Comparison of cell types across life cycle stages of the hydrozoan Clytia hemisphaerica

Anna Ferraioli

Sorbonne Université

Ecole Doctorale Complexité du Vivant (ED515)

Laboratoire de Biologie du Développement de Villefranche-sur-mer (UMR7009) CNRS/Sorbonne

Genome and protein evolution in animals
Mécanismes développementaux des Cnidaires

Comparison of cell types across life cycle stages of the hydrozoan *Clytia hemisphaerica*

Par Anna Ferraioli

Thése de doctorat de Biologie du Développement

Dirigée par Richard Copley
Co-dirigée par Evelyn Houliston

La Thése sera présentée et soutenue publiquement le 19/09/2022

Devant une jury composé de:

Dr. Eric Röttinger	Président du jury et Rapporteur
Dr. Maria Ina Arnone	Rapporteur
Dr. Stéphanie Bertrand	Examinateur
Dr. Mathilde Paris	Examinateur
Dr. Richard Copley	Directeur de thése
Dr. Evelyn Houliston	Co-directeur de thése

Ad Anna Sole,
che non vedo l'ora di incontrare.

La natura è un miracolo costante cha ha un unico difetto,
quello di dare uno spettacolo gratuito, ma cosi gratuito
che ormai non ci facciamo più caso.

Luciano De Crescenzo – Il Dubbio, 1992

ACKNOWLEDGEMENTS

I begin with expressing all my gratitude to my PhD supervisors Evelyn Houliston and Richard Copley for giving me the opportunity of doing a PhD and for the precious contributions to my professional and personal growth. Thanks to them, I have learnt invaluable lessons.

I thank all the PIs and PhD students of the EVOCell network. Being part of this network has been an enriching and exciting experience during which I had the privilege of meeting wonderful people and learning so much.

I would like to thank the members of the Clytia team, Tsuyoshi, Carine, Lucas, Sandra, Julia, Manon, Cat, Bastien, Camille, Julie, Francois and Nicholas for their continuous support in the daily lab life and for their contribution to the development of my PhD project.
I would like to thank the past and present members of the genome team, Coralie and Sami for all their help. Special thanks to Philippe who, since I arrived in Villefranche, never hesitated to find the time to fix codes, programs, computers, cables, chairs, light bulbs, my French, sometimes even before I asked.

Many thanks to Yas for being the kindest tutor. Thanks to all the people of the LBDV that offered me help, support, candies and cookies and contributed to ease my PhD journey.

A special thanks goes to Julia, Manon, Anne and Angelica for creating a supportive and joyful environment which has been essential in getting through this journey. Thanks to the past and present members of the LBDV students community Vitoria, Marie, Marta, Alex, Sophie, Marianne who contributed in many ways to making my PhD unforgettable. Gracias to David who keeps supporting me from the other side of the world.

Finally, I want to thank my family: grazie a mamma e papà ai quali non è ancora ben chiaro cosa faccio nella vita, ma che, nonostante tutto, non smettono di incoraggiarmi e supportarmi ogni giorno. Thanks to my brother who is always there for me even from far away. Thanks to Andrea who, patiently, is always on my side. Thanks to them, I get reminded of the significance of love.

TABLE OF CONTENTS

RÉSUMÉ

Comparaison des types cellulaires entre les stades de vie de l'hydrozoaire *Clytia hemisphaerica*

L'hydrozoaire *Clytia hemisphaerica* présente un cycle de vie triphasique comprenant une colonie de polypes à propagation végétative et la méduse, qui est le stade se reproduisant de façon sexuée. Les méduses mâles et femelles libèrent leurs gamètes quotidiennement et environ un jour après la fécondation, une larve planula ciliée se forme. Après trois jours, la planula se fixe sur un substrat et se métamorphose pour donner naissance à un polype fondateur de la colonie, destiné à l'alimentation, le gastrozoïde. La colonie se propage par extension du stolon et un deuxième type de polype, le gonozoïde, libère des méduses par bourgeonnement. Les analyses précédentes du génome et des transcriptomes des trois principaux stades de vie de *Clytia* ont révélé des programmes d'expression génique spécifiques à chaque étape (Leclère et al. 2019, Nature Ecology & Evolution). J'ai étendu cette comparaison au niveau des types cellulaire en utilisant la technologie du sc-RNA-seq chez la méduse et la larve de *Clytia*. Avec des collègues du LBDV, en collaboration avec les laboratoires de L. Pachter et de D. Anderson à Caltech nous avons généré un premier "atlas" des types cellulaires de la méduse de *Clytia* (Chari et al. 2021, Science Advances). J'ai élargi cet atlas en intégrant des données supplémentaires. L'analyse de l'atlas a révélé huit grands types cellulaires, dont l'épiderme et le gastroderme, les cellules bioluminescentes, les ovocytes et les cellules souches multipotentes des hydrozoaires (i-cells) et leurs dérivés telles que les cellules neurales, les nématocytes et les cellules glandulaires. L'analyse par hybridation in situ des profils d'expression a révélé des sous-types non caractérisés auparavant, dont 14 sous-populations neuronales. L'analyse de la trajectoire de la lignée des nématocystes a révélé deux programmes transcriptionnels distincts, un programme "nématoblaste", caractérisé par la production de la nématocyste, et la phase de différenciation du nématocyte, caractérisée par la production du nématocil.

Pour obtenir les données sc-RNAseq pour la planula j'ai optimisé les protocoles de dissociation, de fixation et de tri des cellules. J'ai obtenu un jeu de données pour la planula de 4370 cellules, regroupées en 19 clusters cellulaires. Après l'analyse des profils d'expression par hybridation in situ de gènes à trois stades de développement de la planula, j'ai pu attribuer des identités cellulaires et regrouper les 19 clusters en 8 grandes classes cellulaires. Celles-ci correspondent à l'épiderme, le gastroderme, les i-cells, les nématocytes, les cellules neurales, les cellules neurosécrétrices aborales et les cellules muqueuses et les cellules excrétrices putatives (PEC). Cet inventaire des types cellulaires de la planula de *Clytia* représente le premier atlas cellulaire d'une larve d'hydrozoaire et fournit la caractérisation de populations cellulaires non décrites auparavant ainsi que des informations supplémentaires sur les types cellulaires déjà connus.

Enfin, j'ai exploité les données sc-RNAseq de la planula et de la méduse pour explorer les signatures transcriptionnelles partagées par type cellulaire. La méduse est la forme la plus complexe des cnidaires. Mon but était de tester si l'augmentation du répertoire moléculaire reflète une expansion des types de cellules chez l'adulte. J'ai comparé les signatures

moléculaires des types de cellules de méduse avec celles de planula et cherché lesquelles étaient partagées ou uniques. Cette étude comparative est préliminaire, mais elle permet de tirer quelques interprétations biologiques. Les transcriptomes de certaines classes de cellules telles que le gastroderme, les i-cells, les cellules neurales et les nématocytes se sont révélés clairement similaires entre méduse et planula. En revanche, d'autres classes ont montré une probabilité plus faible d'être partagées entre les stades, comme l'épiderme, les cellules muqueuses et les cellules excrétrices putatives (PEC) de la planula et l'épiderme et les cellules glandulaires digestives de la méduse. Cette approche comparative fournit des résultats prometteurs, permettant une première compréhension des similitudes entre les étapes à un niveau plus large de classes de cellules. Cependant, les méthodes nécessitent de l'optimisation pour permettre une comparaison beaucoup plus fine au niveau du type de cellule. La classe des nématocytes est le seul exemple où il a été possible d'établir une similarité entre les types cellulaires. Dans l'ensemble, les résultats de cette exploration initiale des programmes de régulation, ainsi que les informations que j'ai pu recueillir à partir des analyses comparatives, ont permis de dessiner des scénarios évolutifs sur l'origine et le développement des types de cellules chez *Clytia*.

ABSTRACT

Cell type comparison across life stages of the hydrozoan *Clytia hemisphaerica*

The hydrozoan *Clytia hemisphaerica* displays a typical tri-phasic hydrozoan life cycle including a vegetatively propagating polyp colony and free-swimming medusa form as the sexually reproductive life stage. Male and female jellyfish spawn daily and after fertilisation a ciliated planula larva forms in about one day. After three days the planula settles and metamorphoses to give rise to a primary feeding polyp, the gastrozooid, founder of the polyp colony. The colony propagates by stolon extension and a second type of polyp, the gonozooid, releases medusa by budding. Previous analysis of the genome and the bulk transcriptome across the three life stages revealed specific gene expression programs for each stage (Leclère et al. 2019, Nature Ecology & Evolution). In this work I have extended this comparison to the level of individual cell types via single-cell RNA transcriptomics of *Clytia* medusa and larva.

Together with colleagues from LBDV and members of the Pachter' and Anderson's lab at Caltech, we generated the first female medusa cell type atlas (Chari et al. 2021, Science Advances). I extended this atlas by generating and integrating further sc-RNAseq data. Analysis of the medusa cell type atlas revealed eight broad cell type classes including epidermis and gastrodermis, bioluminescent cells, oocytes and the hydrozoan multipotent stem cells (i-cells) and their derivatives such as neurons, nematocytes and gland cells. In situ hybridisation analysis of expression patterns revealed previously uncharacterized subtypes including 14 neuronal subpopulations. Trajectory analysis of the nematocyte lineage revealed two distinct transcriptional programs within this cell class, a "nematoblast" phase, characterised by the production of the typical nematocyte capsule, and the nematocyte differentiation phase, characterised by the production of the nematocil apparatus.

Sc-RNAseq for the *Clytia* planula required refinement of cell dissociation, fixation and sorting protocols (collaboration with Arnau Sebé-Pedros' group, CRG, Barcelona). The planula cell atlas consists of 4370 cells grouped in 19 cell clusters. Following in situ hybridisation expression patterns analysis of known and novel genes at three planula developmental stages I could assign cell identities and combine the 19 clusters in 8 broad cell classes. These correspond to the two cnidarian epithelial tissue layers, the epidermis and the gastrodermis, the i-cells, the nematocytes, neural cells, aboral neurosecretory cells and distinct populations of secretory cells, mucous cells and putative excretory cells (PEC). This *Clytia* planula cell atlas is the first cell atlas of an hydrozoan larva and provides characterization of previously undescribed cell populations as well as further information on already known cell types.

Finally, I exploited our cell type atlases of the planula and medusa to explore shared transcriptional signatures at the cell type level. The medusa is the most complex form of cnidarians. The aim of this comparative analysis is to clarify whether the increase in molecular repertoire reflects an expansion of cell types in the adult. I compared the molecular signatures of jellyfish cell types with those of the planula and asked which were shared or unique. This comparative study is still in its early stages, however it has allowed some preliminary biological interpretations. The transcriptomes of certain cell classes such as gastrodermis, i-cells, neural cells and nematocytes showed clear similarity. In contrast, some other classes showed a lower

probability of being shared between stages, such as the epidermis, mucous cells and PECs of the planula and the epidermis and digestive glandular cells of the jellyfish. This comparative approach provides promising results, allowing a first understanding of the similarities between the steps at a broader level of cell classes. However, the methods require further optimisation to allow comparison at the cell type level. Finer level of comparison was revealed so far only within the nematocyte class with stages of nematogenesis showing a high degree of similarity in gene usage between medusa and planula. Overall, these results allowed evolutionary scenarios on the origin and development of cell types in *Clytia* to be drawn.

COLLABORATIONS

My PhD project was undertaken as part of EvoCELL, a Marie Skłodowska-Curie Innovative Training Network (MSCA-ITN) aiming at studying the evolution of cell types and tissues in vertebrates and invertebrates (https://evocell-itn.eu/). In this context I received training on experimental work and computational analysis. Most importantly, I had the opportunity to discuss with the PIs and the other students of the network concerning the best strategies to overcome experimental and computational issues.

My work contributed to the first medusa cell atlas, published in collaboration with the Pachter and Anderson groups at Caltech (Chari, Weissbourd, Gehring, Ferraioli et al., 2021). The sc-RNAseq data for this first atlas was generated at Caltech from adult control and starved *Clytia* medusae. I contributed to the regular discussions on data processing, presentation and interpretation held between the LBDV and Caltech teams. My experimental role in the collaboration was to validate the identity of cell types extracted from single cell data by selecting marker gene candidates and carrying out *in situ* hybridization. In addition, I validated the effect of starvation on adult medusae in the gonads by confocal microscopy and made the corresponding figure for the paper. Together with Lucas Leclère, I also generated all the figures for the paper that include expression pattern images.

Encapsulation experiments that led to the generation of planula single cell transcriptomics data were performed by Marta Iglesias-Garcia (Sebé-Pedros lab) at the CRG genomics facility (Barcelona) by using frozen samples of planula dissociated cells that I prepared in Villefranche and shipped to the platform. Additionally, in collaboration with the Sebé-Pedros group, I tested encapsulation of freshly dissociated cells using the InDrop platform and prepared the cDNA libraries with the help of the Sebé-Pedros group members.

Fixation of specimens, embedding and sectioning of *Clytia* planula and medusa for TEM were performed by Sophie Pagnotta at the Plateforme Commune de Microscopie Électronique, Université Côte d'Azur. I provided the living material and contributed to the imaging and interpretation.

PREFACE

My thesis manuscript is structured in four chapters.

- Chapter 1 includes a general introduction which begins by introducing the concept of the evolution of the cell types. The overview of the single cell transcriptomics methodology follows in Section 1.2. In Section 1.3 I introduce the diversity of cnidarian life cycles and briefly describe the contribution of the two main cnidarian models, *Hydra* and *Nematostella,* to the understanding of cnidarian cell types. A description of *Clytia hemisphaerica* as a model organism used in this study follows in Section 1.4. Finally, the overview of the objectives of this study are described in Section 1.5.

The following chapters include a general overview at the beginning and are organised in four sections each, such as a brief introduction, the results, the methods and finally a discussion.

- In Chapter 2 I present the work I did to generate and validate the cell type atlas of the *Clytia* medusa. This chapter begins with a general introduction to the first medusa cell atlas paper in section 2.1 that included this work in the framework of a collaboration with the Pachter and Anderson groups at Caltech (Chari, Weissbourd, Gehring, Ferraioli et al., 2021). The following sections cover additional results and analysis I did to extend the medusa atlas after the publication. These include an enlarged medusa dataset I compiled and the ultrastructural characterisation of the medusa cell types through electronic microscopy imaging and analysis. This chapter contains the publication concerning the first medusa atlas as an annex.

- In Chapter 3 I introduce the generation and validation of the cell atlas of the planula of *Clytia*. After a short introduction in section 3.1 I present the results in section 3.2 which is subdivided in several subsections where results concerning each cell class are presented. This chapter includes an additional section on troubleshooting the planula dissociation protocol (section 3.4).

- Chapter 4 includes the comparative analysis at the cell type level I performed by exploiting the information included in the available atlas. Following a short introduction which contextualises my comparative study, the results section begins with the outline of our approach in section 4.1. The following sections include considerations we could draw by interpreting the results. Alternative results we obtained with our approach are presented as annex to the chapter.

Finally I summarise my findings and outline evolutionary considerations I could draw from those in the General Discussion. Conclusions and Future perspectives follow in the end.

CHAPTER 1 -

Introduction

1.1 - Cell types as evolutionary units

One important strategy to reconstruct the evolutionary history of structures or traits is to determine homology across species and infer characteristics of the common ancestor. The concept of homology linked to common evolutionary origin was first introduced by Charles Darwin in his historical work *"On the Origin of Species"* :

" We have seen that the members of the same class, independently of their habits of life, resemble each other in the general plan of their organisation. This resemblance is often expressed by the term "unity of type"; or by saying that the several parts and organs in the different species of the class are homologous. [...] If we suppose that the ancient progenitor, the archetype as it may be called, of all mammals, had its limbs constructed on the existing general pattern, for whatever purpose they served, we can at once perceive the plain signification of the homologous construction of the limbs throughout the whole class. "

On the Origin of Species - Charles Darwin (1859)

A way to establish homology is by measuring how similar two structures are. In the past, homology has been determined mainly by morphological observation which has represented a valuable tool for comparing structures. However morphology-based analyses might not be sufficient to discriminate between different cells that co-exist in the same tissue (Hwang et al., 2018).
At present, scientists can exploit a variety of molecular tools to investigate homology (Arendt, 2003, 2005). Indeed, the combination of morphological and molecular approach can facilitate the establishment of homology between cells (Arendt, 2003).

The cell is the biological unit of tissues and organs. Cells with the same functions or similar morphology in a given tissue or organ can be seen as a group forming a "type". A confluence of technology development in microfluidics, sequencing and imaging has led to the delineation of cell types becoming a major focus of the modern era of biology.
A cell type can be defined by the expression of a unique combination of effector genes and transcription factors which determine its phenotype and function (Arendt, 2003, 2005, 2008). This unique combination of genes has been described as the "molecular fingerprint" of a cell type (Arendt, 2005). On this basis, cell type-specific fingerprints can be compared on a molecular level to decipher homologies within and across species, allowing the reconstruction of their evolutionary history. Indeed, considering the relationship between homology and common ancestry, determining homology between two cell types across species implies the existence of a similar cell type in the common ancestor (Arendt, 2008). This requires a clear determination of the identity of a given cell type which can be defined by the identification of the specific regulatory programs that affect gene expression or "*core regulatory complex (CoRC)*" (Arendt et al., 2016).

The core regulatory complex is defined as

" *A protein complex composed of terminal selector transcription factors that enables and maintains the distinct gene expression programme of a cell.* "

Arendt et al. (2016)

Where *"terminal selectors"* are transcription factors that regulate cell type specific gene expression in differentiated cells and control their identity by repressing the emergence of alternative types (Arendt et al., 2016). When, in an organism, the regulatory programs that define the identity of a given cell type undergo modifications during the evolution in a partially independent manner, a cell type can be considered as an *"evolutionary unit" (Arendt et al., 2016)*.

This proposed definition of cell type opens the way to establish comparative evolutionary studies under a new light. An important challenge has been to find experimental approaches for the collection of molecular data at the cell type level. In the next section I will introduce the principle of the methodology of single cell transcriptomics which, over the last few years, has been developing rapidly to face this challenge.

1.2 - Single cell transcriptomics

In the modern genomic era, many approaches to study gene expression have been developed. Sequencing of genomes and bulk transcriptomes have made it possible to address questions in many fields of biology. However, bulk gene expression approaches do not provide the required resolution to investigate the differences between cells in a given tissue or animals (Hwang et al., 2018; Marioni & Arendt, 2017).

During the past decade the single cell based technological approaches have been developed and are nowadays, widespread techniques in continuous and rapid advancement. Among many applications including mapping of epigenetic landscapes and sequencing of genomes of single cells, the most popular remains single-cell RNA sequencing (sc-RNAseq) (Marioni & Arendt, 2017). A variety of methods have been developed to enable the processing of isolated cells to allow cDNA libraries to be produced in which the sequences are barcoded to trace them back to individual cells. Single cells can be sorted into multiwell plates (Deng et al., 2014; Jaitin et al., 2014) or encapsulated in aqueous droplets within an oil-based phase by using droplet based microfluidic devices (Klein et al., 2015; Macosko et al., 2015; https://www.10xgenomics.com/). Plate-based methods allow information concerning the size of the cells and the surface proteins to be obtained (Jaitin et al., 2014), whereas droplet-based systems do not. On the other hand, droplet-based systems ensure the encapsulation of tens of thousands of cells and are more convenient than the plate-based methods, which can profile only hundreds of cells in a single experiment (Macosko et al., 2015; Marioni & Arendt, 2017).

When applied to tissue or whole animals, one of the challenges of single cell transcriptomics is sample preparation (Shapiro et al., 2013). It is critical to establish a good dissociation protocol which allows maximum recovery of cell types and a low percentage of cell death. Sample preparation is even more challenging when downstream processing of dissociated cells is applied to marine animal cells which, in some cases, can undergo osmotic or ionic shock when transferred into solutions with compositions that vary significantly from sea water.

