dependent Body Size Plasticity in Sea Anemones. *Development* **151,** dev202926. doi:10.1242/dev.202926 (2024).

167. He, S. *et al.* An Axial Hox Code Controls Tissue Segmentation and Body Patterning in Nematostella Vectensis. *Science* **361,** 1377–1380. doi:10.1126/science.aar8384 (2018).

168. DuBuc, T. Q., Stephenson, T. B., Rock, A. Q. & Martindale, M. Q. Hox and Wnt Pattern the Primary Body Axis of an Anthozoan Cnidarian before Gastrulation. *Nat Commun* **9,** 2007. doi:10.1038/s41467-018-04184-x. PMID: 29789526 (2018).

169. Renfer, E., Amon-Hassenzahl, A., Steinmetz, P. R. H. & Technau, U. A Muscle-Specific Transgenic Reporter Line of the Sea Anemone, Nematostella Vectensis. *Proc Natl Acad Sci U S A* **107,** 104–108. doi:10.1073/pnas.0909148107. PMID: 20018670 (2010).

170. Leclère, L. & Rentzsch, F. RGM Regulates BMP-Mediated Secondary Axis Formation in the Sea Anemone Nematostella Vectensis. *Cell Reports* **9,** 1921–1930. doi:10.1016/j.celrep.2014.11.009. PMID: 25482565 (2014).

171. Watanabe, H. *et al.* Sequential Actions of -Catenin and Bmp Pattern the Oral Nerve Net in Nematostella Vectensis. *Nat Commun* **5,** 5536. doi:10.1038/ncomms6536 (2014).

172. Genikhovich, G. & Technau, U. On the Evolution of Bilaterality. *Development* **144,** 3392–3404. doi:10.1242/dev.141507 (2017).

173. Cole, A. G. *et al.* Muscle Cell-Type Diversification Is Driven by bHLH Transcription Factor Expansion and Extensive Effector Gene Duplications. *Nat Commun* **14,** 1747. doi:10.1038/s41467-023-37220-6 (2023).

174. Cole, A. G. *et al.* Updated Single Cell Reference Atlas for the Starlet Anemone Nematostella Vectensis. *Front Zool* **21,** 8. doi:10.1186/s12983-024-00529-z. PMID: 38500146 (2024).

175. Bosch, T. C. G. *et al.* Back to the Basics: Cnidarians Start to Fire. *Trends Neurosci* **40,** 92–105. doi:10.1016/j.tins.2016.11.005. PMID: 28041633 (2017).

176. Sprecher, S. G. Neural Cell Type Diversity in Cnidaria. *Frontiers in Neuroscience* **16.** doi:10.3389/fnins.2022.909400. PMID: 35685775 (2022).

177. Layden, M. J. *et al.* MAPK Signaling Is Necessary for Neurogenesis in Nematostella Vectensis. *BMC Biology* **14,** 61. doi:10.1186/s12915-016-0282-1 (2016).

178. Havrilak, J. A. *et al.* Characterization of NvLWamide-like Neurons Reveals Stereotypy in Nematostella Nerve Net Development. *Developmental Biology* **431,** 336–346. doi:10.1016/j.ydbio.2017.08.028 (2017).

179. Fritzsche, S. & Hunnekuhl, V. S. Cell-Specific Expression and Individual Function of Prohormone Convertase PC1/3 in Tribolium Larval Growth Highlights Major Evolutionary Changes between Beetle and Fly Neuroendocrine Systems. *EvoDevo* **12,** 9. doi:10.1186/s13227-021-00179-w (2021).

180. Busengdal, H. & Rentzsch, F. Unipotent Progenitors Contribute to the Generation of Sensory Cell Types in the Nervous System of the Cnidarian Nematostella Vectensis. *Dev. Biol.* **431,** 59–68. doi:10.1016/j.ydbio.2017.08.021. PMID: 28827097 (2017).

181. Lemaître, Q. I. B. *et al.* NvPrdm14d-expressing Neural Progenitor Cells Contribute to Non-Ectodermal Neurogenesis in Nematostella Vectensis. *Nat Commun* **14,** 4854. doi:10.1038/s41467-023-39789-4. PMID: 37563174 (2023).

182. Hauser, F., Koch, T. L. & Grimmelikhuijzen, C. J. P. Review: The Evolution of Peptidergic Signaling in Cnidaria and Placozoa, Including a Comparison with Bilateria. *Front Endocrinol (Lausanne)* **13,** 973862. doi:10.3389/fendo.2022.973862. PMID: 36213267 (2022).

183. Zang, H. & Nakanishi, N. Expression Analysis of Cnidarian-Specific Neuropeptides in a Sea Anemone Unveils an Apical-Organ-Associated Nerve Net That Disintegrates at Metamorphosis. *Front Endocrinol (Lausanne)* **11,** 63. doi:10.3389/fendo.2020.00063. PMID: 32140137 (2020).

184. Zhang, Y. *et al.* Involvement of RFamide Neuropeptides in Polyp Contraction of the Adult Scleractinian Corals *Euphyllia Ancora* and *Stylophora Pistillata. General and Comparative Endocrinology* **314,** 113905. doi:10.1016/j.ygcen.2021.113905 (2021).

185. Babonis, L. S. & Martindale, M. Q. Old Cell, New Trick? Cnidocytes as a Model for the Evolution of Novelty. *Integr Comp Biol* **54,** 714–722. doi:10.1093/icb/icu027. PMID: 24771087 (2014).

186. Babonis, L. S., Martindale, M. Q. & Ryan, J. F. Do Novel Genes Drive Morphological Novelty? An Investigation of the Nematosomes in the Sea Anemone Nematostella Vectensis. *BMC Evol Biol* **16,** 114. doi:10.1186/s12862-016-0683-3. PMID: 27216622 (2016).

187. Columbus-Shenkar, Y. Y. *et al.* Dynamics of Venom Composition across a Complex Life Cycle. *eLife* **7,** e35014. doi:10.7554/eLife.35014 (2018).

188. Babonis, L. S. & Martindale, M. Q. PaxA, but Not PaxC, Is Required for Cnidocyte Development in the Sea Anemone Nematostella Vectensis. *EvoDevo* **8,** 14. doi:10.1186/s13227-017-0077-7. PMID: 28878874 (2017).

189. Moran, Y. *et al.* Analysis of Soluble Protein Contents from the Nematocysts of a Model Sea Anemone Sheds Light on Venom Evolution. *Mar Biotechnol* **15,** 329–339. doi:10.1007/s10126-012-9491-y (2013).

190. Marlow, H., Matus, D. Q. & Martindale, M. Q. Ectopic Activation of the Canonical Wnt Signaling Pathway Affects Ectodermal Patterning along the Primary Axis during Larval Development in the Anthozoan Nematostella Vectensis. *Dev Biol* **380,** 324–334. doi:10.1016/j.ydbio.2013.05.022. PMID: 23722001 (2013).

191. Layden, M. J. & Martindale, M. Q. Non-Canonical Notch Signaling Represents an Ancestral Mechanism to Regulate Neural Differentiation. *Evodevo* **5,** 30. doi:10.1186/2041-9139-5-30. PMID: 25705370 (2014).

192. Babonis, L. S. On the Evolutionary Developmental Biology of the Cell. *Trends in Genetics.* doi:10.1016/j.tig.2024.06.003 (2024).

193. Babonis, L. S. *et al.* Single-Cell Atavism Reveals an Ancient Mechanism of Cell Type Diversification in a Sea Anemone. *Nat Commun* **14,** 885. doi:10.1038/s41467-023-36615-9 (2023).

194. Sunagar, K. *et al.* Cell Type-Specific Expression Profiling Unravels the Development and Evolution of Stinging Cells in Sea Anemone. *BMC Biol.* **16,** 108. doi:10.1186/s12915-018-0578-4. PMID: 30261880 (2018).

195. Gahan, J. M., Kouzel, I. U., Jansen, K. O., Burkhardt, P. & Rentzsch, F. Histone Demethylase Lsd1 Is Required for the Differentiation of Neural Cells in Nematostella Vectensis. *Nat Commun* **13,** 465. doi:10.1038/s41467-022-28107-z. PMID: 35075108 (2022).

196. Ozment, E. *et al.* Cnidarian Hair Cell Development Illuminates an Ancient Role for the Class IV POU Transcription Factor in Defining Mechanoreceptor Identity. *Elife* **10,** e74336. doi:10.7554/eLife.74336. PMID: 34939935 (2021).

197. Baranyk, J. *et al.* Structural, Molecular and Developmental Evidence for Cell-Type Diversity in Cnidarian Mechanosensory Neurons. *Nat Commun* **16,** 1514. doi:10.1038/s41467-025-56115-2. PMID: 39929800 (2025).

198. Gahan, J. M., Leclère, L., Hernandez-Valladares, M. & Rentzsch, F. A Developmental Role for the Chromatin-Regulating CoREST Complex in the Cnidarian Nematostella Vectensis. *BMC Biol* **20,** 184. doi:10.1186/s12915-022-01385-1. PMID: 35999597 (2022).

199. Tournière, O., Gahan, J. M., Busengdal, H., Bartsch, N. & Rentzsch, F. Insm1-Expressing Neurons and Secretory Cells Develop from a Common Pool of Progenitors in the Sea Anemone Nematostella Vectensis. *Sci Adv* **8,** eabi7109. doi:10.1126/sciadv.abi7109. PMID: 35442742.

200. Levy, S. *et al.* Ectopic Activation of GABAB Receptors Inhibits Neurogenesis and Metamorphosis in the Cnidarian Nematostella Vectensis. *Nat Ecol Evol* **5,** 111–121. doi:10.1038/s41559-020-01338-3. PMID: 33168995 (2021).

201. Havrilak, J. A., Al-Shaer, L., Baban, N., Akinci, N. & Layden, M. J. Characterization of the Dynamics and Variability of Neuronal Subtype Responses during Growth, Degrowth, and Regeneration of Nematostella Vectensis. *BMC Biol* **19,** 104. doi:10.1186/s12915-021-01038-9. PMID: 34001126 (2021).

202. Miramón-Puértolas, P., Pascual-Carreras, E. & Steinmetz, P. R. H. A Population of Vasa2 and Piwi1 Expressing Cells Generates Germ Cells and Neurons in a Sea Anemone. *Nat Commun* **15,** 8765. doi:10.1038/s41467-024-52806-4. PMID: 39384751 (2024).

203. Denner, A. *et al.* Nanos2 Marks Precursors of Somatic Lineages and Is Required for Germline Formation in the Sea Anemone Nematostella Vectensis. *Sci Adv* **10,** eado0424. doi:10.1126/sciadv.ado0424. PMID: 39151009 (2024).

204. Pang, K., Matus, D. Q. & Martindale, M. Q. The Ancestral Role of COE Genes May Have Been in Chemoreception: Evidence from the Development of the Sea Anemone, Nematostella Vectensis (Phylum Cnidaria; Class Anthozoa). *Dev Genes Evol* **214,** 134–138. doi:10.1007/s00427-004-0383-7 (2004).