The most widely used method for Sc-RNAseq is currently the commercial Chromium system by 10X Genomics which enables high-throughput profiling of 3' ends of RNAs from single cells along with elevated encapsulation efficiency (Hwang et al., 2018; https://www.10xgenomics.com/). The Chromium system allows the encapsulation of isolated cells and gel beads containing barcoded oligonucleotides into aqueous droplets on a microfluidic device. The formation of the droplets is ensured by a continuous flow of an oil phase. The droplets are termed Gel Beads in Emulsion, or GEMs. Each functional GEM contains a single cell, a single gel bead and reverse transcriptase (RT) reagents. Within each GEM single cells are lysed, the gel bead is dissolved and the identically barcoded oligonucleotides bind the 3' end of the transcripts released by the cell. A cDNA library is generated through reverse transcription of the polyadenylated mRNA molecules bound to the barcoded oligonucleotides. As a result, the cDNA library from a single cell will have the same barcode, allowing the sequencing reads to be mapped back to the cell of origin. A second type of randomised barcode, termed Unique Molecular Identifier (UMI), labels each of the RNA molecules (Fig. 1.1, https://www.10xgenomics.com/). Paired-end sequencing with a single index follows the library preparation. Sequencing reads are associated in pairs, with one read containing barcode and UMI sequences and the second containing the transcript. Sequencing reads are then mapped against a suitable genome and used to generate a matrix of counts vs cells, which is exploited for downstream cluster analysis.

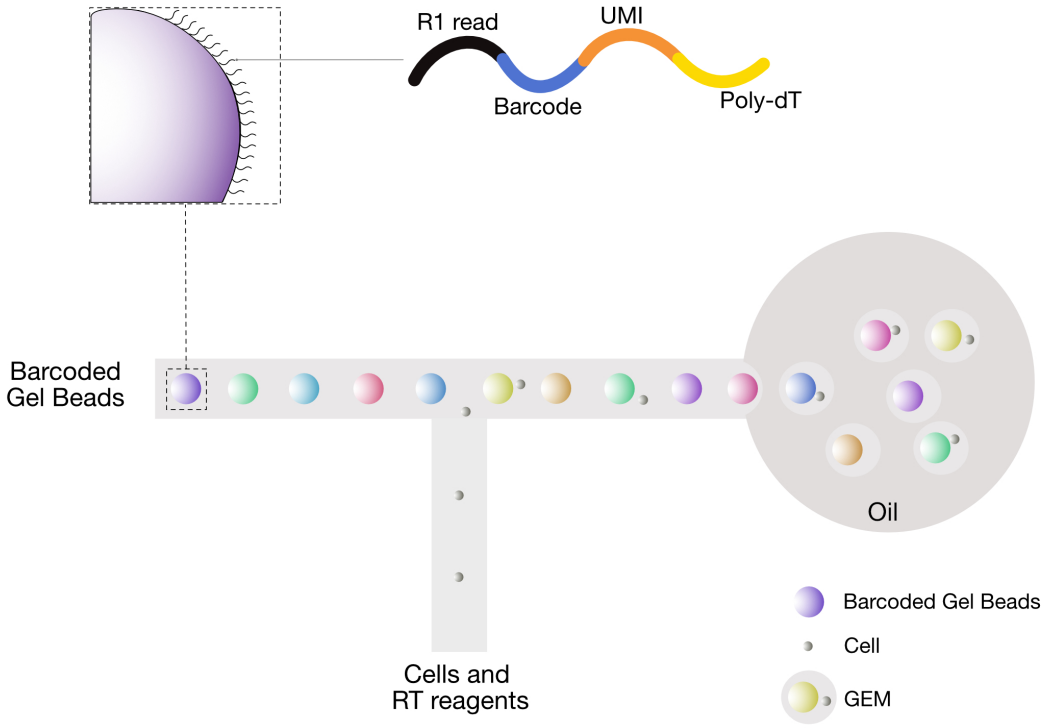


Figure 1.1 - Cell Encapsulation in a microfluidics system
Diagram based on the 10X Genomics Chromium system (https://www.10xgenomics.com/).

1.3 - The phylum Cnidaria and life cycle diversity

The phylum Cnidaria is a group predominantly constituted by marine invertebrates composed of about 11.000 species (Appeltans et al., 2012; Collins, 2009; Zapata et al., 2015). It includes corals, jellyfish, sea anemones, hydroids and the freshwater polyp *Hydra (Zapata et al., 2015)*. The most obvious common trait of the phylum is a particular intracellular structure, the cnidocyst, also known as cnida or nematocyst (Collins, 2009; Zapata et al., 2015). This specialised organelle contains venom and a tightly coiled dart-like structure, and is enclosed in a particular cell type, the stinging cells or cnidocytes (nematocytes).

Cnidarians are a diverse group of animals and raise interest among scientists for many reasons (Zapata et al., 2015). Relatively recent phylogenetic analyses placed Cnidaria as the sister group of Bilateria after a divergence that occurred around 550 million years ago (Fig. 1.2; Dunn et al., 2014). Given this phylogenetic position, knowledge from cnidarians can be very informative regarding metazoan evolution as they share many features with bilaterians such as tissue layers, muscle fibres and sensory organs (Boero et al., 2005; Technau & Steele, 2012).

The large diversity of cnidarian species reflects their variety of life cycles (Collins, 2002). The most representative is a pelago-benthic life cycle with a ciliated planula larva, a benthic polyp and a free swimming medusa as sexually reproductive form (Nielsen, 1998; Technau & Steele, 2012).

Cnidarian species are classified into two evolutionary branches with distinct life cycle characteristics: Anthozoa and Medusozoa (Fig.1.2). These are strongly supported as sister groups on the basis of phylogenetic analysis of molecular and morphological data (Bridge et al., 1995; Collins, 2002; Kayal et al., 2018). Anthozoa include Octocorallia and Hexacorallia, while Medusozoa are constituted by four classes, Hydrozoa, Scyphozoa, Staurozoa and Cubozoa (Fig.1.2).

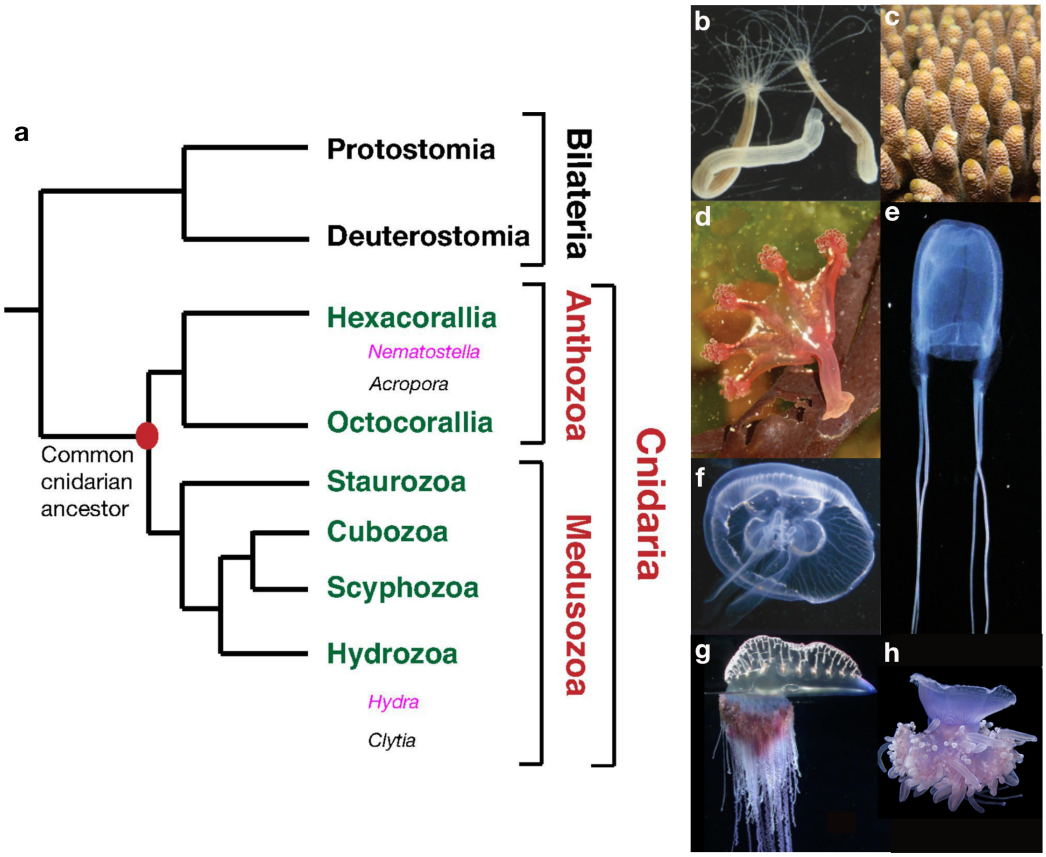


Figure 1.2 - Phylogeny of Cnidaria

(a) Phylogenetic tree showing relationships within the main cnidarian groups and with Bilateria. The name of some of the species with a sequenced genome in pink and black. Taken from Technau & Steele (2012). (b) *Nematostella vectensis* (Anthozoa) adult polyps. Taken from Miller & Ball (2008). (c) *Acropora millepora* (Anthozoa) adult colony. Taken from Technau & Steele (2012). (d) *Calvadosia cruxmelitensis* (Staurozoa). Taken from Ohdera et al. (2019). (e) *Morbakka virulenta* (Cubozoa) adult medusa. Taken from Khalturin et al. (2019). (f) *Aurelia aurita* (Scyphozoa) adult medusa. Taken from Miller & Ball, (2008). (g) *Physalia physalis* (Hydrozoa, Siphonophora). (h) *Cephea cephea* (Scyphozoa). (g) and (h) are taken from Zapata et al. (2015).

Anthozoan life cycles are generally constituted by a planula larva which settles and metamorphoses into an adult polyp (Fig. 1.3; Collins, 2002). On the other hand, Medusozoa are mainly constituted by species forming a free-swimming medusa as well as a polyp stage (Fig. 1.3; Technau & Steele, 2012). The alternation of the polyp stage which is characterised by asexual reproduction and a sexually reproductive medusa stage represents one of the main features of Medusozoa (Brusca & Brusca, 1996). The largest variety of life cycles is observed within Hydrozoa, with many hydrozoan species displaying planula larvae, sessile polyps and medusae, which are released by budding from the polyps (Fig. 1.3; Collins, 2002). Examples of exceptions are siphonophores, which form polymorphic pelagic colonies composed of several polyps and a reduced medusa (Fig. 1.2; Bouillon & Boero, 2000) and the freshwater polyp *Hydra*, which lacks a medusa stage (Fig. 1.3; Bouillon & Boero, 2000; Technau & Steele, 2012). Scyphozoan life cycles generally include a planula larva, a sessile polyp form and a medusa. To produce the juvenile medusae, or ephyrae, the polyps go through a process termed strobilation, which consists in the serial production of juvenile medusa by transverse fission at the oral end of the polyps (Fig. 1.3; Collins, 2002). Cubozoan life cycles generally follow the same alternation between planula, polyp and medusa, however each polyp metamorphoses completely into a single medusa (Fig. 1.3; Collins, 2002). Finally, the enigmatic class of Staurozoa does not produce a free-swimming medusa. Indeed, stauromedusae live on the substrate attached to a stalk. The larvae are non ciliated and crawl on the substrate before transforming into polyps (Fig. 1.3; Collins, 2002).

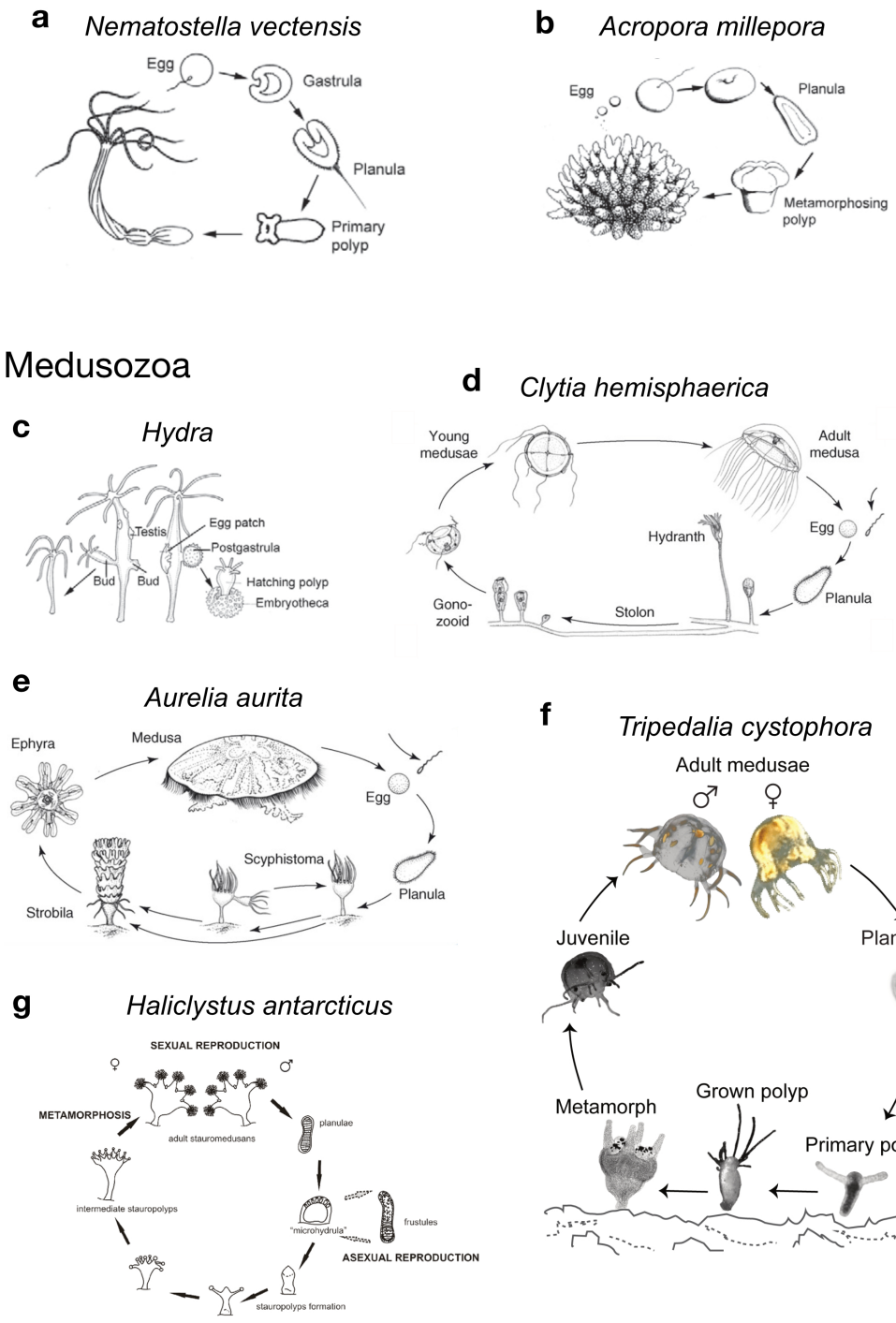


Figure 1.3 - Cnidarian life cycles

(a) Life cycle of *Nematostella vectensis* (Anthozoa). (b) Life cycle of *Acropora millepora* (Anthozoa). (c) Life cycle of *Hydra* (Hydrozoa). (a), (b) and (c) are taken from Technau & Steele (2012). (d) Life cycle of *Clytia hemisphaerica* (Hydrozoa). (e) Life cycle of *Aurelia aurita* (Scyphozoa). (d) and (e) are taken from Miller & Ball (2008). (f) Life cycle of *Tripedalia cystophora* (Cubozoa). Taken from Gurska & Garm (2014). (g) Hypothetical life cycle of *Haliclystus antarcticus (Staurozoa).* Taken from Miranda et al (2010).

The most studied laboratory cnidarian species are the freshwater hydrozoan polyp-only *Hydra (H. vulgaris* and *H. magnipapillata;* Fig. 1.3*)* and the brackish water sea anemone *Nematostella vectensis* (Anthozoa; Fig. 1.2 and 1.3). More recently, a number of other cnidarian experimental models such as *Aurelia aurita* (Fig. 1.2; Khalturin et al., 2019), *Acropora millepora* (Fig. 1.2; Fig. 1.3; Ball et al., 2002)*, Hydractinia echinata* (Frank et al., 2020) and *Clytia hemisphaerica* (Fig. 1.3; Houliston et al., 2010, 2022; Leclère et al., 2019) have been developed. In the following sections I briefly discuss *Hydra* and *Nematostella* model systems*,* highlighting their contribution to the understanding of the cnidarian cell types. Finally I will introduce *Clytia hemisphaerica* as the model used in this study.

1.3.1 - *Hydra* and the interstitial stem cells lineage

Hydra was the first cnidarian model system used for developmental biology, in particular for exploring regenerative capabilities (Technau & Steele, 2012; Trembley et al., 1744). The life cycle of *Hydra* is mainly constituted by a polyp stage which is capable of asexual reproduction and forms new polyps through budding. When sexual reproduction occurs, the embryo develops in a cuticule from which a fully formed polyp hatches after a period of dormancy that can last between 2 and 24 weeks (Martin et al., 1997). *Hydra* is a versatile model system and has proven suitable to address several developmental biology problems, for instance, neurogenesis (Mcconnell, 1932), the phenomenon of polyp induction (Browne, 1909) as well as regeneration (H. R. Bode, 2003), and stemness (David & Murphy, 1977).

In line with the typical cnidarian organisation, *Hydra* polyps are constituted by two tissue layers, the ectoderm and the endoderm, separated by a layer of acellular mesoglea (Fig. 1.4a; Bode et al., 1973) The epithelial cells of the two layers are constantly in a proliferative state (Campbell, 1967) and are considered epithelial stem cells (Bode, 1996). Cell division occurs along the body column. As the cells divide they get displaced along the axis of the polyp, towards the apical region and the foot where they get expelled or become part of a budding polyp (Bode, 1996). All the other cell types are enclosed in the interspace between the two layers and constitute the interstitial cell lineage (Fig. 1.4a; Bode et al., 1973). The interstitial cells (i-cells) are a multipotent stem cell population (Bosch & David, 1987) that can give rise to germline cells (Nishimiya-Fujisawa & Kobayashi, 2012) and somatic derivatives such as neural cells (Davis et al., 1968), nematocytes (David & Challoner, 1974) and secretory cells (Bode et al., 1987; Fig. 1.4b).
A recent study provided detailed transcriptional signatures of the cell types of *Hydra* exploiting single cell transcriptomics (Siebert et al., 2019). Analysis of the gene signatures revealed that the three stem cell lineages exhibit distinct profiles. Analysis of the i-cell lineage revealed the presence of a common interstitial cell state which is a progenitor of gland cells and neural cells. The signature of the gland cell/neural cell precursor is distinct from that of nematocyte precursors suggesting a distinct developmental origin of this cell type with respect to neural and secretory cells. Furthermore, these findings provide a novel interpretation of the i-cell model in which the multipotent progenitor undergoes a first differentiation in either

nematocytes or the neural/gland progenitor. A second decision is then made by the neural/gland progenitor which migrates from the ectodermal layer and can provide neural or gland cells to the endodermal layer (Fig. 1.4b; Siebert et al., 2019).

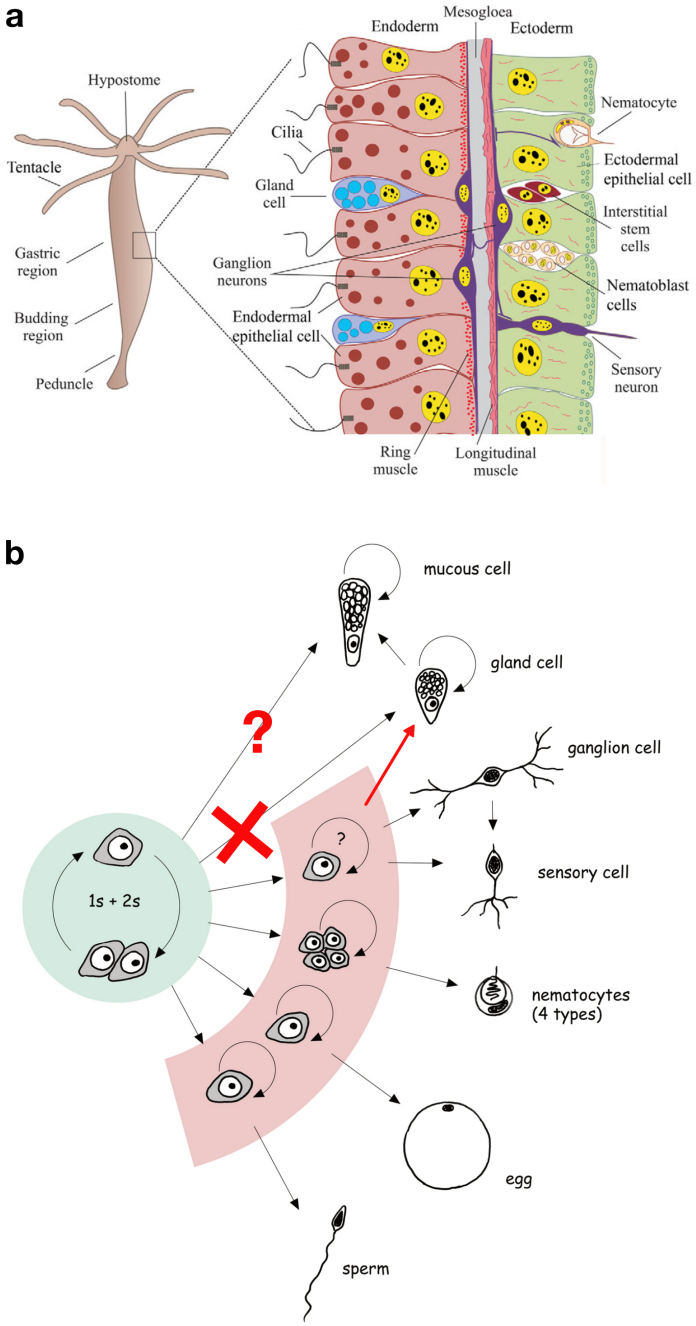


Figure 1.4 - Anatomy of *Hydra* and derivatives of the I-cells.
(a) *Hydra* polyp is a two-layered tube, with a ring of tentacles around the mouth opening. Asexual budding occurs laterally. Details of the two tissue layers show ectoderm and endoderm separated by the mesoglea. Interstitial cells and their derivatives are scattered between endodermal and endodermal cells. Taken from Technau & Steele (2012). (b) Schematic of cell proliferation and differentiation in the i-cell lineage. Multipotent progenitors are highlighted in green. Committed progenitors or transient

amplifying cells are in red. Taken from Hobmayer et al. (2012) and modified acoording to Siebert et al. (2019) (red cross, red arrow, red question mark).

1.3.2 - *Nematostella vectensis* and the evolution of nervous system

Nematostella vectensis was the first cnidarian model species to have its genome sequenced (Putnam et al., 2007; Technau & Steele, 2012). In contrast with *Hydra*, the life cycle of *Nematostella* includes a planula larva as well as a sessile polyp stage. The accessibility of the larvae has allowed detailed description of embryogenesis and larval development (Hand & Uhlinger, 1992; Fig. 1.5a), which have made *Nematostella* informative for evolutionary studies of bilaterian traits. Several studies have addressed axial patterning (Röttinger et al., 2012; Wikramanayake et al., 2003), and mesoderm evolution (Martindale, 2005) aiming to establish a comparison with Bilateria and gain insight on the evolution of those features. Other studies of *Nematostella* have focused on the nervous system to gain knowledge on the neural architecture of the common Cnidaria/Bilateria ancestor. A degree of conservation has been uncovered as explained more below.

The neural cells of *Nematostella* are organised in a nerve net along the oral-aboral axis and distributed in the two tissue layers. Distinct neural populations are detected in specific regions of the body, such as the pharyngeal and oral nerve rings, tentacle tips and larval apical tuft (Marlow et al., 2009; Fig. 1.5b).
Neurogenesis occurs in the ectoderm at gastrula stage. Analysis of bHLH and HMG-box *Sox* transcription factor families indicate many similarities with the neurogenesis described in Bilateria, suggesting conserved functions in neurogenic pathways (Layden et al., 2012; Marlow et al., 2009). Furthermore, Notch signalling during *Nematostella* development inhibits downstream neurogenesis pathways in a likely-conserved role to that in Bilateria (Richards & Rentzsch, 2015).
A recent study provided molecular signatures at single cell level of the larva and the adult of *Nematostella* (Sebé-Pedrós, Saudemont, et al., 2018). Analysis of the cell type atlases revealed two broad groups of neural cells including several subtypes with potential specialised functions. Analysis of the larval neural cells identified a novel transcriptional state putatively involved in neural function which indicates the identification of a novel neural population specific to the planula stage (Sebé-Pedrós, Saudemont, et al., 2018). The high resolution of the molecular signatures permitted a comparison with some bilaterian species, in particular with the nematode *C.elegans*. Comparative analysis of broad neural signatures revealed a degree of similarity of co-expressed modules across *Nematostella* and *C.elegans* supporting the existence of a neural-like cell type in the common Cnidaria/Bilateria ancestor. On the other hand, no clear correlation across species was observed when comparing gene signatures of specialised neural subtypes. This suggests that the assembly of neural molecular components might occur independently in each lineage and coincide with the emergence of specific neural subtypes (Sebé-Pedrós, Saudemont, et al., 2018).

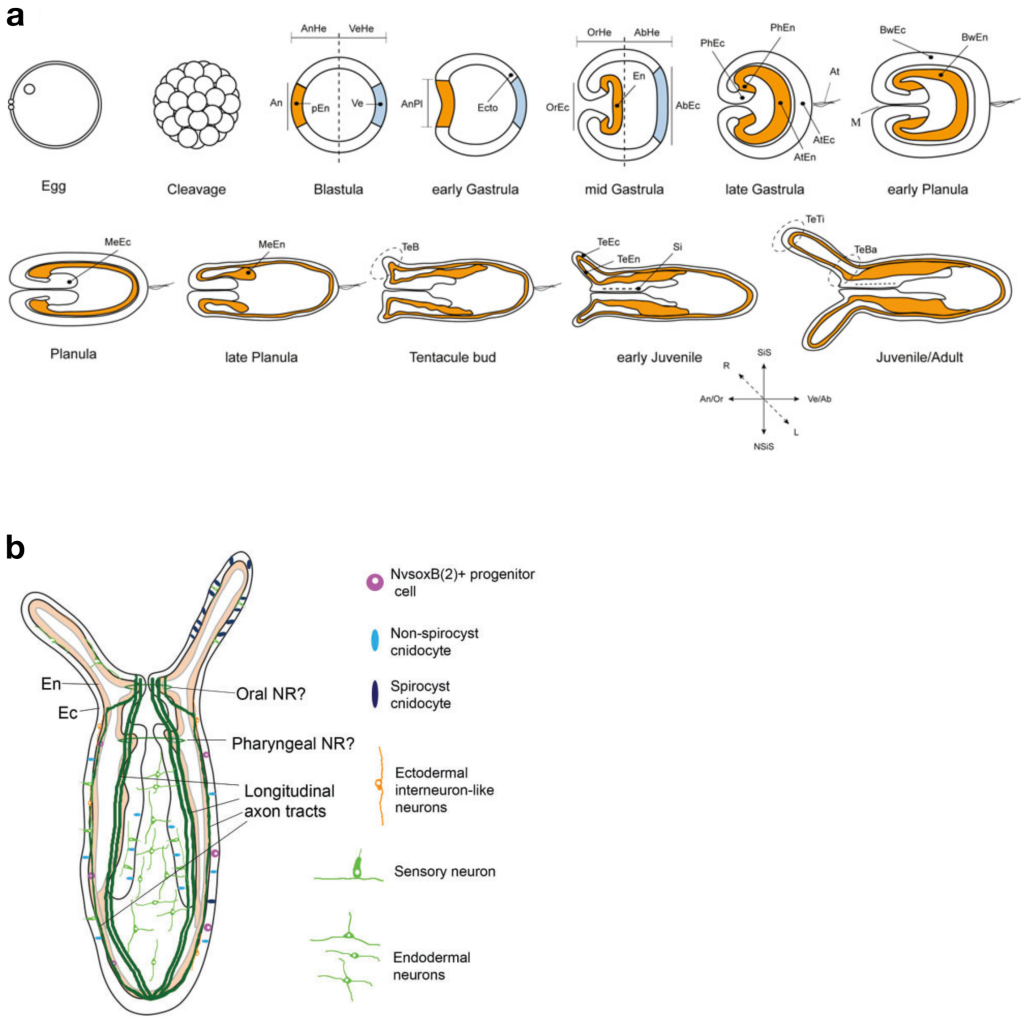


Figure 1.5 - Development of *Nematostella* and anatomy of the nerve net

(a) Schematics of the development of *Nematostella vectensis* from embryo to adult. Abbreviations are: Animal Plate(AnPl) Animal pole (An). Aboral Ectoderm(AbEc) ,Aboral Hemisphere (AbHe), Animal pole (An), Animal Hemisphere (AnHe), Animal/Oral (An/Or),Animal Plate (AnPl),Apicaltuft (At), Apical tuft Ectoderm (AtEc), Apical tuft Endoderm (AtEn), Body wall Ectoderm (BwEc), Body wall Endoderm (BwEn), Endoderm (En), Ectoderm (Ecto), Left (L), Mouth (M), Mesentery Ectoderm (MeEc), Mesentery Endoderm (MeEn), Non-Siphonoglyph Side (NSiS), Oral Ectoderm (OrEc), Oral Hemisphere (OrHe), presumptive Endoderm (pEn), Pharyngeal Ectoderm (PhEc), Pharyngeal Endoderm (PhEn),Right (R),Siphonoglyph (Si), Siphonoglyph Side (SiS), Tentacle Bud (TeB), Tentacle Base (TeBa),Tentacle Ectoderm (TeEc), Tentacle Endoderm (TeEn), Tentacle Tip (TeTi), Vegetalpole (Ve),Vegetal/Aboral (Ve/Ab), Vegetal Hemisphere (VeHe). Taken from Ormestad et al. (2011). (b) Schematics of the organisation of the nerve net of *Nematostella*. Taken from Layden et al. (2016).

1.4 - *Clytia hemisphaerica* and the evolution of the hydrozoan life cycle

Clytia hemisphaerica was selected as an experimental model species around 15 years ago for its suitability in addressing embryonic axis specification (Houliston et al., 2022; Momose et al., 2008; Momose & Houliston, 2007).
Clytia belongs to Hydrozoa and unlike its cousin *Hydra* exhibits the 'complete' hydrozoan life cycle including a free-swimming medusa stage. The life cycle of *Clytia* comprises three main life forms. The medusa is the sexually reproductive stage. Male and female medusae release gametes into the sea water triggered by light on a daily basis (Quiroga Artigas et al., 2018). From fertilised eggs, a ciliated planula larva forms in about one day. In laboratory conditions, after three days the larvae undergo a drastic metamorphosis during which the body flattens against the substrate and it gives rise to a sessile primary polyp, or gastrozooid, in about 24 hours. Once fully formed, the primary polyp is capable of catching prey and feeding using the tentacular system which surrounds the mouth. The primary polyp is the founder of the polyp colony and it propagates by stolon extension. The polyp colony includes a second form of polyp, the gonozooid, which is a specialised form of polyp that releases clonally identical baby medusae by budding. The polyp colony is potentially immortal due to the constant vegetative growth and replacement of the old parts. Once released by the gonozooid, the baby medusae reach sexual maturity in two-three weeks and can live for about one month in laboratory conditions (Fig. 1.6A and B; Houliston et al., 2010; Lechable et al., 2020).

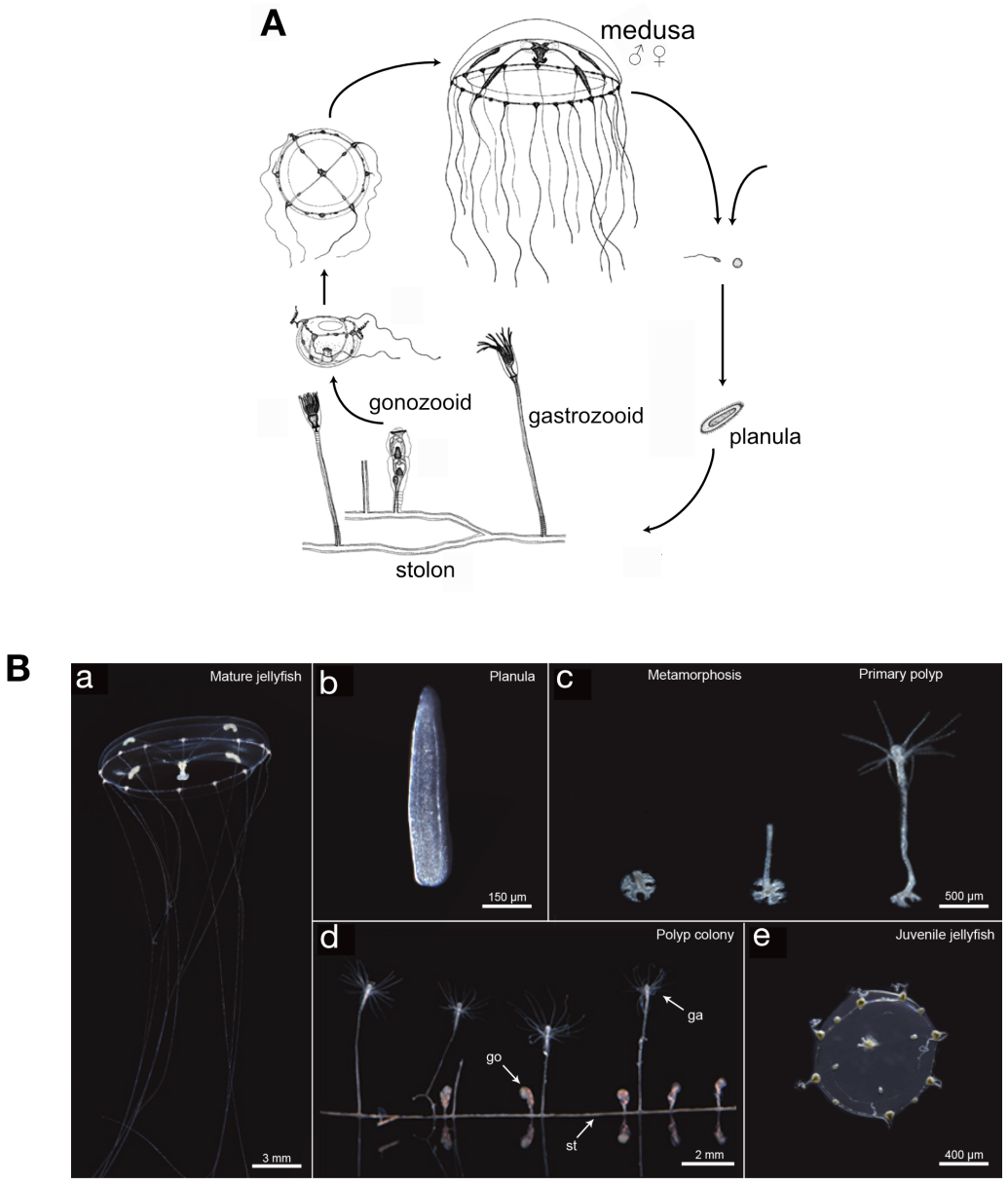


Figure 1.6 - The life cycle of *Clytia*
(A) Schematics of the life cycle of *Clytia*. Edited from Leclère et al. (2019). (B) Images of the life cycle stages: (a) adult medusa. (b) planula larva; the aboral pole at the bottom. (c) Steps of metamorphosis and formation of the primary polyp. (d) polyp colony. Arrows indicate gonozooid (go), gastrozooid (ga) and stolon (st). (e) juvenile medusa. Edited from Lechable et al. (2020).

All the life cycle stages, including the eggs, exhibit transparency (Fig. 1.6B) and the entirety of the life cycle can be reproduced in the laboratory. Over the years a culture protocol has been established. Crucial aspects are the use of different tanks for each life stage and ensuring a continuous water flow. Baby medusae are kept in “nursery” tanks with a continuous and circular water flow to provide a constant movement and avoid the sinking of the medusae.

Similarly, adult medusae are kept in modified “Kreisel tanks”. Spawning of adult medusae is achieved daily by light cues (see Chapter 2). Embryos are cultured at 17 degrees in small glass dishes with sea water and a mix of antibiotics. Polyp colonies are raised on glass slides suspended in commercial zebrafish tanks. Larval metamorphosis naturally occurs after 3 days post fertilisation upon uncharacterised bacterial cues. It can be induced by adding synthetic GLWamide peptide to the sea water, allowing different strains to be produced and maintained easily (Lechable et al., 2020).

Established experimental methods and resources for *Clytia* include *in situ* hybridisation, transmission electron microscopy, genomic resources (for instance an EST collection), and genetic tools, for instance gene knockdown and a very recent transgenesis method (reviewed in Houliston et al., 2022).
In 2019 the genome of *Clytia* along with transcriptomic data for each stage became available (Leclère et al., 2019). After *Hydra* (Chapman et al., 2010), *Clytia* is the second hydrozoan species for which the genome has been sequenced and the first one with a complete life cycle (Leclère et al., 2019). The availability of molecular data for all stages allowed a detailed characterisation of differentially expressed genes across the life cycle, showing greater complexity of gene expression within polyp and medusa stages. The medusa represents an evolutionary novelty of Medusozoa but its evolution is not associated with the evolution of new genes (Gold et al., 2019; Khalturin et al., 2019; Leclère et al., 2019). Indeed, comparative analysis across stages of *Clytia* of taxa-restricted genes do not show a preferential enrichment in the medusa stage. On the other hand, analysis of transcription factors showed a larger number associated with the medusa stage, with a subset of those shared between medusa and polyp. Fewer transcription factors were associated with the planula stage (Fig. 1.7a). Notably, a large number of transcription factors that are conserved within Bilateria, known to be involved in development, have been detected in different sites of the well-organised medusa nervous system, but not at planula stage (Fig. 1.7b). This suggests a possible association with diverse cell types in the medusa. In contrast gene families associated with neurogenesis have been detected across all the life stages. This indicates that a shared molecular repertoire is involved in the genesis and maintenance of neural cell types. However, differentiated cell types at planula and medusa stages might rely on distinct molecular and cellular programs (Leclère et al., 2019). The association of the expanded expression of transcription factors with the emergence of putative novel cell types in the medusa stage is addressed in this thesis.

In the following subsections I present the general anatomy of the medusa and the planula of *Clytia,* which are the main focus of my study, along with the state of knowledge concerning their cell types prior to this study.

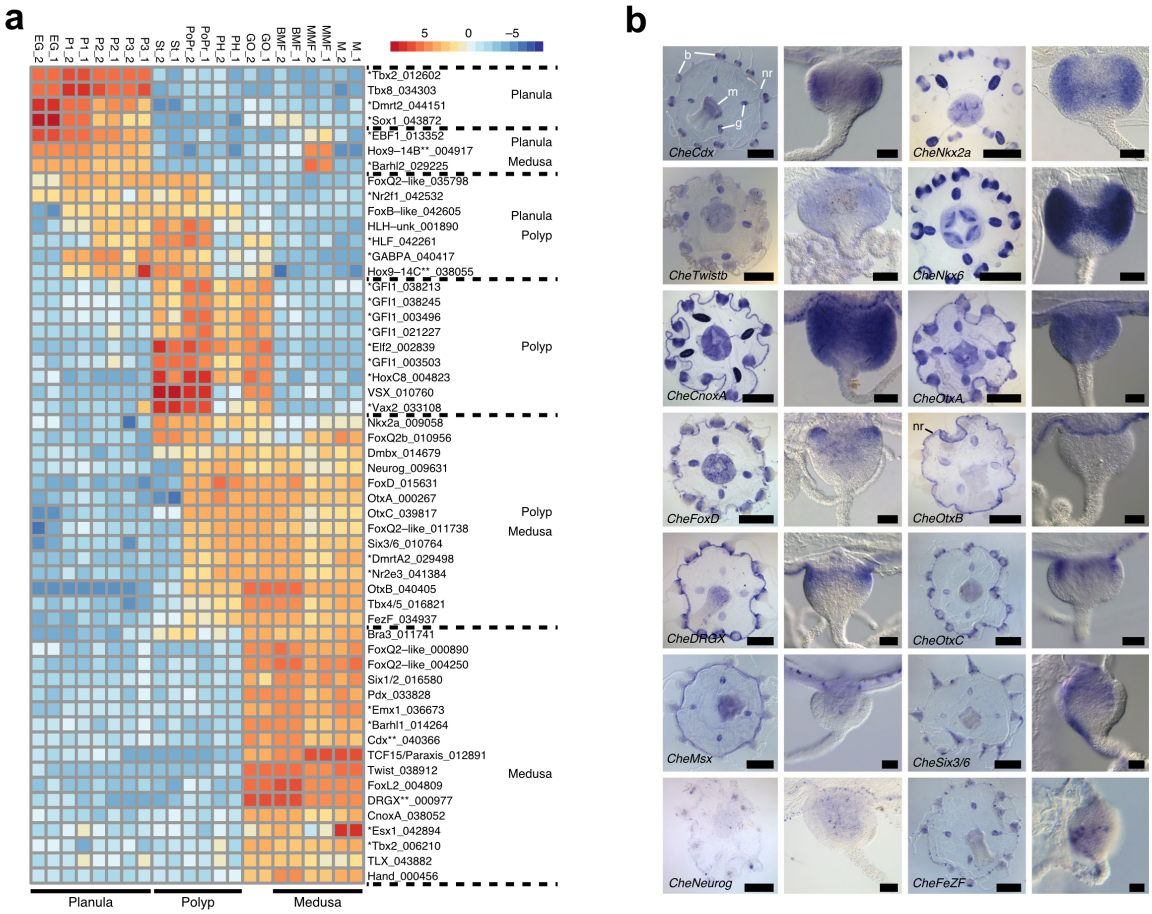


Figure 1.7 - Analysis of transcription factors and expression patterns
(a) Heatmap of differentially expressed transcription factors across stages. Abbreviation on the top of the heatmap are; EG, early gastrula; P1/P2/P3, planula at 24 h/48 h/72 h after fertilisation; PoPr, primary polyp; St, stolon; GO, gonozooid; PH, gastrozooid/polyp head; BMF, baby female medusae 1 day old; MMF, mature female medusa; M, mature male medusa. (b) In situ hybridization of whole female medusa (left, scale bar represents 500 µm) and tentacle bulb (right, scale bar represents 50 µm) of medusa-enriched transcription factors. Both (a) and (b) are taken and edited from Leclère et al. (2019).