205. Finnerty, J. R. Cnidarians Reveal Intermediate Stages in the Evolution of Hox Clusters and Axial Complexity1. *American Zoologist* **41,** 608–620. doi:10.1093/icb/41.3.608 (2001).

206. Layden, M. J., Röttinger, E., Wolenski, F. S., Gilmore, T. D. & Martindale, M. Q. Microinjection of mRNA or Morpholinos for Reverse Genetic Analysis in the Starlet Sea Anemone, Nematostella Vectensis. *Nat Protoc* **8,** 924–934. doi:10.1038/nprot.2013.009. PMID: 23579781 (2013).

207. Ikmi, A., McKinney, S. A., Delventhal, K. M. & Gibson, M. C. TALEN and CRISPR/Cas9-mediated Genome Editing in the Early-Branching Metazoan Nematostella Vectensis. *Nat Commun* **5,** 5486. doi:10.1038/ncomms6486 (2014).

208. Schwaiger, M. *et al.* Evolutionary Conservation of the Eumetazoan Gene Regulatory Landscape. *Genome Res.* **24,** 639–650. doi:10.1101/gr.162529.113. PMID: 24642862 (2014).

209. Moran, Y. *et al.* Cnidarian microRNAs Frequently Regulate Targets by Cleavage. *Genome Res.* **24,** 651–663. doi:10.1101/gr.162503.113. PMID: 24642861 (2014).

210. Marlow, H., Roettinger, E., Boekhout, M. & Martindale, M. Q. Functional Roles of Notch Signaling in the Cnidarian Nematostella Vectensis. *Dev Biol* **362,** 295–308. doi:10.1016/j.ydbio.2011.11.012. PMID: 22155407 (2012).

211. Röttinger, E., Dahlin, P. & Martindale, M. Q. A Framework for the Establishment of a Cnidarian Gene Regulatory Network for "Endomesoderm" Specification: The Inputs of ß-Catenin/TCF Signaling. *PLoS Genet* **8.** doi:10.1371/journal.pgen.1003164. PMID: 23300467 (2012).

212. Lee, P. N., Kumburegama, S., Marlow, H. Q., Martindale, M. Q. & Wikramanayake, A. H. Asymmetric Developmental Potential along the Animal–Vegetal Axis in the Anthozoan Cnidarian, *Nematostella Vectensis*, Is Mediated by Dishevelled. *Developmental Biology* **310,** 169–186. doi:10.1016/j.ydbio.2007.05.040 (2007).

213. Tritschler, S. *et al.* Concepts and Limitations for Learning Developmental Trajectories from Single Cell Genomics. *Development* **146,** dev170506. doi:10.1242/dev.170506 (2019).

214. Nowotschin, S. & Hadjantonakis, A.-K. Use of KikGR a Photoconvertible Green-to-Red Fluorescent Protein for Cell Labeling and Lineage Analysis in ES Cells and Mouse Embryos. *BMC Dev. Biol.* **9,** 49. doi:10.1186/1471-213X-9-49. PMID: 19740427 (2009).

215. Tsutsui, H., Karasawa, S., Shimizu, H., Nukina, N. & Miyawaki, A. Semi-Rational Engineering of a Coral Fluorescent Protein into an Efficient Highlighter. *EMBO Rep* **6,** 233–238. doi:10.1038/sj.embor.7400361. PMID: 15731765 (2005).

216. Patterson, G. H. & Lippincott-Schwartz, J. A Photoactivatable GFP for Selective Photolabeling of Proteins and Cells. *Science* **297,** 1873–1877. doi:10.1126/science.1074952 (2002).

217. Bastide, S. *et al.* TATTOO-seq Delineates Spatial and Cell Type–Specific Regulatory Programs in the Developing Limb. *Science Advances* **8,** eadd0695. doi:10.1126/scidv.add0695 (2022).

218. Plessier, F., Schmutz, S., Novault, S. & Marlow, H. Single-Cell Transcriptomic Analysis in the Regenerating Cnidarian Nematostella Vectensis. *Methods Mol Biol* **2450,** 565–581. doi:10.1007/978-1-0716-2172-1_30. PMID: 35359329 (2022).

219. Pachitariu, M. & Stringer, C. Cellpose 2.0: How to Train Your Own Model. *Nat Methods* **19,** 1634–1641. doi:10.1038/s41592-022-01663-4 (2022).

220. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: A Generalist Algorithm for Cellular Segmentation. *Nat Methods* **18,** 100–106. doi:10.1038/s41592-020-01018-x (2021).

221. Walt, S. van der *et al.* Scikit-Image: Image Processing in Python. *PeerJ* **2,** e453. doi:10.7717/peerj.453 (2014).

222. Amiel, A. R. *et al.* Characterization of Morphological and Cellular Events Underlying Oral Regeneration in the Sea Anemone, Nematostella Vectensis. *Int J Mol Sci* **16,** 28449–28471. doi:10.3390/ijms161226100. PMID: 26633371 (2015).

223. Bradshaw, B., Thompson, K. & Frank, U. Distinct Mechanisms Underlie Oral vs Aboral Regeneration in the Cnidarian Hydractinia Echinata. *eLife* **4,** e05506. doi:10.7554/eLife.05506 (2015).