1.4.1 - The adult medusa of *Clytia*

In the past few years, several medusozoan genomes have become available (Gold et al., 2019; Khalturin et al., 2019; Kim et al., 2019; Leclère et al., 2019) and interest concerning the complexity of medusae has increased. The fast-swimming behaviour of the medusa requires specialised muscle cells and a coordination of those by a relatively organised nervous system (Khalturin et al., 2019). The tentacular systems bear stinging cells which are used as weapons to actively catch prey for feeding (Condamine et al., 2019). Food is processed in a relatively organised digestive system. In cubozoan and scyphozoan medusae specialised sensory

organs termed rhopalia have been described. Cubozoan medusae have the most complex ones, each containing six eyes, two of which are sophisticated camera-eyes, and a statocyst for orienting the rhopalia (Nilsson et al., 2005). Some hydrozoan medusae also have statocysts and even sometimes eyes (Singla, 1975; Suga et al., 2010).

The medusa of *Clytia* develops within the specialised gonozooids of the polyp colony and is released by lateral budding (Houliston et al., 2010). The budding process starts with an outgrowth of the ectodermal and gastrodermal layers of the polyp (Kraus et al., 2015). A layer of cells delaminate and form a transient structure called the entocodon, which will develop into smooth and striated muscles of the sub-umbrella. The ectodermal protrusion will generate the exumbrella, the external layer of the velum and the tentacle epidermis. The gastro digestive system will generate from the endodermal bud (Kraus et al., 2015). The gelatinous body of the medusa of *Clytia* is organised according to a tetra-radial symmetry. The bell-shaped umbrella is constituted by two layers, the outer exumbrella and the inner sub-umbrella separated by a thick layer of acellular mesoglea (Fig. 1.8; Houliston et al., 2010). The mouth, or manubrium is placed in the centre of the subumbrella. The base of the manubrium is constituted by the gastric cavity, or stomach, and it is connected to four tubular-shaped radial canals (Fig. 1.8). The radial canals run along the subumbrella and are connected to the gonads and to a circular canal which runs around the bell margin (Fig. 1.8). The canals serve to distribute nutrients from the manubrium to the other organs. The circular canal connects the tentacular system, more precisely joining the tentacle bulbs which produce the tentacles (Fig. 1.8). When baby medusae bud from the gonozooids, they have four bulbs linked to the four radial canals. As the medusae grow, the number of bulbs increases to 32 (Houliston et al., 2010; Schmid et al., 1974). Statocyst sensory structures are located on the bell margin between each pair of tentacle bulbs (Singla, 1975). The nervous system consists of a diffuse nerve net which innervates the subumbrella and the organs, and two condensed nerve rings, motor and sensory nerve rings, which run round the bell margin and are connected with the statocysts (Fig. 1.8; Houliston et al., 2010; Weissbourd et al., 2021).

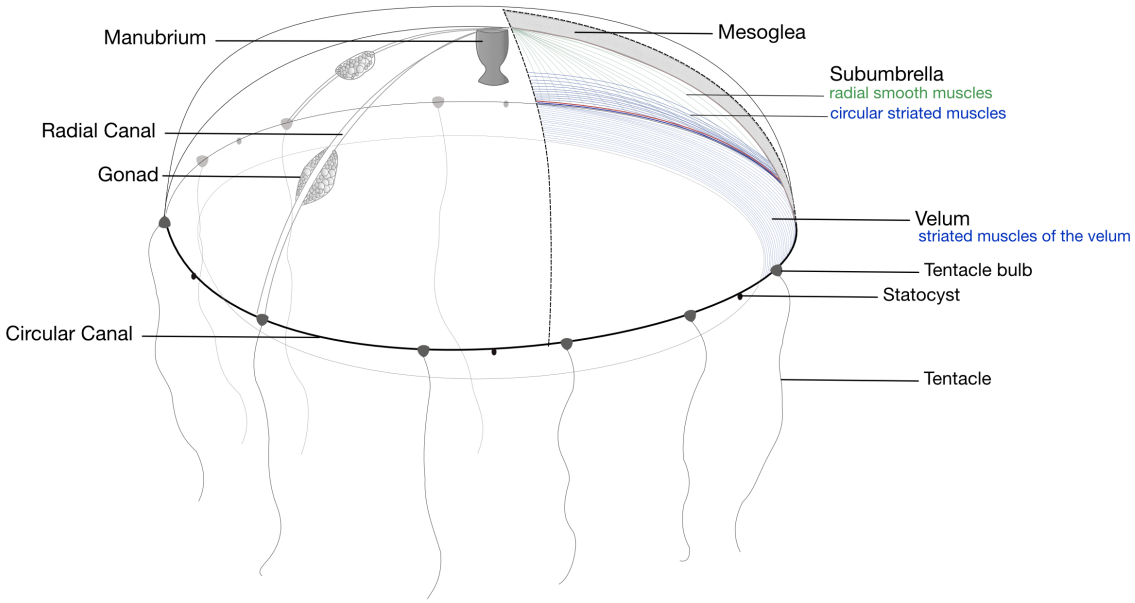


Figure 1.8 - Diagram of the medusa of *Clytia*
Edited from Chari et al. (2021).

The medusa exhibits great ability to regenerate organs and part of the bell (Sinigaglia et al., 2020). Even when dissected, the manubrium, the gonad and the tentacle bulbs remain functional for some time, for instance dissected tentacle bulbs will continue to produce tentacles. This is in part due to the population of hydrozoan i-cells, which are well described in *Hydra* (see above; Sinigaglia et al., 2020). I-cell pools in the medusa are present in the gonads and in the proximal region of the tentacle bulb ectoderm (Leclère et al., 2012).

In the *Clytia* medusa the relationships between i-cells and their derivatives, such as neural cells and nematocytes, has been partly resolved by analysis of sc-RNAseq data (Chari et al., 2021) and is discussed in Chapter 2. The generation of nematocytes from i-cell progenitors , a process highly active in the tentacles, was already well characterised prior to the generation of the cell atlas. Indeed, the well-ordered tentacular system of the medusa, described as a cellular conveyor belt, has allowed a precise molecular characterisation of nematocytes development (Condamine et al., 2019; Denker et al., 2008). The pool of stem cells in the proximal region of the ectoderm of the bulb proliferates and differentiates into mature nematocytes while migrating towards the tentacle (Condamine et al., 2019; Fig. 1.9). Analysis of transcriptomic data obtained from isolated parts of the tentacle, notably the proximal, middle and distal regions, revealed differentially expressed genes across the samples. The Wnt signalling pathway potentially plays a role in their regulation given the spatially distinct expression of Wnt-related genes in the tentacle and tentacle bulb (Condamine et al., 2019).

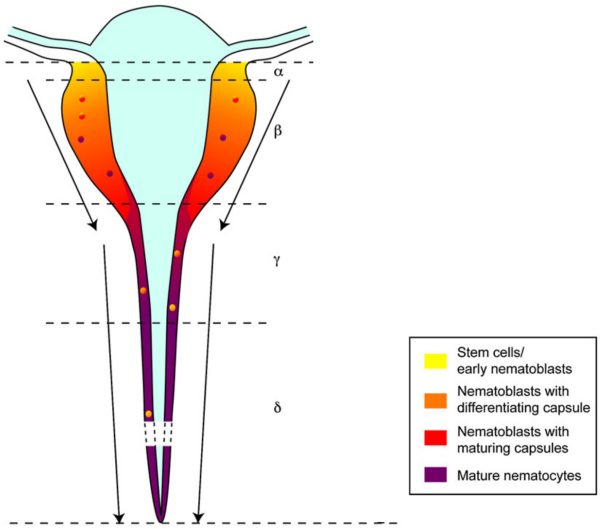


Figure 1.9 - Model of nematogenesis in *Clytia* medusa tentacles
Diagram of the tentacle bulb of a *Clytia* medusa with area of progressive nematogenesis highlighted. Greek letters indicate portions of the tentacle bulb and tentacle. Arrows indicate the direction of the differentiation. Edited from Denker et al. (2008).

As already mentioned, fast contractile muscle fibres are needed for the swimming movement of the medusa. Muscle fibres lie in the subumbrella layer, which includes the endodermal plate associated with the mesoglea, an overlapping layer of radially oriented smooth epitheliomuscular cells and a layer of circularly oriented striated epitheliomuscular fibres. The striated fibres are only found in the most peripheral part of the bell and are responsible for the fast contractions during swimming (Leclère & Röttinger, 2017; Sinigaglia et al., 2020; Fig. 1.8).

Previous studies dating back to 1966 have described at least three types of glandular digestive cells in the medusa of *Clytia* based on morphological observation (Bouillon, 1966). Two of those were mainly observed at the apical side of the manubrium while the third one was mainly observed in the region of the stomach. Sc-RNAseq data allowed the characterisation of additional digestive gland cell types (Chari et al., 2021) which are described in detail in Chapter 2.

1.4.2 - The planula larva of *Clytia*

The planula larva of *Clytia* has a diploblastic structure with two tissue layers, and a simple body plan organised along an oral-aboral axis (Bodo & Bouillon, 1968; Freeman, 2005). The anterior end or aboral pole is more rounded and the posterior end or oral pole, more pointed

(Fig.1.10). The larva swims by directional beating of the ectodermal cilia, moving with the aboral pole in front.

The two main tissue layers forming the planula are the outer ectoderm and the inner endoderm (Fig.1.10). At the cellular level, the formation of these layers begins at the blastula stage, when an initial epithelialization of the blastodermal cells occurs. During this phase of embryogenesis blastodermal cells start to change shape and each becomes polarised according to an apical-basal axis. As the polarisation proceeds, blastodermal cells develop cilia that start beating. With the progression of the blastula stage the cells become columnar, although the shape of individual cells is not uniform. At the late blastula stage the blastoderm becomes thicker in a specific region that will be the future oral pole. At the onset of gastrulation, some cells at the oral pole adopt a particular shape exhibiting an elongated apical-basal axis and enlarged basal sides. Because of this peculiar shape, they are defined as *"bottle cells"* (Kraus et al., 2020). The formation of the bottle cells is associated with Epithelial-mesenchymal transition, or EMT (Shook & Keller, 2003). Indeed, following the progressive enlargement of the basal side and the restriction of the apical domain, these cells, the presumptive endoderm, detach from the epithelium and adopt a mesenchymal morphology starting to migrate into the blastocoel. At the mid gastrula stage the site of ingression becomes thinner and the embryo exhibits a more rounded shape at the aboral pole and becomes more pointed at the oral pole. Bottle cells are still present at the oral pole of the late gastrula indicating that in *Clytia* endoderm is generated during the whole gastrulation period. At 24 hpf, cell ingression terminates and ingressed cells start to form the endodermal epithelial layer. At 48hpf the formation of the planula is complete and the distinctive shapes of the oral and aboral pole become very clear. Further cell differentiation continues until 72hpf (Fig. 1.10; Kraus et al., 2020).

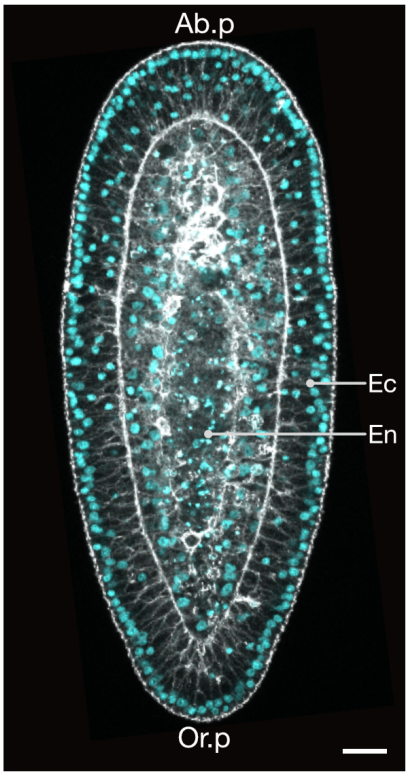


Figure 1.10 - The planula larva of *Clytia*
Confocal image of 48hpf planula larva. Abbreviations are: Aboral pole (Ab.p), Oral pole (Or.p), Ectoderm (Ec), Endoderm /Gastrodermis (En). Nuclei in cyan (Hoechst); Cell boundaries in grey (Phalloidin). Maximum projection of two Z-planes from a confocal stack. Scale bar 50μm.

The planula stage, as well as embryonic stages, have proven suitable to address questions concerning the evolution of axial patterning systems. The developmental patterning of the planula is regulated by evolutionary-conserved programs, notably the Wnt pathway, which defines the oral-aboral axis (Momose et al., 2008; Momose & Houliston, 2007). Potential patterning genes in the planula acting downstream of Wnt signalling and affected by PCP were identified by an unbiased transcriptome comparison (bulk RNA seq) between early gastrula stage embryos in which these pathways were experimentally up or downregulated (Lapébie et al., 2014). The transcriptional responses of the identified genes correlated with regionalised expression in the embryo and larva. These genes, which include both members of known families of developmental regulators as well as poorly characterised taxon-restricted genes, are candidates for patterning roles but also as differentiation genes for different cells (Lapébie et al., 2014).

Early embryonic expression of genes belonging to the Forkhead subfamilies (Fox) suggests an evolutionary conserved role of these genes during embryogenesis in *Clytia* (Chevalier et al., 2006). Expression during planula development as well as specific structures in the medusa suggest that members of the Fox subfamilies might play an additional role in cell type

development that has been lost in Bilateria or acquired in Cnidaria. For instance, expression of FoxB and FoxQ2 family genes is detected in early stages of embryogenesis (eggs and blastula and gastrula). This is consistent with involvement in early embryonic pattern systems as is also seen in Bilateria. *Clytia* FoxB is also detected later in the development of the planula and is associated with a putative nematogenic site as well as three distinct regions in the adult medusa suggesting a possible involvement of this gene in the development of neural cell types (Chevalier et al., 2006).
Comparison with other cnidarian as well as bilaterian species shows that some aspects of the gene expression and function are also conserved in *Clytia* for the Sox family genes. The expression of some Sox genes in *Clytia* show consistency across life stages at the cell type level. For instance, four Sox genes contribute to the i-cell lineage and five are expressed in nematoblasts in both planula and medusa. On the other hand, other Sox genes show poor correlation of gene expression across life stages within potentially similar cell types. For instance, *Sox10* in the medusa is expressed in a subpopulation of neurons as well as i-cells and maturing oocytes suggesting a role in i-cell-derived neurogenesis. In contrast in the planula, *Sox10* expression is associated with putative neurogenic cells forming an inner layer of the ectoderm at the gastrula stage, distinct from the i-cells (Jager et al., 2011; Kraus et al., 2020). This suggests a potential role in the diversification of planula specific cells and/or a possible association with a pathway for neurogenesis in the planula not mediated by i-cells. This will be discussed further in Chapter 3 and Chapter 4.
Phylogenetic analysis of Hox, Parahox and Hox-related genes have identified orthologous genes in *Clytia*. Analysis of the expression patterns of seven genes revealed that those belonging to Hox9-14 and CDX groups showed expression along the oral-aboral axis during planula development suggesting a potential conserved role in axial patterning. However, other Hox genes were observed to be expressed only at medusa stage in putative specialised regions. For instance, *cheGSX* and *cheHOX1* show expression respectively in tentacle bulb nematogenic cells and in statocysts, implying a potential involvement in cell type specification (Chiori et al., 2009).
These findings indicate that some well known patterning genes potentially play evolutionary conserved roles in the development of *Clytia*, while others show expression in the medusa stage of the hydrozoan life cycle in particular cell types .

In this context, the characterization of the cell types is crucial to gain insights on the evolution of conserved pathways and to determine the possible role of these genes in the development of specific cell types.

1.4.2.1 - Cell types of the planula of *Clytia*

Prior to the present work, our knowledge of the *Clytia* planula cell types has relied on histological studies.

The first study on the planula larva of *Clytia* dates back to 1886 when Elie Metchnikoff provided a description of jellyfish he collected at Napoli, Trieste and Villefranche-sur-mer. In his work

"Embryologische Studien An Medusen" (Metchnikoff, 1886) he published his studies on cnidarian embryology in which he describes the development from eggs to adult stages of some cnidarian species, including *Clytia*. Almost a century later, in 1968, Jean Bodo and France Bouillon published their comparative studies of hydrozoan larvae belonging to five species collected at Roscoff, including *Clytia hemisphaerica*, referred to as *Phalidum hemisphaericum.* This study represents the first characterization of the cell types of *Clytia* the planula based on histology and morphological observations. Following classical histological approaches involving fixation, embedding of the specimens, mainly in paraffin, for sectioning and finally classical stainings highlighting different kinds of macromolecules, the authors were able to discriminate between different cells and provide the first detailed characterization of cell types with a particular focus on glandular cells and nematocytes (Fig. 1.11; Bodo & Bouillon, 1968).

Two types of glandular cells were identified. The first type is reported as 'cellules glandulaires spumeuses' (termed 'foamy' mucous cells by Campbell & Bode (1983)) differentiating from the aboral ectodermal cells already at the gastrula stage. These are characterised by the presence of multiple irregular vacuoles containing foamy secretion materials that are described as of an acid mucopolysaccharide nature (Bodo & Bouillon, 1968). The nucleus is mainly located on the basal side of the cells while the vacuoles are concentrated in the apical side (Fig. 1.11). A morphologically similar cell type has also been observed in adult polyps and medusae of several species, notably in the hypostomal region of the polyps and in the manubrium (mouth) of medusae (Bouillon, 1966). In the adults such foamy gland cells differentiate mainly within the endoderm.
A second type of glandular cell was described as 'cellules sphéruleuses' (termed 'granular' mucous cells by Campbell & Bode (1983)) which are characterised by multiple small granules or droplets containing muco-proteinic secretion materia (Bodo & Bouillon, 1968). These cells appear to differentiate from the aboral ectoderm in late planula stages and are reported to be less numerous than the foamy glandular cells. The shape and the secretory mode of action of these cells appears to be similar to the spheroulous glandular cells described in adult polyps and medusae (Bouillon, 1966). However, the latter is exclusively localised in the endoderm of the hypostome and the manubrium of the adults (Bouillon, 1966).

The ectodermal cells located along the flanks of the larva on the median band surrounding the larval body are characterised by the presence of small granules. Bodo & Bouillon (1968) suggested a relationship between these granules and the formation of the 'perisarc' such as the typical chitinous layer produced by the hydrozoans as a protection for the soft parts of the body, notably the 'theca' (Fig. 1.11).

A cell type reported as of endodermal origin is the i-cells, characterised by a rounded shape and a large nucleus and identified predominantly in the endoderm or located between endodermal cells and the mesoglea (Fig. 1.11).

According to Bodo and Bouillon (1968), mature nematocytes are mainly found in the ectoderm at the oral (posterior) pole of the planula (Fig. 1.11). Early stages of nematocytes, the nematoblasts, can be distinguished as early as the gastrula stage, persisting throughout later planula stages in the endoderm (Fig. 1.11). Nematoblasts migrate through the mesoglea,

towards the ectoderm and completely differentiate following migration. Nematoblasts were observed earlier in development than i-cells suggesting that these might both originate directly from the differentiation of some ‘cellules endodermique banales’, or classical epithelial endodermal cells. Furthermore, in some older larvae, after at least four days post fertilisation, nematoblasts were observed to continue to proliferate and differentiate but migration towards the ectoderm was not reported (Bodo & Bouillon, 1968).
Another type of cell characterised by irregular granules in the cytoplasm was described, mainly in the ectoderm at the oral pole. The granules were defined as ‘excretion granules’, possibly corresponding to residues of digestion of the larval yolk. However these types of cells were only found rarely and only in older larvae (Bodo & Bouillon, 1968).

Potential planula neurosensory and ganglion cell ultrastructures were described in another study published in the following years, in which transmission electron microscopy (TEM) was used to describe cell types in the *Clytia* planula (Thomas et al., 1987). Neurosensory cells were observed mainly at the aboral pole and laterally in the ectoderm. These cells are characterised by the presence of a distinctive invagination of the apical cytoplasm surrounding the base of the cilium and by a long ciliary rootlet. Neurosecretory-like vesicles were observed around the nucleus along with neurites closely associated with the basal contractile processes of the epithelio-muscular cells. Ganglion cells were identified by the unique ovoid shape and by free ribosomes in the cytoplasm. These cells were mainly found at the base of the epidermis layer. Neurite-like projections of ganglion cells were described as forming a network with other neurites from the neurosensory cells at the level of the mesoglea and epithelio-muscular cells (Thomas et al., 1987).

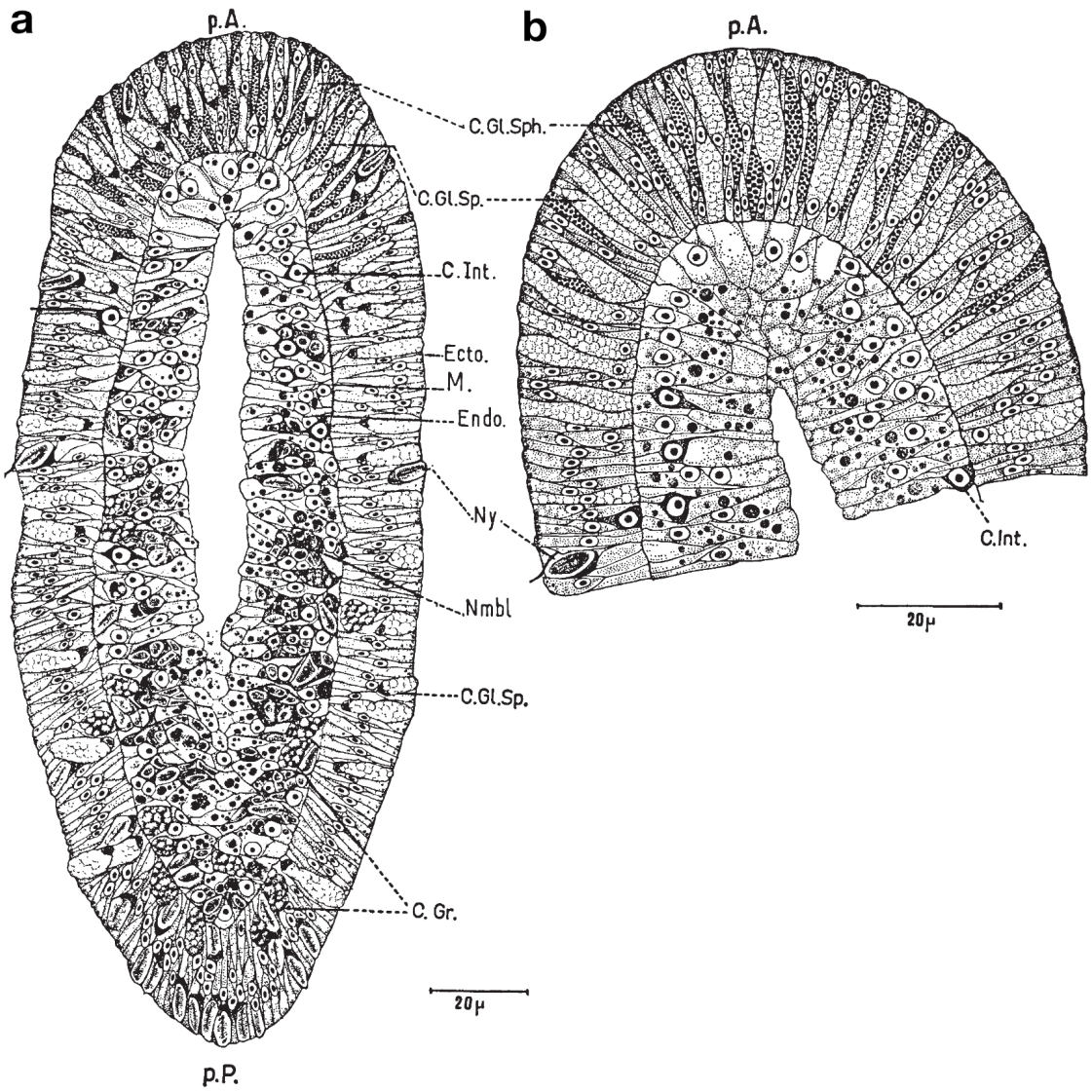


Figure 1.11 - View of the cells of the planula

(a) Diagram of a longitudinal histological section of the planula in a terminal phase of the development. (b) Detail of the anterior region of the planula with the two types of aboral glandular cells highlighted. Abbreviations are: glandular “foamy” cells (C.Gl.Sp), granular mucous cells (C.Gl.Sph), Ectodermal granular cells (C.Gr), Interstitial cells (or I-cells, C.Int.), ectoderm (Ecto), endoderm (Endo), mesoglea (M), nematoblasts (Nmbl), nematocytes (Ny), aboral or anterior pole (p.A), oral or posterior pole (p.P). Taken from Bodo & Bouillon (1968).

1.5 - Objectives

In this Introduction chapter I have illustrated the concept of the evolution of the cell types which proposes a novel perspective to trace the evolutionary history by using the cells as units. In the framework of large comparative studies, Cnidaria is well placed as an “early branching” taxon, sister group to Bilateria. I introduced the model species used in this project, *Clytia hemisphaerica*, and illustrated what is known concerning its cell types prior to this study. The characterisation of the cell types of *Clytia* can contribute to expanding the knowledge on the evolution of metazoan cell types in a broader comparative context. Furthermore, *Clytia* belongs to Hydrozoa which, among Cnidaria, is considered the group with the largest diversity of life cycles. The accessibility of the adult medusa and the planula larva permits the application of innovative single cell transcriptomics approaches to characterise molecular signatures of cell types. Besides representing a valuable resource for the *Clytia* and the cnidarian community in general, those data serve as a foundation to establish a comparative analysis of the cell types within a complete hydrozoan life cycle allowing the complexity of the medusa to be compared to the much simpler planula.

Based on these concepts my thesis work has three main objectives:

- The characterisation of the cell types of the adult medusa. This required the establishment of a dissociation protocol and generation of single cell transcriptomics data. The validation of the data via *in situ* hybridization and the interpretation of gene expression patterns permitted the generation of the cell atlas of the medusa. (Chapter 2).

- The characterisation of the cell types of the planula larva. As for the previous objective, this required establishment of a dissociation protocol, validation and interpretation of the data and finally the generation of the cell type atlas of the planula larva. (Chapter 3).

- Comparative analysis of cell type signatures by exploiting the cell atlases of the medusa and the planula of *Clytia* as a foundation. The ultimate aim of my PhD project was to resolve similarities and differences across life stages of *Clytia* and to start to address the complexity of the medusa in relation to a putative expansion of cell types. (Chapter 4).

CHAPTER 2 -

Cell types of the adult medusa of *Clytia*

Overview of the chapter

The free-swimming medusa stage, or jellyfish, is the fundamental innovation of Medusozoa, which is considered the sister group of Anthozoa. Medusozoa includes Hydrozoa, Scyphozoa, Staurozoa and Cubozoa.

The medusa represents the most complex body form of the cnidarian lineage. The body plan is organised with radial symmetry and shows distinct organs such as the manubrium (mouth/stomach) and the gonads. The medusa actively catches prey using the typical cnidarian weapons, the cnidocytes or nematocysts. The gelatinous bell is constituted by various tissues along with the acellular mesoglea, derived from the classical cnidarian epitheliomuscular cells of the ectoderm and endoderm layers. These include on the subumbrellar side striated muscle fibres that provide fast contractions for swimming. Swimming and other sophisticated behaviours unique to the medusa stage are coordinated by a well organised nervous system involving two condensed nerve rings running around the bell margin, which in some species are linked to additional organs such as statocysts and even complex camera-eyes.

In the past few years, several medusozoan genomes have become available, raising interest in understanding which cell types contribute to the complexity of the medusa stage. In this Chapter I introduce the cell type atlas of the medusa of *Clytia* which I generated exploiting single cell transcriptomics. I obtained the data using freshly dissociated cells from individuals that were just about to reach sexual maturity. I integrated the data from these cells with a recently published dataset generated by our collaborators in the Pachter and Anderson groups. In the context of that collaboration, I contributed to the interpretation and the validation of the data by mining marker gene lists and performing *in situ* hybridisation of selected candidate genes. Using existing information about genes expressed in each cell cluster and additional *in situ* hybridisation analysis we could assign cell type identities to 36 clusters, which we grouped into seven broad cell classes. To address the complexity of the medusa nervous system we isolated and reclustered the neural cells and selected neural-specific candidate genes to carry out *in situ* hybridisations. Analysis of the expression patterns revealed 14 putative neural subpopulations. We charted the origin and development of neural cells and nematocytes from the i-cells by performing pseudotime analysis. Trajectory analysis of the nematocyte lineage revealed two distinct transcriptional programs, which we characterised as the “nematoblast” phase, defined by the production of the nematocyte capsule, and the “nematocyte differentiation” phase, distinguished by the production of the nematocil apparatus.

My final integrated dataset consists of 25819 cells grouped in 40 cell types. I transferred the annotation we proposed for the first atlas achieving broadly equivalent cell types, with a slight increase in the numbers of neural cells and an additional cluster which likely represents technical artefacts. To further characterise the medusa cell types, I investigated their ultrastructure by electron microscopy, deducing identity from the spatial distributions and molecular signatures obtained in the atlas analysis.

This medusa cell atlas provides a valuable tool to address the complexity of the medusa stage at the cell type level. Along with available atlases of other cnidarian species, it opens the way to comparative analyses concerning cell type evolution within this group.

2.1 - INTRODUCTION

The central aim of my project is to establish a comparison at cell type level across life cycle stages of *Clytia*.

As described in Chapter 1, *Clytia* belongs to Hydrozoa and exhibits a 'complete' hydrozoan life cycle including three main life forms, the planula larva, the polyp colony and the adult medusa or jellyfish (Houliston et al., 2010). It displays a complex body organisation based on tetra-radial symmetry and proper organs such as the manubrium (feeding organ comprising mouth and stomach), the gonads and tentacle bulbs. Smooth and striated muscle fibres are present in the subumbrella, along with a nervous system condensed into two nerve rings, and diffuse nerve nets associated with tissue layers of the subumbrella and organs (see Chapter 1).

We initially characterised the cell types of the medusa of *Clytia* in the context of a collaboration with the Pachter and Anderson groups at Caltech. The resulting publication of the first cell type atlas of *Clytia* medusa, is included as an annex to this chapter (Chari et al., 2021).
I contributed to this collaboration mainly by interpreting and validating the identity of cell types by *in situ* hybridisation and confocal microscopy. The Pachter lab developed a multiplexed single cell sequencing approach which allows several samples treated with different conditions to be pooled into a single experimental run. As a proof of concept that this could be used to compare cellular responses within whole animals, they applied this approach to *Clytia* medusa fixed dissociated cells. For the main dataset, they pooled dissociated cells from animals in control condition and after a period of starvation. An additional experimental approach involved exposing medusae to multiple ionic stimuli with the aim to predict the expression of immediate early genes in *Clytia* which are described to be expressed in neurons that respond to stimuli (Sheng & Greenberg, 1990). These experimental designs resulted in the generation of two medusa cell type atlases, the "Fed and Starved" and the "Stimulation" atlases (Chari et al., 2021). We exploited lists of cluster-specific marker genes computed from an integrated "Fed and Starved" and "Stimulation" atlases. We attributed tentative identities using previously characterised genes (Balasubramanian et al., 2012; Condamine et al., 2019; Denker et al., 2008; Gillespie & Müller, 2009; Leclère et al., 2012; Leclère & Röttinger, 2017; Quiroga Artigas et al., 2018; Tucker, 2010) and selecting novel candidates by picking highly specific markers from the signature of the clusters. I then selected cluster specific candidates and carried out *in situ* hybridisation to assign identities. Analysis of the expression patterns revealed 36 cell types which we manually grouped in seven cell classes: the two epitheliomuscular layers (epidermis and gastrodermis), bioluminescent cells, germ cells grouped with the i-cells, and three classes of i-cell-derivatives notably, the nematocytes, the neural cells and the digestive gland cells (Chari et al., 2021). Comparisons of the scRNAseq data obtained from control and starved medusae revealed some specific changes in cell type transcriptomes. In particular, the oocytes and the gastrodermis cells were the most affected by starvation. This reflects a marked visible reduction in gonad size as well as in the overall decrease of the medusa body mass during starvation. I explored in more detail the consequences of starvation on the cellular organisation within the gonad by phalloidin/hoechst staining and confocal microscopy, by comparing control and starved medusae. This imaging analysis confirmed that during

starvation a substantial reorganisation of gastrodermal cells occurs along with depletion of the majority of medium sized oocytes (Fig.7 in Chari et al., 2021). In addition, we exploited the cell atlas to address the origin of neural cells and nematocytes from the i-cells, as has been shown in *Hydra*. Pseudotime analysis revealed continuity between i-cells to neural cells, and i-cells to nematocytes. Indeed, we were able to distinguish two phases of nematogenesis associated with two distinct but consecutive transcriptional programs. Additionally, we identified 14 putative medusa neural subpopulations by reclustering analysis of isolated neural cells. This analysis has been particularly useful for predicting novel neuropeptide precursors by exploring the gene signatures of the putative subpopulations (Chari et al., 2021).

In parallel to this collaboration on the inital atlas, I generated and analysed two additional female medusa single cell datasets using freshly dissociated cells from sexually immature jellyfish. I integrated these two datasets with the ones we published with our collaborators.
The integrated medusa cell atlas is presented in the section below along with the validation by *in situ* hybridisation and micrographs of the identified cell types.

2.2 - RESULTS

2.2.1 - A *Clytia* Medusa Cell Type Atlas

Clytia medusae can live for around one or two months in laboratory conditions (Lechable et al., 2020). To generate single cell transcriptomics data I chose to use 10-to-14 day-old jellyfish that had not yet reached sexual maturity. At maturation, gametes are continuously produced in the gonads to allow daily spawning. Before maturity, the numbers of these cells is lower, reducing the risk of predominantly sampling gametes.
I generated two single cell libraries by performing encapsulation experiments at two different institutes (at the France Genomics platform at the IPMC in Nice, and at the EMBL genomics core facility in Heidelberg, see Methods) using 10X Genomics technology (https://www.10xgenomics.com). For each experiment I loaded a maximum of 10000 freshly dissociated female medusa cells into the 10X Chromium controller. I generated two sets of data, respectively “IPMC” and “EMBL” datasets, that I integrated with the “Fed and Starved” medusa data from Chari et al (2021) and corrected for the batch effect using Harmony (Korsunsky et al., 2019; Fig. 2.1b).

My integrated dataset consists of 25819 cells grouped in 40 clusters that I refer to as cell types throughout (Fig 2.1a). Verification of batch effect correction demonstrated homogeneous mixing of the cells from the three batches into each cluster as well as some interesting differences across those (Fig 2.1b). For instance, the “IPMC” data were generated from a medusa that was just about to reach sexual maturity, so this dataset is dominated by oocyte mRNA signatures. Indeed, in the integrated atlas there is a large contribution to the oocyte class from this dataset. The medusae that I used to generate the “EMBL” data had gone through travelling and I couldn’t ensure the best conditions of feeding and seawater temperature once I arrived in Heidelberg. Therefore, at the time of the experiments, the medusae were starving and likely physiologically stressed. This resulted in a greater number of epidermis cells and very few oocytes (Fig. 2.1b).
I assigned cell types identities on the basis of the medusa atlas we published in the context of the collaboration. Cell types of the integrated medusa atlas could largely be re-assigned from the Chari et al., (2021) dataset by using the same markers (Fig. 2.1c and 2.2d).
Our integrated cell altas allowed the identification of three additional neural cell types and larger numbers of cells per cluster with respect to the dataset published in Chari et al., (2021), providing a more solid classification of medusa cell types. An additional cluster that we identified in our integrated medusa atlas is cluster 13. I initially classified this cluster as an additional i-cell type given the expression of Piwi and Vasa in these cells. However, analysis of the markers indicate that many more markers from other cell types are expressed by these cells which might ultimately represent a technical artefact (Fig. 2.1c). I do not yet understand their identity and for the moment I named them “mixed profile” and I do not consider those further in the chapter.

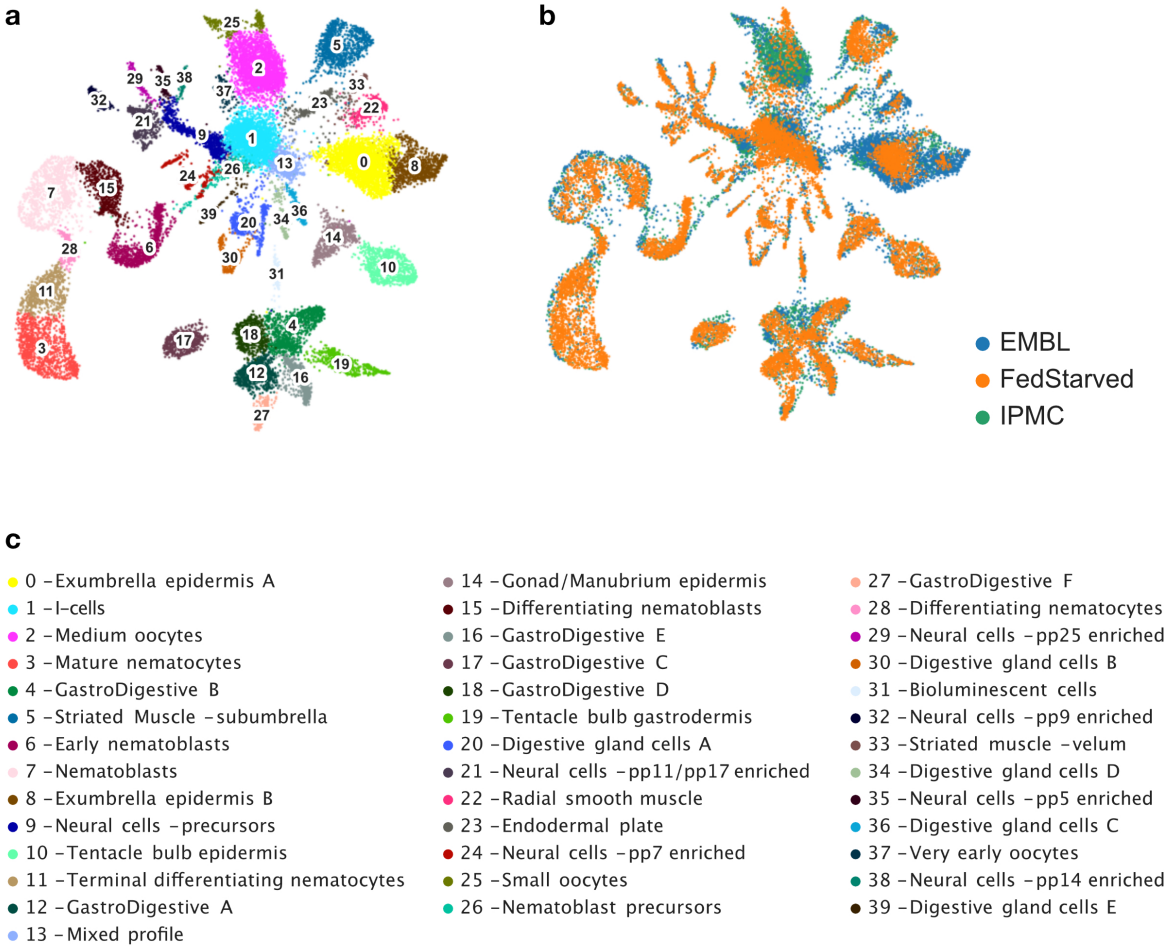


Figure 2.1 - Integrated Medusa Cell Atlas
(a) 2D-UMAP plot embedding according to the Leiden algorithm annotation of the clusters. (b) 2D-UMAP plot embedding of the cell atlas showing cells belonging to the 3 libraries following batch effect correction. Libraries or batches are labelled with custom names according to the platform where the experiments were performed respectively “IPMC" and “EMBL”. The label “FedStarved” refers to the “Fed and Starved” dataset published in Chari et al., 2021. (c) Cell type annotation .