224. Rentzsch, F. & Technau, U. Genomics and Development of Nematostella Vectensis and Other Anthozoans. *Current Opinion in Genetics & Development. Developmental Mechanisms, Patterning and Evolution* **39,** 63–70. doi:10.1016/j.gde.2016.05.024 (2016).

225. Amiel, A. R., Foucher, K., Ferreira, S. & Röttinger, E. Synergic Coordination of Stem Cells Is Required to Induce a Regenerative Response in Anthozoan Cnidarians. *bioRxiv,* 2019.12.31.891804. doi:10.1101/2019.12.31.891804 (2019).

226. Schaffer, A. A., Bazarsky, M., Levy, K., Chalifa-Caspi, V. & Gat, U. A Transcriptional Time-Course Analysis of Oral vs. Aboral Whole-Body Regeneration in the Sea Anemone Nematostella Vectensis. *BMC Genomics* **17,** 718. doi:10.1186/s12864-016-3027-1. PMID: 27605362 (2016).

227. Warner, J. F. *et al.* NvERTx: A Gene Expression Database to Compare Embryogenesis and Regeneration in the Sea Anemone Nematostella Vectensis. *Development* **145.** doi:10.1242/dev.162867. PMID: 29739837 (2018).

228. Warner, J. F., Amiel, A. R., Johnston, H. & Röttinger, E. Regeneration Is a Partial Redeployment of the Embryonic Gene Network. *bioRxiv,* 658930. doi:10.1101/658930 (2019).

229. Watanabe, H., Hoang, V. T., Mättner, R. & Holstein, T. W. Immortality and the Base of Multicellular Life: Lessons from Cnidarian Stem Cells. *Semin. Cell Dev. Biol.* **20,** 1114–1125. doi:10.1016/j.semcdb.2009.09.008. PMID: 19761866 (2009).

230. Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C. & Teichmann, S. A. The Technology and Biology of Single-Cell RNA Sequencing. *Molecular Cell* **58,** 610–620. doi:10.1016/j.molcel.2015.04.005 (2015).

231. Ziegenhain, C. *et al.* Comparative Analysis of Single-Cell RNA Sequencing Methods. *Molecular Cell* **65,** 631–643.e4. doi:10.1016/j.molcel.2017.01.023 (2017).

232. Jaitin, D. A. *et al.* Massively Parallel Single Cell RNA-Seq for Marker-Free Decomposition of Tissues into Cell Types. *Science* **343,** 776–779. doi:10.1126/science.1247651. PMID: 24531970 (2014).

233. Kivioja, T. *et al.* Counting Absolute Numbers of Molecules Using Unique Molecular Identifiers. *Nat Methods* **9,** 72–74. doi:10.1038/nmeth.1778 (2012).

234. Baran, Y. *et al.* MetaCell: Analysis of Single-Cell RNA-seq Data Using K-nn Graph Partitions. *Genome Biol* **20.** doi:10.1186/s13059-019-1812-2. PMID: 31604482 (2019).

235. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177,** 1888–1902.e21. doi:10.1016/j.cell.2019.05.031. PMID: 31178118 (2019).

236. Kiselev, V. Y. *et al.* SC3: Consensus Clustering of Single-Cell RNA-seq Data. *Nat Methods* **14,** 483–486. doi:10.1038/nmeth.4236 (2017).

237. Keren-Shaul, H. *et al.* MARS-seq2.0: An Experimental and Analytical Pipeline for Indexed Sorting Combined with Single-Cell RNA Sequencing. *Nat Protoc* **14,** 1841–1862. doi:10.1038/s41596-019-0164-4 (2019).

238. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial Reconstruction of Single-Cell Gene Expression Data. *Nat. Biotechnol.* **33,** 495–502. doi:10.1038/nbt.3192. PMID: 25867923 (2015).

239. Rodrigues, O. R. & Monard, S. A Rapid Method to Verify Single-Cell Deposition Setup for Cell Sorters. *Cytometry Part A* **89,** 594–600. doi:10.1002/cyto.a.22865 (2016).

240. Plessier, F. & Marlow, H. Cellular and Transcriptional Trajectories of Neural Fate Specification in Sea Anemone Uncover Two Modes of Adult Neurogenesis. *Nat Commun.* doi:10.1038/s41467-026-68802-9 (2026).

241. Srivastava, M. Beyond Casual Resemblance: Rigorous Frameworks for Comparing Regeneration Across Species. *Annual Review of Cell and Developmental Biology* **37,** 415–440. doi:10.1146/annurev-cellbio-120319-114716 (Volume 37, 2021 2021).

242. Layden, M. J. in *Zic Family: Evolution, Development and Disease* 27–39 (Springer, Singapore, 2018). doi:10.1007/978-981-10-7311-3_2.

243. Sebens, K. P. THE ECOLOGY OF INDETERMINATE GROWTH IN ANIMALS. *Annual Review of Ecology, Evolution, and Systematics* **18,** 371–407. doi:10.1146/annurev.es.18.110187.002103 (Volume 18, 1987 1987).

244. Baguñà, J. *et al.* in *Experimental Embryology in Aquatic Plants and Animals* 129–162 (Springer US, Boston, MA, 1990). doi:10.1007/978-1-4615-3830-1_7.

245. Labusch, M., Mancini, L., Morizet, D. & Bally-Cuif, L. Conserved and Divergent Features of Adult Neurogenesis in Zebrafish. *Front Cell Dev Biol* **8,** 525. doi:10.3389/fcell.2020.00525. PMID: 32695781 (2020).