I grouped the resulting 40 cell types into cell classes according to the analysis of expression patterns as previously described in Chari et al., (2021). Indeed, because of the large correspondence of the cell types across datasets, our classes are consistent with the ones assigned previously. In detail, I distinguished seven cell classes corresponding to the epidermis, the gastrodermis, the bioluminescent cells, the germ cells/i-cells, which includes i-cells and oocytes, the neural cells, the digestive gland cells and the nematocytes (Fig. 2.2b). The spatial distribution of the cell classes in the medusa is shown by the diagram in Figure 2.2a.
The integrated medusa cell atlas is presented in detail in the sections below. Each class is treated in a separate section along with the cell types included in each of them. Oocytes and i-cells belong to the same class of stem-cells/germ-cells although are treated into separate sections. The heatmap in Figure 2.2d shows the distribution across cell types and classes of the markers I used to assign identity and additional diagnostic markers that are discussed

below. The research article which includes our first medusa cell atlas is presented as an annex to this chapter.

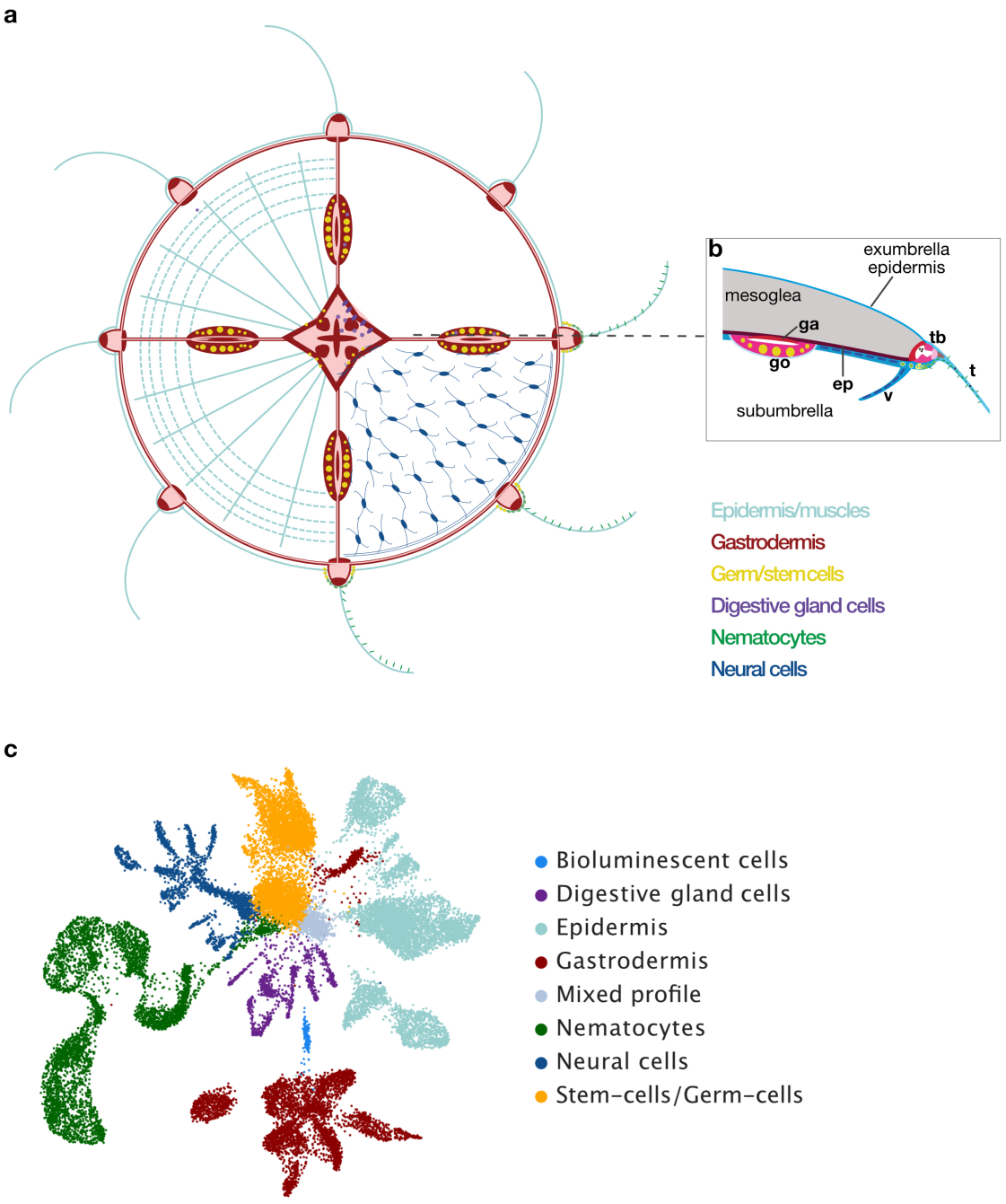


Figure 2.2 - Medusa Cell Type Classes
(a) Simplified diagram of adult medusa showing the spatial distribution of the broad cell classes (Adapted from Weissbourd et al., 2021). Bioluminescent cells are not shown. Colour code is according to (c). (b) Diagram showing a portion of the umbrella and the organisation of the layers of the epidermis and the gastrodermis. Abbreviations are: gastrodermis (ga), gonad (go), endodermal plate (ep), tentacle bulb (tb), tentacle (t), velum (v). (c) 2D-UMAP plot embedding showing the seven broad cell classes. (d) (on the following page) Heatmap of in situ candidates and diagnostic markers grouped by cell classes. Expression of genes is standardised between 0 and 1 using 'standard_scale = 'var'' option of the Scanpy function 'sc.pl.heatmap'.

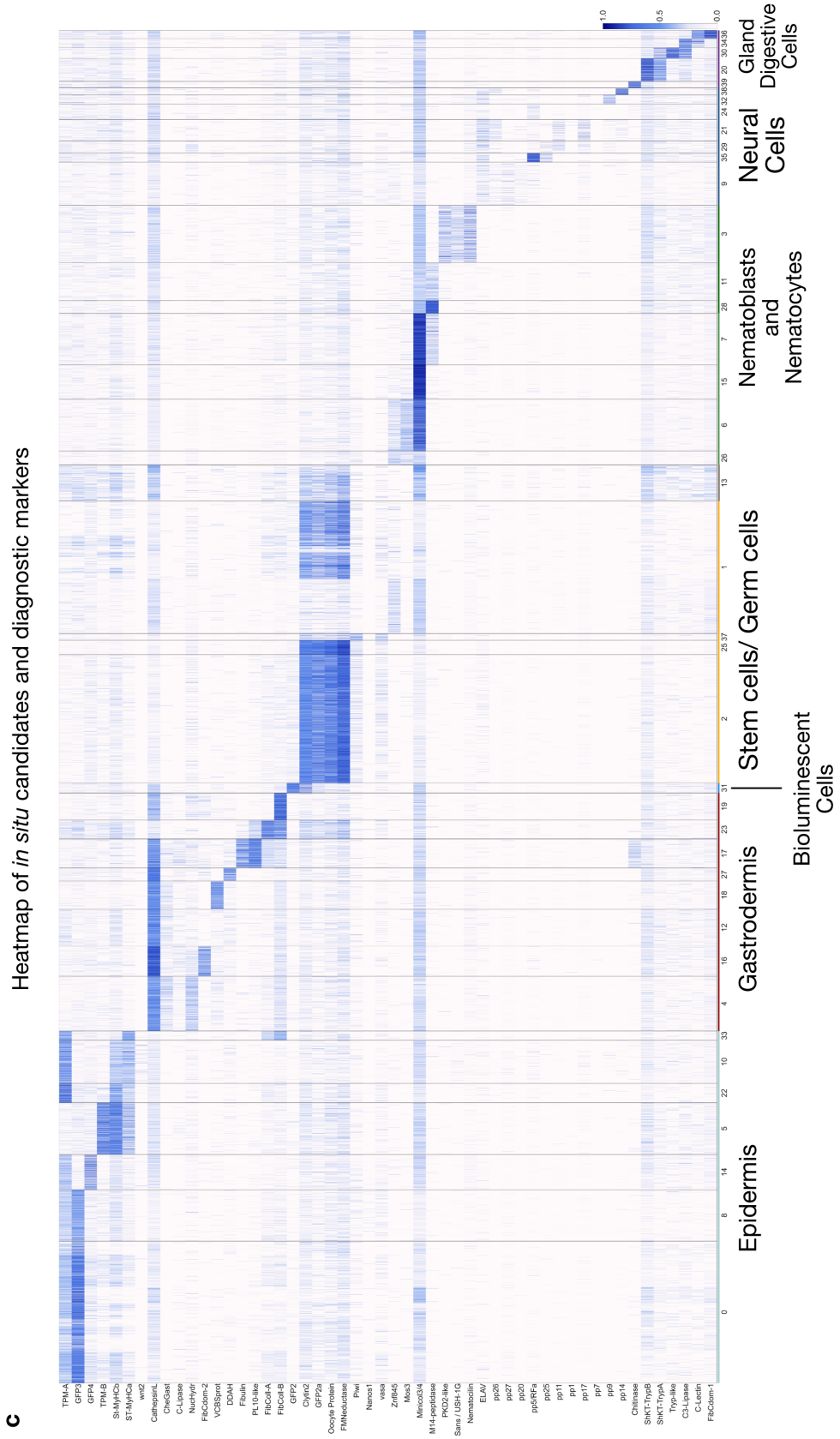


Figure 2.2d - Medusa Cell Classes

2.2.1.1 - Epitheliomuscular cells (epidermis)

The epidermis cell class covers a set of variations on the classical cnidarian epitheliomuscular cell type, which are characterised by having their apical poles facing externally. These include the medusa striated muscle type, which along with radially aligned smooth muscle lines the subumbrellar surface (Leclère & Röttinger, 2017).

Analysis of the gene expression in our integrated atlas revealed an additional cluster of exumbrella epidermis compared to the initial atlas (Chari et al., 2021). As mentioned above, the “EMBL " dataset was dominated by epidermal cells due to the physiological condition of the medusae prior to the encapsulation experiment. Indeed, this additional cluster is generated by the relatively large number of epidermal cells in the “EMBL” dataset. Another difference with the published atlas is that cells of the gonad epidermis and the manubrium epidermis are merged into a single cluster.

The exumbrella epidermis (cluster 0 and 8) is characterised by the expression of GFP3 (Fig. 2.2d; Fourrage et al., 2014) and forms a monolayer of cells overlaying the outer part of the bell (Fig. 2.2b).

We assigned cluster 22 as radial smooth muscle based on strong expression of the marker Tropomyosin-A (TPMA; Fig. 2.3b and c). In *Clytia* medusa the smooth muscle fibres lie in the subumbrella and extend radially from the manubrium to the bell margin (Fig. 2.2a). The muscle fibres of these cells are on the external side of the cells of the striated muscle layer (Fig. 2.4). The identity of the striated muscle type is assigned based on the expression of Tropomyosin-B, ST-MyHCa and ST-MyHCb (Myosin heavy chain - striated type I and II; Fig. 2.3b and c). These muscle fibres are oriented circularly in the subumbrella, organised in a band at the periphery of the bell (Fig. 2.2a) and are responsible for the folding of the bell and swimming (Leclère & Röttinger, 2017). We assigned cluster 33 as a second striated muscle type of the velum, characterised by the expression of ST-MyHCa, ST-MyHCb and Peroxidase (Fig. 2.3b and c).

Each of the medusa organs (i.e. manubrium, gonads and tentacle bulbs), as well as the bell (umbrella) are covered by an outer epidermal layer. We deduced that cluster 14 includes gonad and manubrium epidermis, both characterised by the expression of GFP4 (Fig. 2.2d; Fourrage et al., 2014). Tentacle bulb epidermis was assigned based on the expression of Wnt2 (Fig. 2.2c; Condamine et al., 2019) among others such as TPMA, ST-MyHCa and ST-MyHCb (Fig. 2.3b and c).

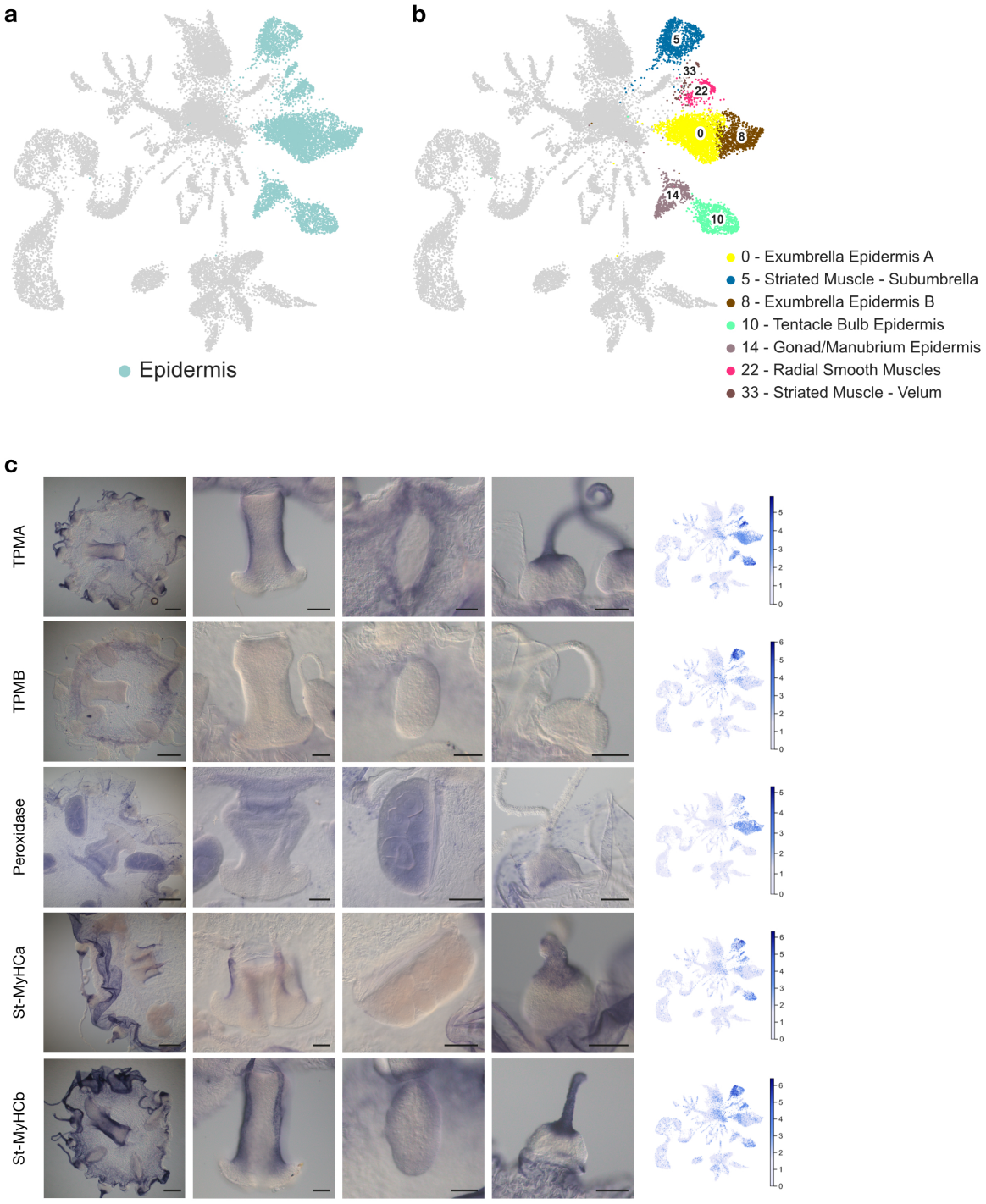


Figure 2.3 - Validation of epidermis and muscle cell types
(a) 2D-UMAP plot coloured by epidermis cell class annotation. (b) 2D-UMAP plot coloured by epidermis cell type annotation, (c) *In situ* hybridization of epidermis marker genes of 10-14 days-old medusa edited from Chari et al., 2021, Supplementary Figures. Images from left to right: whole medusa, manubrium, gonad, tentacle bulb. Right column: 2D-UMAP plots representing the expression for each of the marker genes on the cell atlas. Scale bars in the whole medusa images represent 200 µm; in the manubrium, gonad and tentacle bulb images they represent 100 µm.

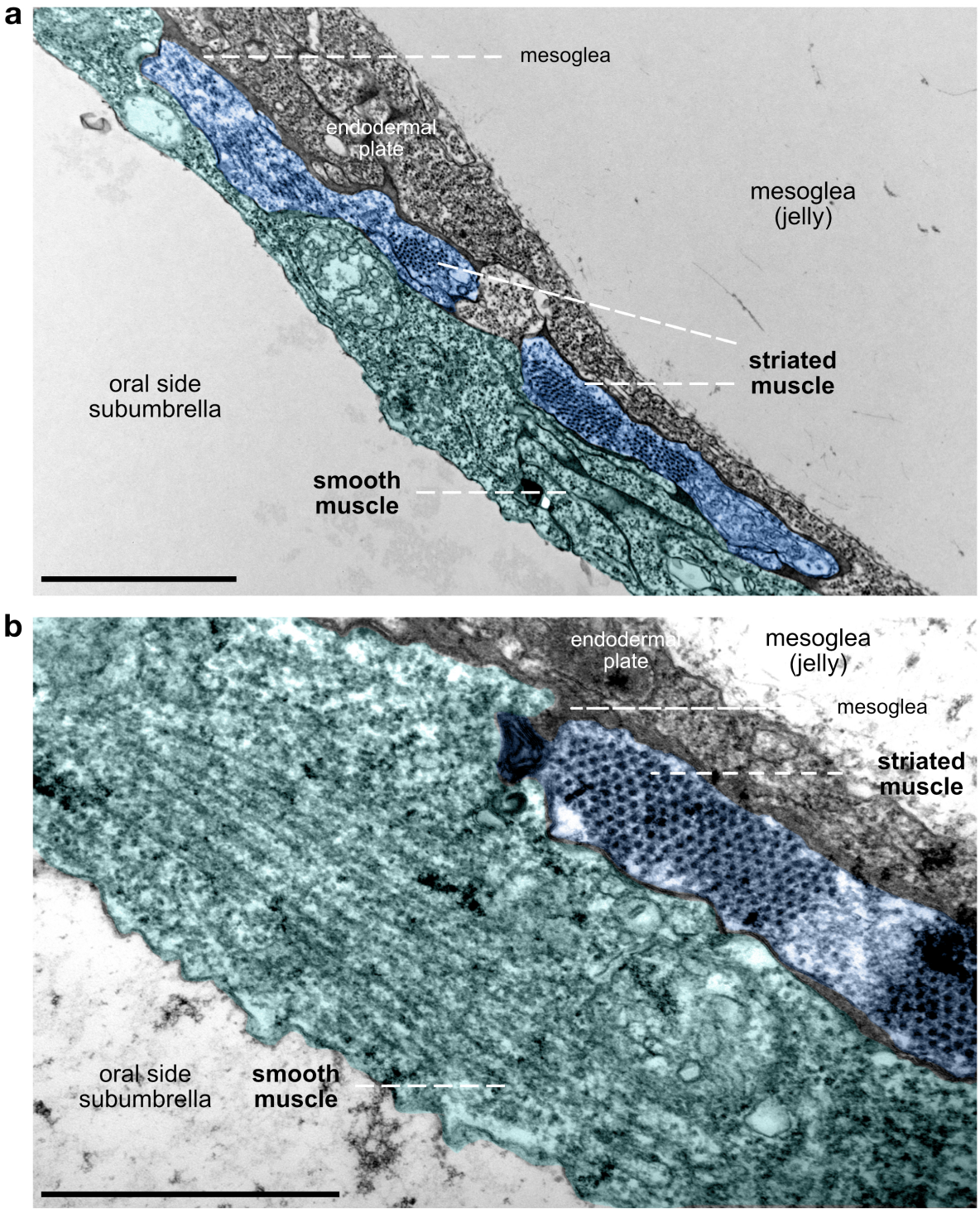


Figure 2.4 - Ultrastructure of the subumbrella epidermis

Electron micrographs of sections through the subumbrella of a *Clytia* medusa. (a) Organisation of smooth muscle and striated muscle in the subumbrella of the medusa. Smooth muscle layer (cyan) overlies the striated muscle (blue). (b) Higher magnification of another section of the subumbrella showing distinct smooth muscle fibres. Scale bar in (a) represents 2µm. A layer of mesoglea divides the muscle cell from the endodermal plate cells. Scale bar in (b) represents 1µm.