246. Wyss, L. S., Bray, S. R. & Wang, B. Cellular Diversity and Developmental Hierarchy in the Planarian Nervous System. *Current Opinion in Genetics & Development* **76,** 101960. doi:10.1016/j.gde.2022.101960 (2022).

247. Van Wolfswinkel, J. C. Piwi and Potency: PIWI Proteins in Animal Stem Cells and Regeneration. *Integrative and Comparative Biology* **54,** 700–713. doi:10.1093/icb/icu084 (2014).

248. Fritz, A. E., Ikmi, A., Seidel, C., Paulson, A. & Gibson, M. C. Mechanisms of Tentacle Morphogenesis in the Sea Anemone Nematostella Vectensis. *Development* **140,** 2212–2223. doi:10.1242/dev.088260. PMID: 23633514 (2013).

249. Siebert, S. *et al.* Stem Cell Differentiation Trajectories in Hydra Resolved at Single-Cell Resolution. *Science* **365,** eaav9314. doi:10.1126/science.aav9314. PMID: 31346039 (2019).

250. Li, X. *et al.* Ether-à-Go-Go Family Voltage-Gated K+ Channels Evolved in an Ancestral Metazoan and Functionally Diversified in a Cnidarian–Bilaterian Ancestor. *J Exp Biol* **218,** 526–536. doi:10.1242/jeb.110080. PMID: 25696816 (2015).

251. Thorington, G. U. & Hessinger, D. A. Activated L-Type Calcium Channels Inhibit Chemosensitized Nematocyst Discharge from Sea Anemone Tentacles. *The Biological Bulletin* **238,** 180–192. doi:10.1086/708947 (2020).

252. Langfelder, P. & Horvath, S. WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinformatics* **9,** 559. doi:10.1186/1471-2105-9-559 (2008).

253. Paul, V. *et al.* Scratch2 Modulates Neurogenesis and Cell Migration through Antagonism of bHLH Proteins in the Developing Neocortex. *Cereb Cortex* **24,** 754–772. doi:10.1093/cercor/bhs356. PMID: 23180754 (2014).

254. Dattoli, A. A. *et al.* Domain Analysis of the Nematostella Vectensis SNAIL Ortholog Reveals Unique Nucleolar Localization That Depends on the Zinc-Finger Domains. *Sci Rep* **5,** 12147. doi:10.1038/srep12147. PMID: 26190255 (2015).

255. Talice, S. *et al.* Candidate Stem Cell Isolation and Transplantation in Hexacorallia. *Cell Reports* **43,** 114944. doi:10.1016/j.celrep.2024.114944 (2024).

256. Pascual-Carreras, E., Garschall, K. & Steinmetz, P. R. H. Prolonged Starvation Deepens Quiescence in Vasa2/Piwi1-expressing Cells of a Sea Anemone. *PLoS Biol* **23,** e3003525. doi:10.1371/journal.pbio.3003525. PMID: 41359646 (2025).

257. Varley, Á., Horkan, H. R., McMahon, E. T., Krasovec, G. & Frank, U. Pluripotent, Germ Cell Competent Adult Stem Cells Underlie Cnidarian Regenerative Ability and Clonal Growth. *Curr Biol* **33,** 1883–1892.e3. doi:10.1016/j.cub.2023.03.039. PMID: 37028430 (2023).

258. Hobmayer, B. *et al.* Stemness in Hydra - a Current Perspective. *Int J Dev Biol* **56,** 509–517. doi:10.1387/ijdb.113426bh. PMID: 22689357 (2012).

259. Raz, A. A., Wurtzel, O. & Reddien, P. W. Planarian Stem Cells Specify Fate yet Retain Potency during the Cell Cycle. *Cell Stem Cell* **28,** 1307–1322.e5. doi:10.1016/j.stem.2021.03.021. PMID: 33882291 (2021).

260. Holguera, I. & Desplan, C. Neuronal Specification in Space and Time. *Science* **362,** 176–180. doi:10.1126/science.aas9435. PMID: 30309944 (2018).

261. Engel, U. *et al.* A Switch in Disulfide Linkage during Minicollagen Assembly in Hydra Nematocysts. *EMBO J* **20,** 3063–3073. doi:10.1093/emboj/20.12.3063. PMID: 11406583 (2001).

262. Tursch, A., Mercadante, D., Tennigkeit, J., Gräter, F. & Özbek, S. Minicollagen Cysteine-Rich Domains Encode Distinct Modes of Polymerization to Form Stable Nematocyst Capsules. *Sci Rep* **6,** 25709. doi:10.1038/srep25709 (2016).

263. David, C. N. & Gierer, A. Cell Cycle Kinetics and Development of Hydra Attenuata: III. Nerve and Nematocyte Differentiation. *J Cell Sci* **16,** 359–375. doi:10.1242/jcs.16.2.359 (1974).

264. Lara, A., Simonson, B. T., Ryan, J. F. & Jegla, T. Genome-Scale Analysis Reveals Extensive Diversification of Voltage-Gated K+ Channels in Stem Cnidarians. *Genome Biology and Evolution* **15,** evad009. doi:10.1093/gbe/evad009 (2023).

265. Kavyanifar, A., Turan, S. & Lie, D. C. SoxC Transcription Factors: Multifunctional Regulators of Neurodevelopment. *Cell Tissue Res* **371,** 91–103. doi:10.1007/s00441-017-2708-7. PMID: 29079881 (2018).

266. Ryan, J. F. *et al.* The Cnidarian-Bilaterian Ancestor Possessed at Least 56 Homeoboxes: Evidence from the Starlet Sea Anemone, Nematostella Vectensis. *Genome Biology* **7,** R64. doi:10.1186/gb-2006-7-7-r64 (2006).

267. Reilly, M. B., Cros, C., Varol, E., Yemini, E. & Hobert, O. Unique Homeobox Codes Delineate All the Neuron Classes of C. Elegans. *Nature* **584,** 595–601. doi:10.1038/s41586-020-2618-9. PMID: 32814896 (2020).

268. Smith, J. J. *et al.* A Molecular Atlas of Adult C. Elegans Motor Neurons Reveals Ancient Diversity Delineated by Conserved Transcription Factor Codes. *Cell Rep* **43,** 113857. doi:10.1016/j.celrep.2024.113857. PMID: 38421866 (2024).

269. Hobert, O. Homeobox Genes and the Specification of Neuronal Identity. *Nat Rev Neurosci* **22,** 627–636. doi:10.1038/s41583-021-00497-x (2021).

270. Philippidou, P. & Dasen, J. S. Hox Genes: Choreographers in Neural Development, Architects of Circuit Organization. *Neuron* **80,** 12–34. doi:10.1016/j.neuron.2013.09.020. PMID: 24094100 (2013).

271. Hamilton, N. E. & Ferry, M. Ggtern: Ternary Diagrams Using Ggplot2. *Journal of Statistical Software* **87,** 1–17. doi:10.18637/jss.v087.c03 (2018).

272. Reinke, A. W., Baek, J., Ashenberg, O. & Keating, A. E. Networks of bZIP Protein-Protein Interactions Diversified Over a Billion Years of Evolution. *Science* **340,** 730–734. doi:10.1126/science.1233465 (2013).

273. Street, K. *et al.* Slingshot: Cell Lineage and Pseudotime Inference for Single-Cell Transcriptomics. *BMC Genomics* **19,** 477. doi:10.1186/s12864-018-4772-0 (2018).

274. Cheung, S. *et al.* Systemic Coordination of Whole-Body Tissue Remodeling during Local Regeneration in Sea Anemones. *Developmental Cell* **60,** 780–793.e7. doi:10.1016/j.devcel.2024.11.001 (2025).

275. Hariharan, I. K., Wake, D. B. & Wake, M. H. Indeterminate Growth: Could It Represent the Ancestral Condition? *Cold Spring Harb Perspect Biol* **8,** a019174. doi:10.1101/cshperspect.a019174. PMID: 26216720 (2016).

276. David, C. N. Interstitial Stem Cells in Hydra: Multipotency and Decision-Making. *Int. J. Dev. Biol.* **56,** 489–497. doi:10.1387/ijdb.113476cd. PMID: 22689367 (6-7-8 2012).

277. Primack, A. S. *et al.* Differentiation Trajectories of the Hydra Nervous System Reveal Transcriptional Regulators of Neuronal Fate. *bioRxiv,* 2023.03.15.531610. doi:10.1101/2023.03.15.531610. PMID: 36993575 (2023).

278. Brown, D. D. R. & Pearson, B. J. A Brain Unfixed: Unlimited Neurogenesis and Regeneration of the Adult Planarian Nervous System. *Front Neurosci* **11,** 289. doi:10.3389/fnins.2017.00289. PMID: 28588444 (2017).

279. Andreoni-Pham, R. *et al.* Whole Body Regeneration Deploys a Rewired Embryonic Gene Regulatory Network Logic. *Nat Commun.* doi:10.1038/s41467-025-67196-4 (2025).

280. Morris Little, H. *et al.* A Molecular, Spatial and Regulatory Atlas of the Hydra Vulgaris Nervous System. *Development* **152,** dev204983. doi:10.1242/dev.204983 (2025).

281. Noro, Y. *et al.* Regionalized Nervous System in *Hydra* and the Mechanism of Its Development. *Gene Expression Patterns* **31,** 42–59. doi:10.1016/j.gep.2019.01.003 (2019).

282. Holtmann, M. & Thurm, U. Mono- and Oligo-Vesicular Synapses and Their Connectivity in a Cnidarian Sensory Epithelium (Coryne Tubulosa). *Journal of Comparative Neurology* **432,** 537–549. doi:10.1002/cne.1118 (2001).

283. Slautterback, D. B. & Fawcett, D. W. The Development of the Cnidoblasts of Hydra. *The Journal of Cell Biology* **5,** 441–452. doi:10.1083/jcb.5.3.441 (1959).

284. Price, K. L. *et al.* Examination of Germline and Somatic Intercellular Bridges in Hydra Vulgaris Reveals Insights into the Evolutionarily Conserved Mechanism of Intercellular

Bridge Formation. *Mol Biol Evol* **42,** msaf248. doi:10.1093/molbev/msaf248. PMID: 41032677 (2025).

285. Westfall, J. A. The Differentiation of Nematocysts and Associated Structures in the Cnidaria. *Zeitschrift für Zellforschung* **75,** 381–403. doi:10.1007/BF00336871 (1966).

286. Gahan, J. M. *et al.* Functional Studies on the Role of Notch Signaling in Hydractinia Development. *Dev Biol* **428,** 224–231. doi:10.1016/j.ydbio.2017.06.006. PMID: 28601529 (2017).

287. Käsbauer, T. *et al.* The Notch Signaling Pathway in the Cnidarian *Hydra*. *Developmental Biology* **303,** 376–390. doi:10.1016/j.ydbio.2006.11.022 (2007).

288. Lövy, A. *et al.* The Notch Pathway in Metazoa: A Comparative Analysis across Cnidarians and Beyond. *EvoDevo* **16,** 8. doi:10.1186/s13227-025-00246-6 (2025).