2.2.1.2 - Gastro-digestive cells (gastrodermis)

The Gastrodermis is the second typical epithelium of cnidarian (Steinmetz, 2019) and in the medusa of *Clytia* also represents the tissue committed to digestion of food. The manubrium, the gonads and the tentacle bulbs exhibit gastro-digestive cell types likely involved in distinct functions not yet characterised.

As in Chari et al., (2021), in our integrated atlas we identified six gastrodermis subtypes likely involved predominantly in digestion (GastroDigestive types A, B, C, D, E, F; Fig.2.5a and b which I refer throughout as GD-A, GD-B GD-C, GD-E and GD-F) as well as two types likely to be heavily involved in extracellular matrix formation (Endodermal Plate, Tentacle Bulb Gastrodermis; Fig. 2.5a and b).

Gastrodermis cells are characterised by the expression of the enzyme CathepsinL, involved in intracellular digestion, detected throughout the eight clusters (Fig. 2.5c). Single cell data allowed the characterisation of previously undescribed subtypes of gastrodermis in the medusa by analysis of the expression patterns of cluster-specific candidates (Fig. 2.5b; Chari et al., 2021). We could not identify any specific expressed marker genes for two of the subtypes, GD-A and B, due to the largely shared gene signature with the other subtypes. However, CheGast (*Clytia* specific gastrodermal marker), an otherwise uncharacterised gene detected in GD-A, B and D, is expressed in the distal segment of the manubrium (Fig. 2.5c), suggesting that a possible specialised compartmentalisation exists.

Analysis of *in situ* expression patterns of the extracellular protease BP10-like (BP10-like Zinc metalloprotease) and the ECM glycoprotein Fibulin (Fibulin family protein) allowed the localisation of GD-C to the region of contact between the canals and the three organs (Fig. 2.5b and c; Fig. 2.6c and d). Furthermore, BP10-like is also expressed in the endodermal plate cell type, consistent with a structural function in the formation and/or remodelling of the mesoglea in all these cell types (Fig. 2.5c, 2D-UMAP plot; Fig. 2.6a and b).

The GD-F cells are characterised by the expression of DDAH (dimethylarginine dimethylaminohydrolase), which is involved in Nitric Oxide signalling and in vascular functions. The localisation of this cell type at the base of the manubrium is complementary to BP10-like, corresponding to cells that are on the distal side of the expanded gastric region of the radial canal (Fig. 2.5c). Similarly, DDAH is expressed in a particular region of the tentacle bulb gastrodermis likely corresponding to a specialised group of cells of which the function remains to be determined (Fig. 2.5c).

In situ hybridisation expression patterns for FibCdom-2 (Fibrinogen-C domain protein 2) and VCBSprot (VCBS repeat-containing protein) show a diffuse signal in the gastrodermis of gonads manubrium and tentacle bulb (Fig. 2.5c). The expression of these two genes shows high specificity for GD-D and E respectively (Fig. 2.5c, 2D-UMAP plots). Analysis of the molecular signature revealed that GD-D contains a module associated with cell-cell junctions (Chari et al.,2021). However, the exact role of these two cell types and whether they are part of a distinct gastrodermis compartment is not clear and would require further investigation.

We assigned two further gastrodermis types based on the expression of two fibrillar collagen genes (FibColl-A and FibColl-B; Fig.2.5b and c) as endodermal plate and tentacle bulb gastrodermis. The Endodermal plate cell type is organised into a thin pseudo bilayer within the subumbrella and is responsible for producing the collagen-rich jelly of the bell (Fig.2.6a and b). The tentacle bulb gastrodermis is characterised by the expression of the same genes, although the expression FibColl-B in a distal portion of the tentacle bulb gastrodermis, which is not interested by the expression of FibColl-A is diagnostic for this cell type (Fig.2.5b and c).

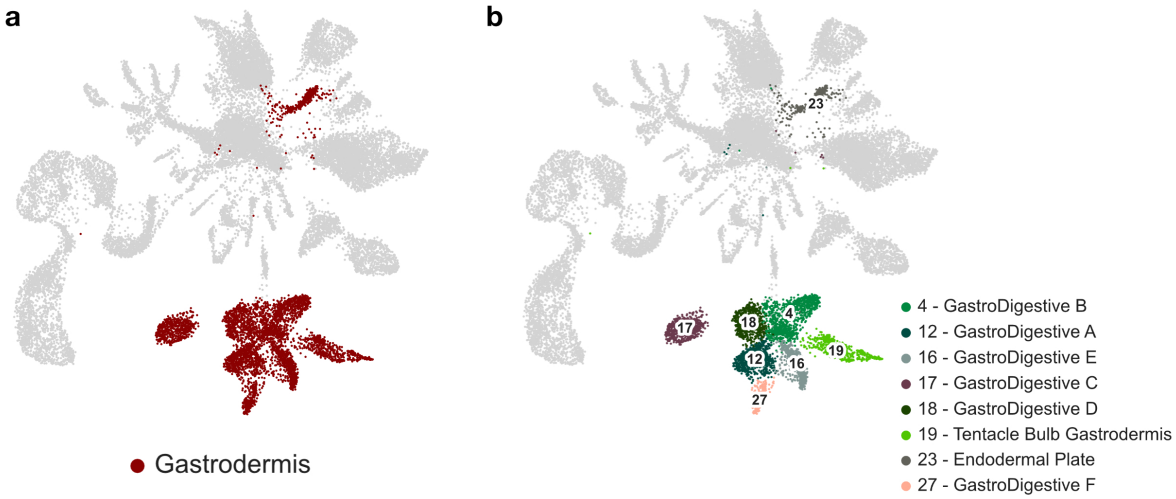


Figure 2.5 - Validation of Gastrodermis cell types
(a) 2D-UMAP plot coloured by gastrodermis cell class annotation. (b) 2D-UMAP plot coloured by gastrodermis cell types annotation. (c) *In situ* hybridization of gastrodermis marker genes of 10-14 day-old medusae edited from Chari et al., 2021, Supplementary Figures (on the following page). Images from left to right: whole medusa, manubrium, gonad, tentacle bulb. Right column: 2D-UMAP plots representing the expression for each of the marker genes on the cell atlas. Scale bars in the whole medusa images represent 200 µm; in the manubrium, gonad and tentacle bulb images they represent 100 µm.

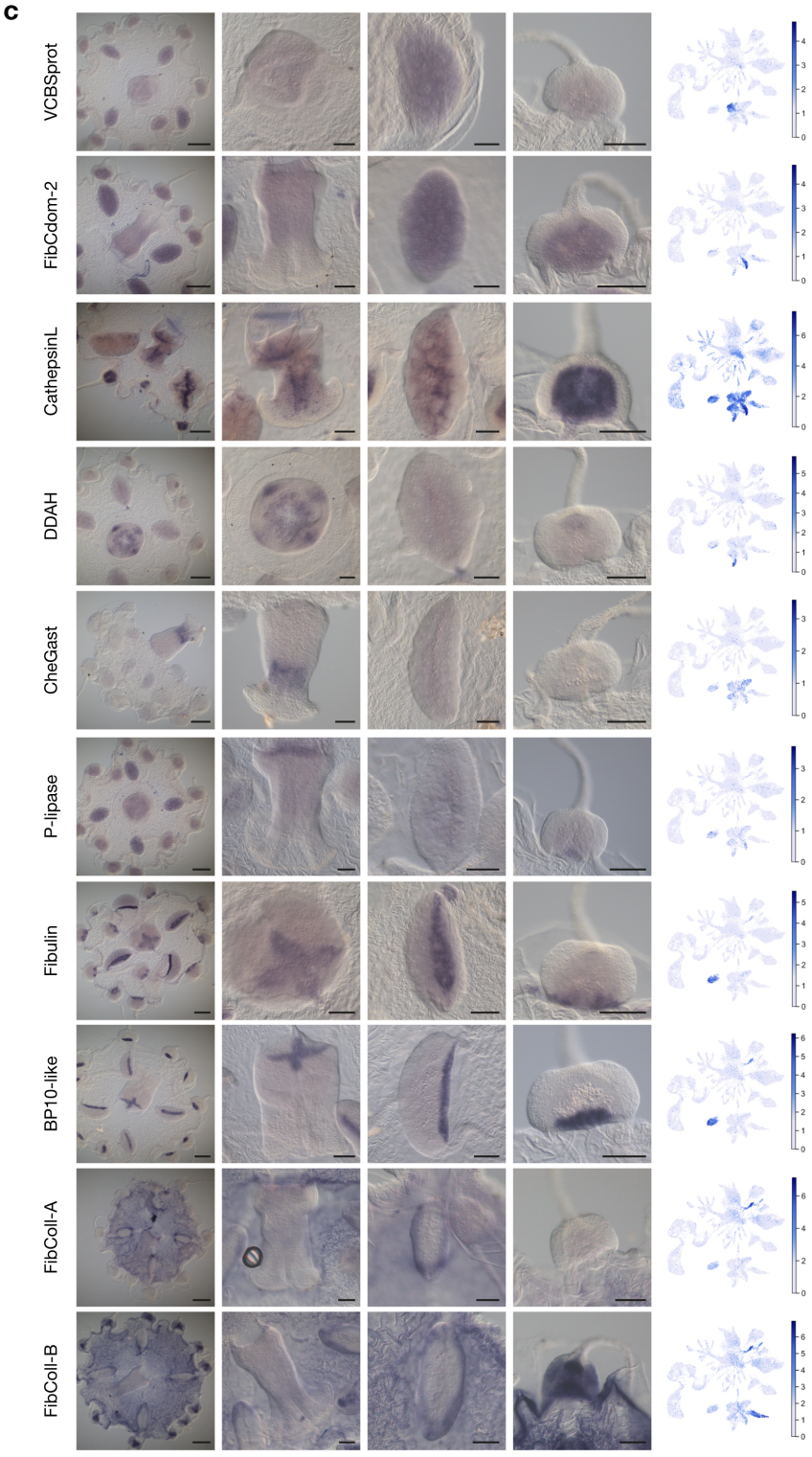


Figure 2.5c - Validation of Gastrodermis cell types

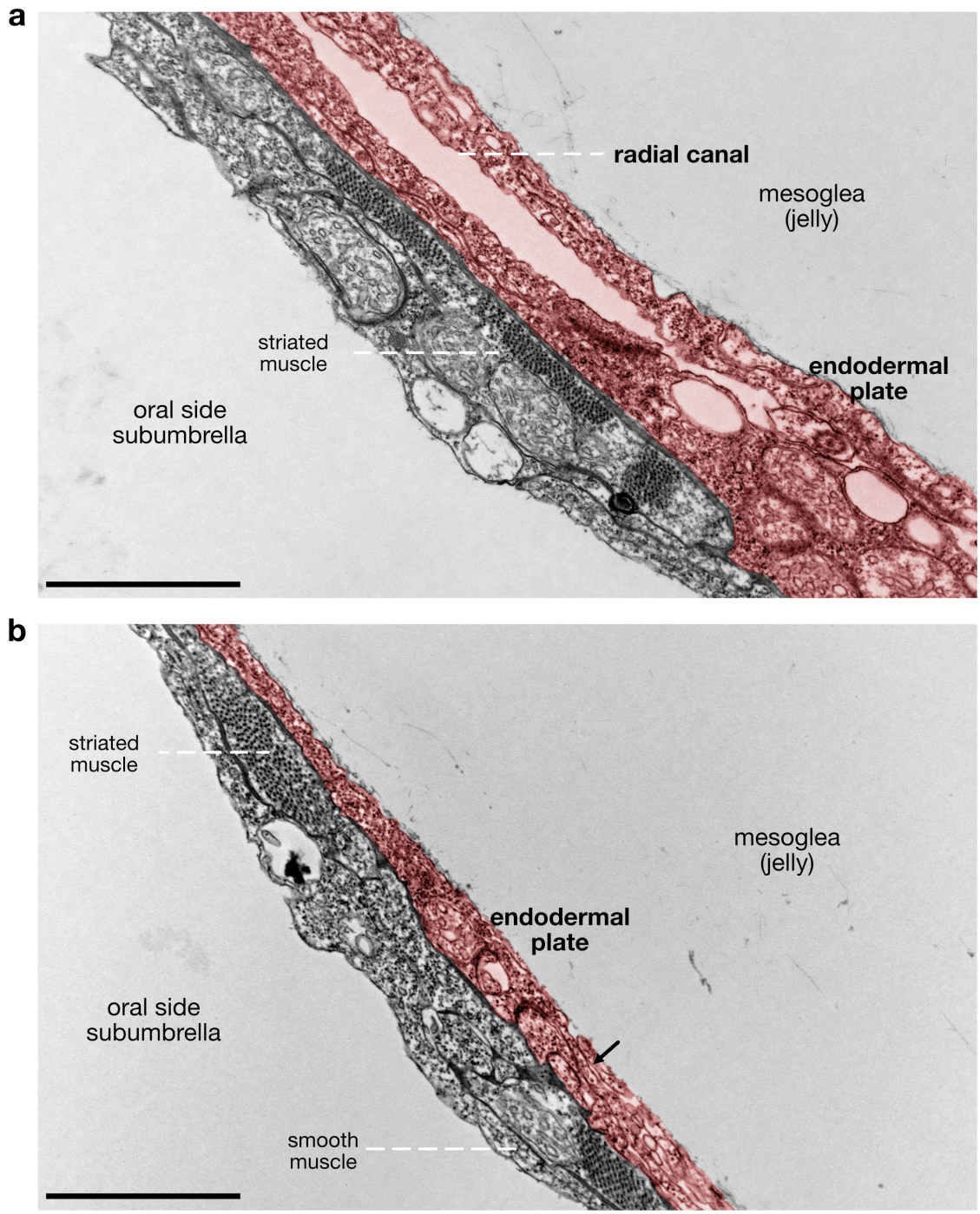


Figure 2.6 - Ultrastructure of gastrodermis cell types
Micrographs of sections through the subumbrella of the medusa of *Clytia.* (a) Organisation of the endodermal plate (red) around the radial canal. (b) organisation of the endodermal plate (red) in another section of the subumbrella where the pseudo bilayer is visible (arrow). Scale bars represent 2µm.

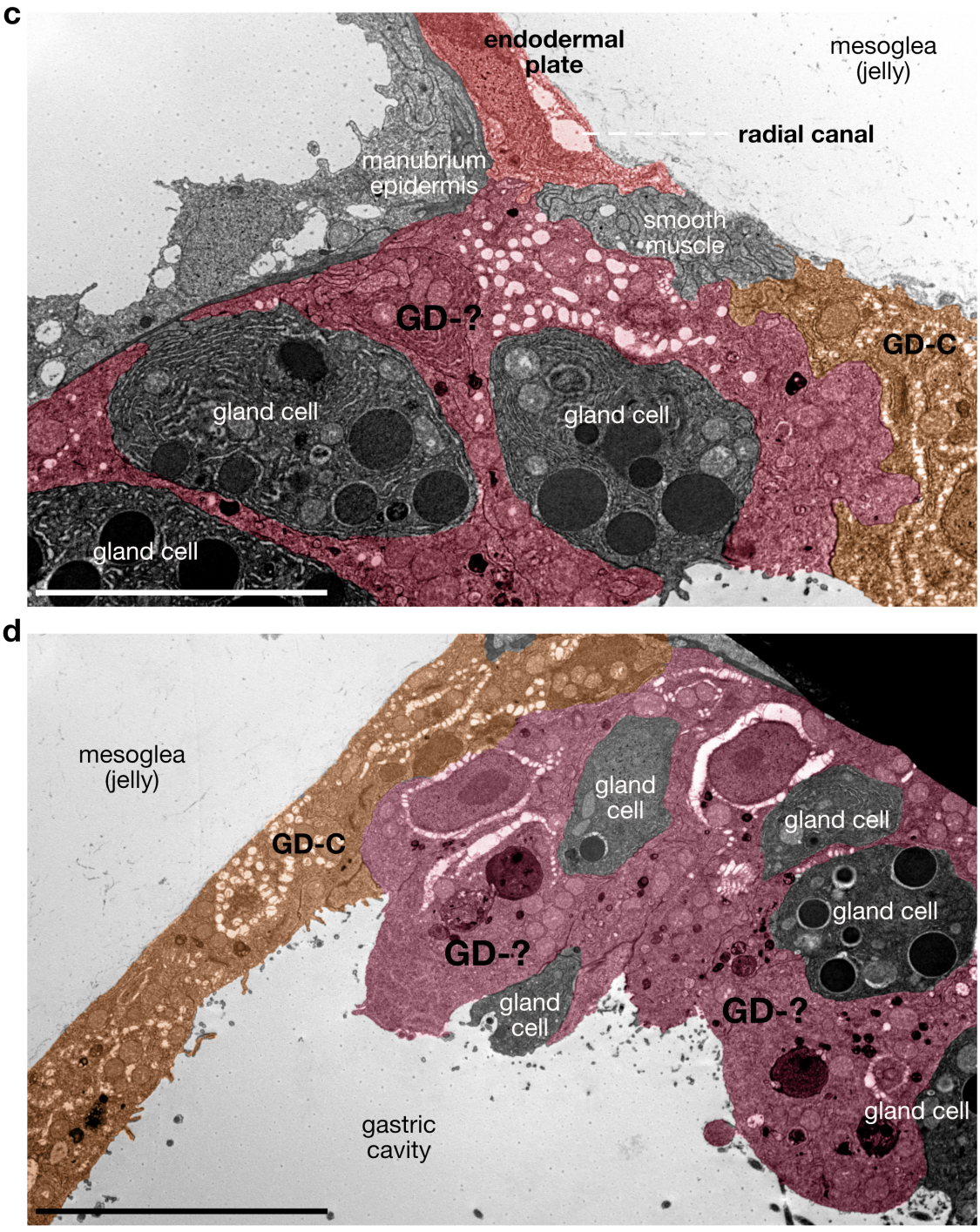


Figure 2.6 c-d - Ultrastructure of gastrodermis cell types
Micrographs of sections through the manubrium of the medusa of *Clytia.* (a) Organisation of the gastrodermis in the portion of the manubrium linked to a radial canal. (b) Section through the manubrium with gastrodermis cell types facing the gastric cavity. Endodermal plate (red), unidentified gastrodermis type (darker red); Basal gastrodermis or GD-C (orange) Scale bars represent 10µm.

2.2.1.3 - Bioluminescent Cells

Bioluminescence is a process that results in the generation of light through a biochemical reaction in which a photoprotein catalyses the oxidation of a substrate. In *Clytia* Bioluminescence has been already described along with the characterisation of genes for the calcium-activated photoprotein Clytin, which belongs to the *Aequorin* family of photoproteins, and four fluorescent proteins (GFPs) that can convert the blue Clytin bioluminescence into green flashes. The different *Clytin* and *GFP* genes are expressed in distinct tissues in medusa and planula (Fourrage et al., 2014; Leclère et al., 2019).

To characterise medusa cell types we exploited the different *GFP* and *Clytin* gene expressions as diagnostic markers for exumbrella epidermis (GFP3; Fig. 2.2d), gonad and manubrium epidermis (GFP4; Fig. 2.2d) and oocytes (GFP2a and Clytin 2; Fig. 2.2d).

In our cell atlas we assigned Cluster 31 as GFP2 positive bioluminescent cells (Fig. 2.7a and b) which are scattered in the tentacle and tentacle bulbs of the medusa (Fig. 2.7c). These cells are GFP positive and so detectable by fluorescent microscopy following excitation with blue light as shown in Fig.2.7c and previously described by Fourrage et al., (2014).