289. Berger, C., Renner, S., Lüer, K. & Technau, G. M. The Commonly Used Marker ELAV Is Transiently Expressed in Neuroblasts and Glial Cells in the Drosophila Embryonic CNS. *Developmental Dynamics* **236,** 3562–3568. doi:10.1002/dvdy.21372 (2007).

290. Gavgani, F. M. *et al. Ectopic Head Regeneration after Nervous System Ablation in a Sea Anemone* https://www.biorxiv.org/content/10.1101/2025.02.04.636403v1 (2025). Pre-published.

291. Fincher, C. T., Wurtzel, O., de Hoog, T., Kravarik, K. M. & Reddien, P. W. Cell Type Transcriptome Atlas for the Planarian Schmidtea Mediterranea. *Science* **360,** eaaq1736. doi:10.1126/science.aaq1736 (2018).

292. Zeng, A. *et al.* Prospectively Isolated Tetraspanin+ Neoblasts Are Adult Pluripotent Stem Cells Underlying Planaria Regeneration. *Cell* **173,** 1593–1608.e20. doi:10.1016/j.cell.2018.05.006. PMID: 29906446 (2018).

293. Hulett, R. E. *et al.* Acoel Single-Cell Atlas Reveals Expression Dynamics and Heterogeneity of Adult Pluripotent Stem Cells. *Nat Commun* **14,** 2612. doi:10.1038/s41467-023-38016-4. PMID: 37147314 (2023).

294. Chandra, V. *et al.* Distributed Neural Computation and the Evolution of the First Brains. *bioRxiv,* 2025.10.03.680388. doi:10.1101/2025.10.03.680388. PMID: 41256607 (2025).

295. Álvarez-Campos, P. *et al.* Annelid Adult Cell Type Diversity and Their Pluripotent Cellular Origins. *Nat Commun* **15,** 3194. doi:10.1038/s41467-024-47401-6 (2024).

296. Juliano, C. E., Swartz, S. Z. & Wessel, G. M. A Conserved Germline Multipotency Program. *Development* **137,** 4113–4126. doi:10.1242/dev.047969 (2010).

297. Funayama, N., Nakatsukasa, M., Mohri, K., Masuda, Y. & Agata, K. Piwi Expression in Archeocytes and Choanocytes in Demosponges: Insights into the Stem Cell System in Demosponges. *Evolution & Development* **12,** 275–287. doi:10.1111/j.1525-142X.2010.00413.x (2010).

298. Reddien, P. W., Oviedo, N. J., Jennings, J. R., Jenkin, J. C. & Sánchez Alvarado, A. SMEDWI-2 Is a PIWI-like Protein That Regulates Planarian Stem Cells. *Science* **310,** 1327–1330. doi:10.1126/science.1116110. PMID: 16311336 (2005).

299. Palakodeti, D., Smielewska, M., Lu, Y.-C., Yeo, G. W. & Graveley, B. R. The PIWI Proteins SMEDWI-2 and SMEDWI-3 Are Required for Stem Cell Function and piRNA Expression in Planarians. *RNA* **14,** 1174–1186. doi:10.1261/rna.1085008. PMID: 18456843 (2008).

300. Juliano, C. E. *et al.* PIWI Proteins and PIWI-interacting RNAs Function in Hydra Somatic Stem Cells. *Proceedings of the National Academy of Sciences* **111,** 337–342. doi:10.1073/pnas.1320965111 (2014).

301. Alié, A. *et al.* Somatic Stem Cells Express *Piwi* and *Vasa* Genes in an Adult Ctenophore: Ancient Association of "Germline Genes" with Stemness. *Developmental Biology* **350,** 183–197. doi:10.1016/j.ydbio.2010.10.019 (2011).

302. Shao, P.-J., Chiu, Y.-L., Tsai, P.-H. & Shikina, S. Possible Germline Progenitor Cells in Extra-Gonadal Tissues of the Sea Anemone, Exaiptasia Diaphana. *Front. Mar. Sci.* **10.** doi:10.3389/fmars.2023.1278022 (2023).

303. Praher, D. *et al.* Characterization of the piRNA Pathway during Development of the Sea Anemone Nematostella Vectensis. *RNA Biol* **14,** 1727–1741. doi:10.1080/15476286.2017.1349048. PMID: 28783426 (2017).

304. Modepalli, V., Fridrich, A., Agron, M. & Moran, Y. The Methyltransferase HEN1 Is Required in Nematostella Vectensis for microRNA and piRNA Stability as Well as Larval Metamorphosis. *PLoS Genet* **14,** e1007590. doi:10.1371/journal.pgen.1007590. PMID: 30118479 (2018).

305. Morgan, T. H. Growth and Regeneration in Planaria Lugubris. *Archiv für Entwicklungsmechanik der Organismen* **13,** 179–212. doi:10.1007/BF02161982 (1901).

306. Baguñà, J., Saló, E. & Auladell, C. Regeneration and Pattern Formation in Planarians: III. Evidence That Neoblasts Are Totipotent Stem Cells and the Source of Blastema Cells. *Development* **107,** 77–86. doi:10.1242/dev.107.1.77 (1989).

307. Sato, K. *et al.* Identification and Origin of the Germline Stem Cells as Revealed by the Expression of Nanos-Related Gene in Planarians. *Development, Growth & Differentiation* **48,** 615–628. doi:10.1111/j.1440-169X.2006.00897.x (2006).

308. Wang, Y., Zayas, R. M., Guo, T. & Newmark, P. A. Nanos Function Is Essential for Development and Regeneration of Planarian Germ Cells. *Proc Natl Acad Sci U S A* **104,** 5901–5906. doi:10.1073/pnas.0609708104. PMID: 17376870 (2007).

309. Nishimiya-Fujisawa, C. & Kobayashi, S. in *Reproductive and Developmental Strategies: The Continuity of Life* 123–155 (Springer Japan, Tokyo, 2018). doi:10.1007/978-4-431-56609-0_7.

310. Nishimiya-Fujisawa, C. *et al. An Ancient Split of Germline and Somatic Stem Cell Lineages in Hydra* https://www.biorxiv.org/content/10.1101/2023.07.04.546637v1 (2025). Pre-published.

311. Solana, J. Closing the Circle of Germline and Stem Cells: The Primordial Stem Cell Hypothesis. *EvoDevo* **4,** 2. doi:10.1186/2041-9139-4-2. PMID: 23294912 (2013).

312. Kimura, J. O., Bolaños-Rodriguez, M., Ricci, L. & Srivastava, M. Embryonic Origins of Adult Pluripotent Stem Cells. *Cell* **185,** 4756–4769.e13. doi:10.1016/j.cell.2022.11.008. PMID: 36493754 (2022).

313. Leclère, L. *et al.* Maternally Localized Germ Plasm mRNAs and Germ Cell/Stem Cell Formation in the Cnidarian *Clytia. Developmental Biology* **364,** 236–248. doi:10.1016/j.ydbio.2012.01.018 (2012).

314. Ruggiero, A. *et al. Dual Origins for Neural Cells during Development of the Clytia Planula Larva* https://www.biorxiv.org/content/10.1101/2025.11.17.688882v1 (2026). Pre-published.

315. Elek, A. *et al.* Decoding Cnidarian Cell Type Gene Regulation. *Nat Ecol Evol* **10,** 140–153. doi:10.1038/s41559-025-02906-1 (2026).

316. Kraus, J. E. M., Busengdal, H., Kraus, Y., Hausen, H. & Rentzsch, F. Doublecortin-like Kinase Is Required for Cnidocyte Development in Nematostella Vectensis. *Neural Dev* **19,** 11. doi:10.1186/s13064-024-00188-0. PMID: 38909268 (2024).

317. Stone, M. C., Kothe, G. O., Jegla, T. J. & Rolls, M. M. Neurite Branching Is Associated with Mixed Microtubule Polarity in Sea Anemone Neurons. *J Exp Biol* **228,** jeb.251169. doi:10.1242/jeb.251169. PMID: 40931983 (2025).

318. Westfall, J. A., Yamataka, S. & Enos, P. D. ULTRASTRUCTURAL EVIDENCE OF POLARIZED SYNAPSES IN THE NERVE NET OF HYDRA. *J Cell Biol* **51,** 318–323. doi:10.1083/jcb.51.1.318 (1971).

319. Westfall, J. A., Elliott, S. R., MohanKumar, P. & Carlin, R. W. Immunocytochemical Evidence for Biogenic Amines and Immunogold Labeling of Serotonergic Synapses in Tentacles of Aiptasia Pallida (Cnidaria, Anthozoa). *Invertebrate Biology* **119,** 370–378. doi:10.1111/j.1744-7410.2000.tb00105.x (2000).

320. Chari, T. *et al.* Whole-Animal Multiplexed Single-Cell RNA-seq Reveals Transcriptional Shifts across Clytia Medusa Cell Types. *Science Advances* **7,** eabh1683. doi:10.1126/sciadv.abh1683 (2021).

321. Denker, E., Bapteste, E., Le Guyader, H., Manuel, M. & Rabet, N. Horizontal Gene Transfer and the Evolution of Cnidarian Stinging Cells. *Current Biology* **18,** R858–R859. doi:10.1016/j.cub.2008.07.031 (2008).

322. Hager, G. & David, C. N. Pattern of Differentiated Nerve Cells in Hydra Is Determined by Precursor Migration. *Development* **124,** 569–576. doi:10.1242/dev.124.2.569 (1997).

323. Schüler, A. *et al.* The Rise and Fall of TRP-N, an Ancient Family of Mechanogated Ion Channels, in Metazoa. *Genome Biology and Evolution* **7,** 1713–1727. doi:10.1093/gbe/evv091 (2015).

324. Li, X. *et al.* Major Diversification of Voltage-Gated K+ Channels Occurred in Ancestral Parahoxozoans. *Proc Natl Acad Sci U S A* **112,** E1010–E1019. doi:10.1073/pnas.1422941112. PMID: 25691740 (2015).

325. Tarashansky, A. J., Xue, Y., Li, P., Quake, S. R. & Wang, B. Self-Assembling Manifolds in Single-Cell RNA Sequencing Data. *eLife* **8,** e48994. doi:10.7554/eLife.48994. PMID: 31524596.

326. Forero-Mejia, A. C., Duchatelet, L. & Östman, C. Cnida Morphology as Taxonomic Tools within Tube-Dwelling Anemones (Ceriantharia, Cnidaria). *Oceans* **5,** 491–521. doi:10.3390/oceans5030029 (2024).

327. Babonis, L. S., Ryan, J. F., Enjolras, C. & Martindale, M. Q. Genomic Analysis of the Tryptome Reveals Molecular Mechanisms of Gland Cell Evolution. *EvoDevo* **10,** 23. doi:10.1186/s13227-019-0138-1. PMID: 31583070 (2019).