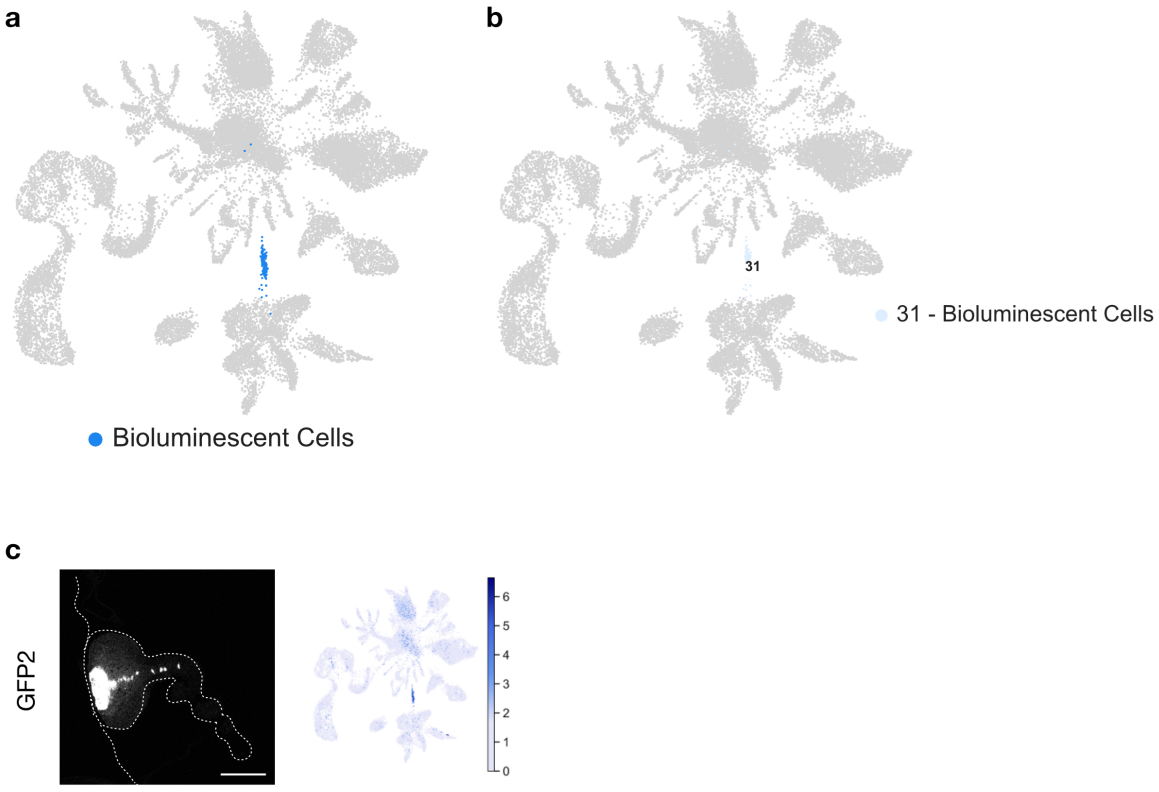


Fig.2.7 - Bioluminescent cells
(a) 2D-UMAP plot coloured by bioluminescent cell class annotation. (b) 2D-UMAP plot coloured by bioluminescent cell type annotation. (c) Confocal microscopy image of GFP2 in cells in the tentacle and tentacle bulb edited from Chari et al., 2021, Supplementary Figures. Scale bar represents 100 µm.

2.2.1.4 - Oocytes

As described above, the three datasets that I combined to generate the integrated medusa cell atlas were generated using dissociated cells from female medusae.

The medusa is the sexually reproductive form of *Clytia* and spawning of female and male jellyfish is triggered every morning by light (Houliston et al., 2010; Lechable et al., 2020). Immediately before spawning, the female gonads contain oocytes at all stages of development (Fig. 2.8a and b).

Oocyte meiotic maturation, followed by spawning, is regulated by Maturation Inducing Hormones (MIH) which are released from the gonad upon light signals (Freeman & Ridgway, 1988; Ikegami et al., 1978). Previous analysis of transcriptomic data from dissected gonads revealed the expression of several opsin genes, among which Opsin9 was the most highly expressed. Opsin9 and PP4, an MIH (maturation inducing hormone) precursor of *Clytia*, were found to be co-expressed in the same neural-type cells in the medusa gonad ectoderm. Knockout of the Opsin9 gene led to failure of both oocyte maturation and spawning upon light stimulation, demonstrating that Opsin9 mediates light induced maturation and spawning (Quiroga Artigas et al., 2018). Additionally, this study generated transcriptome data for dissected mid and late stages of oocytes which were used to assign oocyte identities in the atlas.

In my integrated medusa cell atlas I assigned three clusters as oocytes, which we designated as very early, small and medium oocytes (Fig.2.8a and b). As mentioned above, the majority of cells populating the oocyte clusters derive from the “IPMC” dataset. However, fully grown oocytes (around 200 µm in diameter) are likely absent from our data since in all cases the cell suspension was filtered through a 40µm cell strainer which is smaller than the size of those cells (see Methods). Medium oocytes are characterised by the expression of previously described marker genes such as *Clytin2,* also expressed in bioluminescent cells, and *GFP2a* (Fourrage et al., 2014) among others, while very early oocytes express meiotic recombination proteins such as Spo11 and Scyp1 which are part of the synaptonemal complex (Fig.2.8c; Munro et al., 2022).

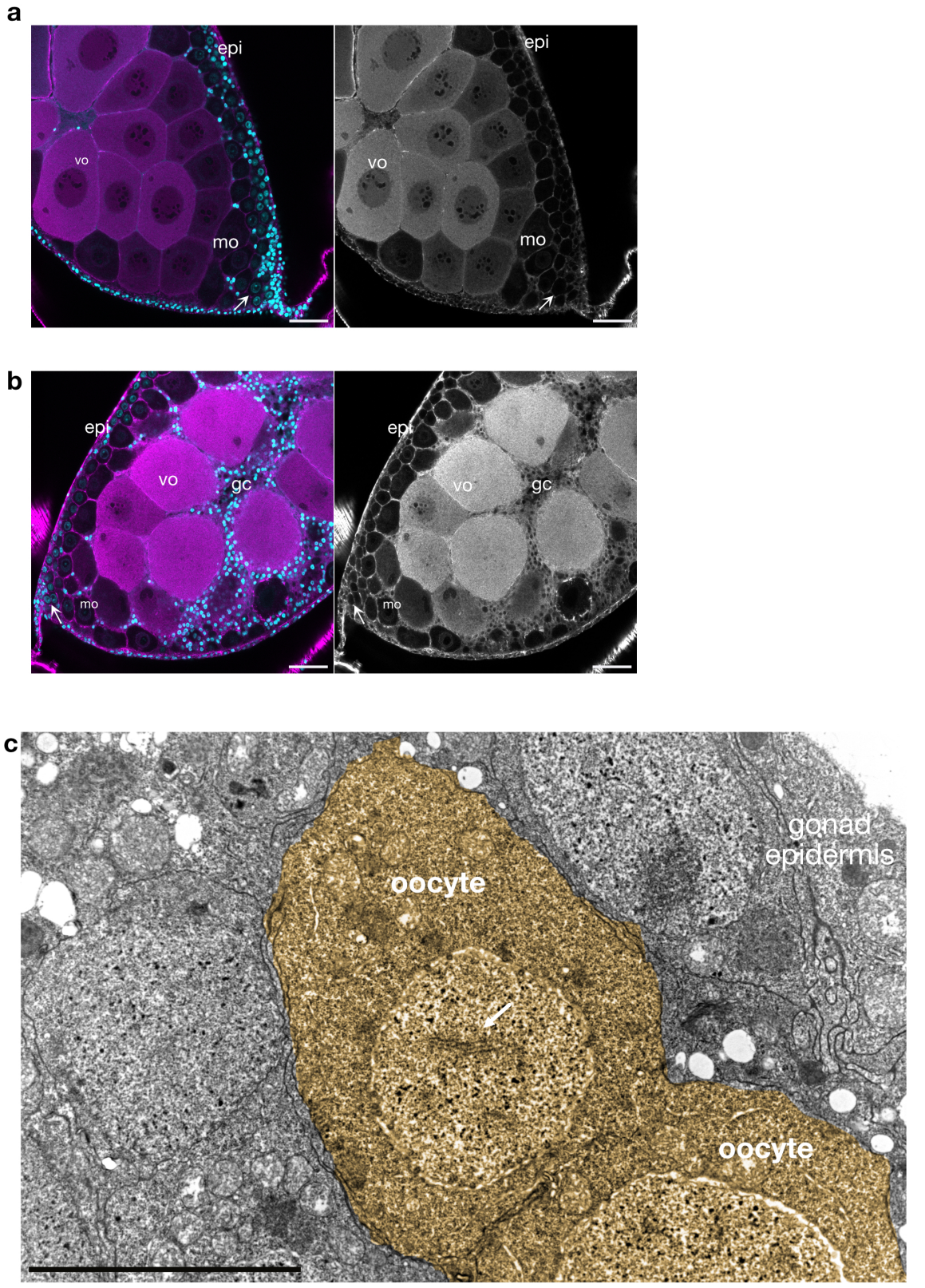


Figure 2.8 - Oocytes
(a, b) Confocal sections through gonads from 2 weeks-old medusae with cell morphology revealed by phalloidin staining of cell boundaries in magenta and grey. The first panel of a and b shows co-staining of nuclei with Hoechst (cyan). Each panel shows co-existence of several stages of oocyte development

in two different gonads: epidermis (epi), vitellogenic oocytes (vo), medium oocytes (mo) and early oocytes (arrows). Very early stage oocytes can be distinguished near the arrows (Munro et al., 2022). Scale Bars represent 50 µm. (c) Electron Micrograph showing a section through the gonad of a 10 day-old medusa. Very early oocytes are highlighted in yellow. The arrow indicates a synaptonemal complex. Scale bar represents 5µm

2.2.1.5 - Interstitial cells (i-cells)

Interstitial cells, or i-cells, are the hydrozoan stem cell population that gives rise to germ cells as well as somatic derivatives. In *Hydra* they give rise to neural cells, nematocysts, and gland cells (Bode, 1996; Watanabe et al., 2009; see Chapter 1).

In my integrated cell atlas I identified one cluster of i-cells, cluster 1. (Fig. 2.9 a and b). I assigned the identity on the basis of the diagnostic expression of previously described i-cell marker genes, such as *Piwi*, *Vasa* and *Nanos1* (Leclère et al., 2012; Fig. 2.9 c and d).

In medusa, the i-cells are predominantly found in the ectoderm and are most abundant in the gonads and tentacle bulbs. Populations of i-cells are present in the proximal portion of the tentacle bulb ectoderm and also at the manubrium base as shown by the *in situ* hybridisation for Nanos1 in Fig.2.9c. In the medusa atlas Nanos1 is detected only in a small number of cells (Fig. 2.9c, atlas plot). The expression of Piwi and Vasa however, argues in favour of our annotation (Fig. 2.9d).

Pseudotime analysis including i-cells, neural cells and nematocytes showed that single cell data support the developmental origin of neural cells and nematocytes from the i-Cell population (Chari et al., 2021), as is described in *Hydra* (Bode, 1996).
A question that still remains to be resolved is whether the digestive gland cells of *Clytia* are also derived from i-cells according to the model in *Hydra* (Bode et al., 1987). We addressed this question in the published analyses of the initial medusa atlas. We were not able to identify a direct link between i-cells and Digestive Gland Cells (Chari et al., 2021). One possibility is that digestive gland cells are generated from i-cells very early during medusa development and so the link is not detectable in more advanced stages. Under this scenario digestive gland cells would go through self renewal following differentiation leading to the maintenance of a very distinct transcriptional profile (Siebert et al., 2019). Alternatively, digestive gland cells derive from a different precursor which has no i-cell-like features. At present, the available medusa datasets still do not provide potential candidate regulatory genes for functional analysis which would allow these hypotheses to be tested. The increased number of cells in the integrated atlas would potentially enable the identification of putative candidates by repeating pseudotime analyses on the basis of a larger set of information. However, computational constraints are caused by the batch effect correction and this analysis requires further refinement (see below in Methods).

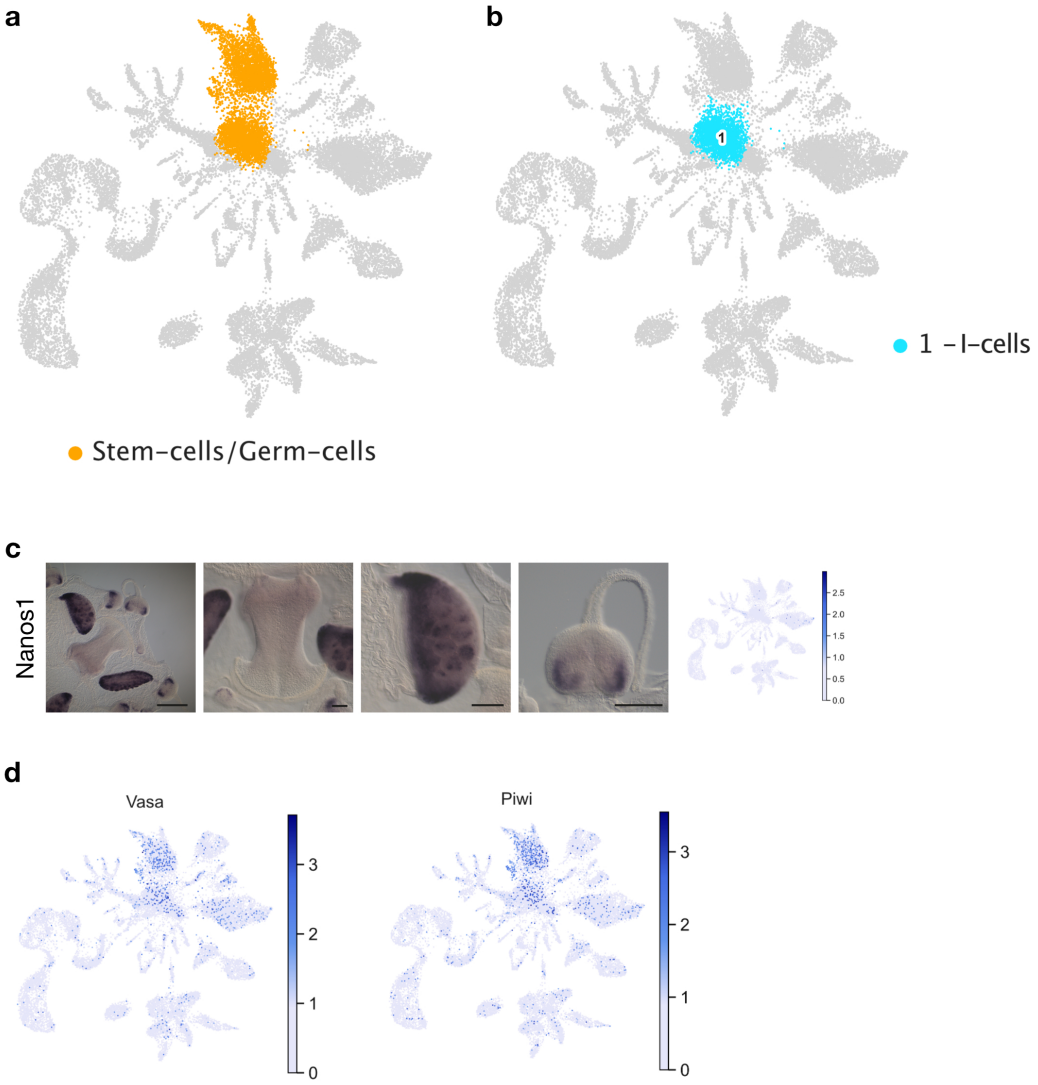


Fig.2.9 - Validation of the interstitial cells (i-cells)
(a) 2D-UMAP plot coloured by i-cells annotation. (b) 2D-UMAP plot coloured by i-cell type annotation (c) *In situ* hybridization of i-cells marker gene, Nanos1, of 10-14 days-old medusa edited from Chari et al., 2021, Supplementary Figures. Images from left to right: whole medusa, manubrium, gonad, tentacle bulb. Right column: 2D-UMAP plots representing the expression for Nanos1 on the cell atlas. Scale bars in the whole medusa images represent 200 µm; in the manubrium, gonad and tentacle bulb images they represent 100 µm. (d) Expression of Vasa and Piwi on the cell atlas.

2.2.1.6 - Nematocytes

The cnidarian stinging cells, known as nematocytes or cnidocytes, are a unique feature of the phylum.
Nematocytes are highly concentrated in the tentacles and in the tentacle bulbs of the medusa of *Clytia* as well being distributed widely across other sites including the rim of the manubrium. The medusa tentacular system is particularly useful for studying nematocyte production as it can be structurally defined as a cellular conveyor belt in which cellular proliferation and differentiation occur in an ordered fashion along the tentacle (Condamine et al., 2019). Proliferation of precursor cells takes place in the tentacle bulb epidermis (Condamine et al., 2019; Denker et al., 2008). The proximal pole of the tentacle bulb epidermis includes a pool of i-cells expressing *Nanos1* (see above) and *Piwi* (Denker et al., 2008). Additional genes characterising nematogenesis are expressed in a spatially ordered progression in the intermediate and the distal part of the tentacle bulb epidermis, for example Minicollagen 3/4 and Nowa which encode for specialised proteins responsible for the formation of the nematocyte capsule (Condamine et al., 2019).

Analyses of the single cell data of *Clytia* medusae support the previous knowledge regarding nematocyte development and bring to light additional molecular information (Fig.2.10a and b; Chari et al., 2021). We identified two distinct transcriptional programs within this cell class. Using previously described and novel marker genes, we could show that these correspond to two largely separate phases of nematogenesis. To distinguish these we adopted the terms “nematoblasts” for the clusters of the first phase and “nematocytes” for clusters of the second. The initial phase, nematoblast development, notably includes production of the nematocyst (i.e. the highly specialised stinging capsule; Fig. 2.11a). This phase is characterised by early expression of previously described genes *Znf845* and *Mos3* (Lapébie et al., 2014) not originally associated with this cell class, as well as known nematocyst genes such as *Minicollagen 3/4* (Condamine et al., 2019). The second phase of nematogenesis is characterised by the expression of proteins relating to mechanosensory function, including Nematocilin, a component of the elaborated structure around the nematocyte cilium, the nematocil, that triggers the capsule discharge upon stimulation (Fig. 2.11b; Balasubramanian et al., 2012). The nematocil is formed following the placement of the mature nematocyte in the epidermis of the tentacle (Fig.2.11b). Nematocytes also strongly express markers orthologous to protein components of vertebrate vibration-sensitive “hair cells'' such as Harmonin, Whirlin and Sans-USH-1G (Chari et al., 2021; Gillespie & Müller, 2009; Fig. 2.10c). In contrast they show no detectable expression of nematoblast phase gene. The mapping of the two sets of nematogenesis genes onto the cell atlas highlighted the clear successive stages of nematocyte development (Fig. 2.10c atlas plots). To establish the spatial ordering of these stages along the tentacle I compared *in situ* hybridisation patterns along the tentacle bulbs and tentacles of the medusa (Chari et al., 2021; Fig.2.10c). I selected candidate gene markers for each of the nematoblast and nematocyte cell clusters using the cell atlas data and performed *in situ* hybridisation. I found that the transition between the two phases corresponded precisely to the junction between the tentacle bulb and the tentacle, and also that mature nematocytes are organised in two rows along the axis of the tentacles (Chari et

al., 2021; Fig.2.10c). The two distinct phases are linked by the expression of M14-peptidase, a cluster specific marker (cluster 28, integrated medusa atlas, Fig. 2.10b; cluster 10 in Chari et al., 2021), previously overlooked. Indeed, the large availability of the cells for this cluster and the expression of M14-peptidase in differentiating nematocytes at the base of the tentacle bulb were enlightening and enabled the two transcriptional programs to be identified (Chari et al., 2021). For instance, in the *Hydra* cell type atlas this link is not obvious as clusters corresponding to the two phases are disconnected (Siebert et al., 2019).

The succession during nematogenesis of two distinct transcriptional programs within the cell class are demonstrated by pseudotime analysis (Chari et al., 2021). The nematocyte developmental trajectory described in Chari is thus corroborated by our integrated medusa atlas (Fig 2.10a, b and c, 2D-UMAP plots). This represents a particularly striking example of a drastic switch in transcriptional regulation during cell type differentiation.

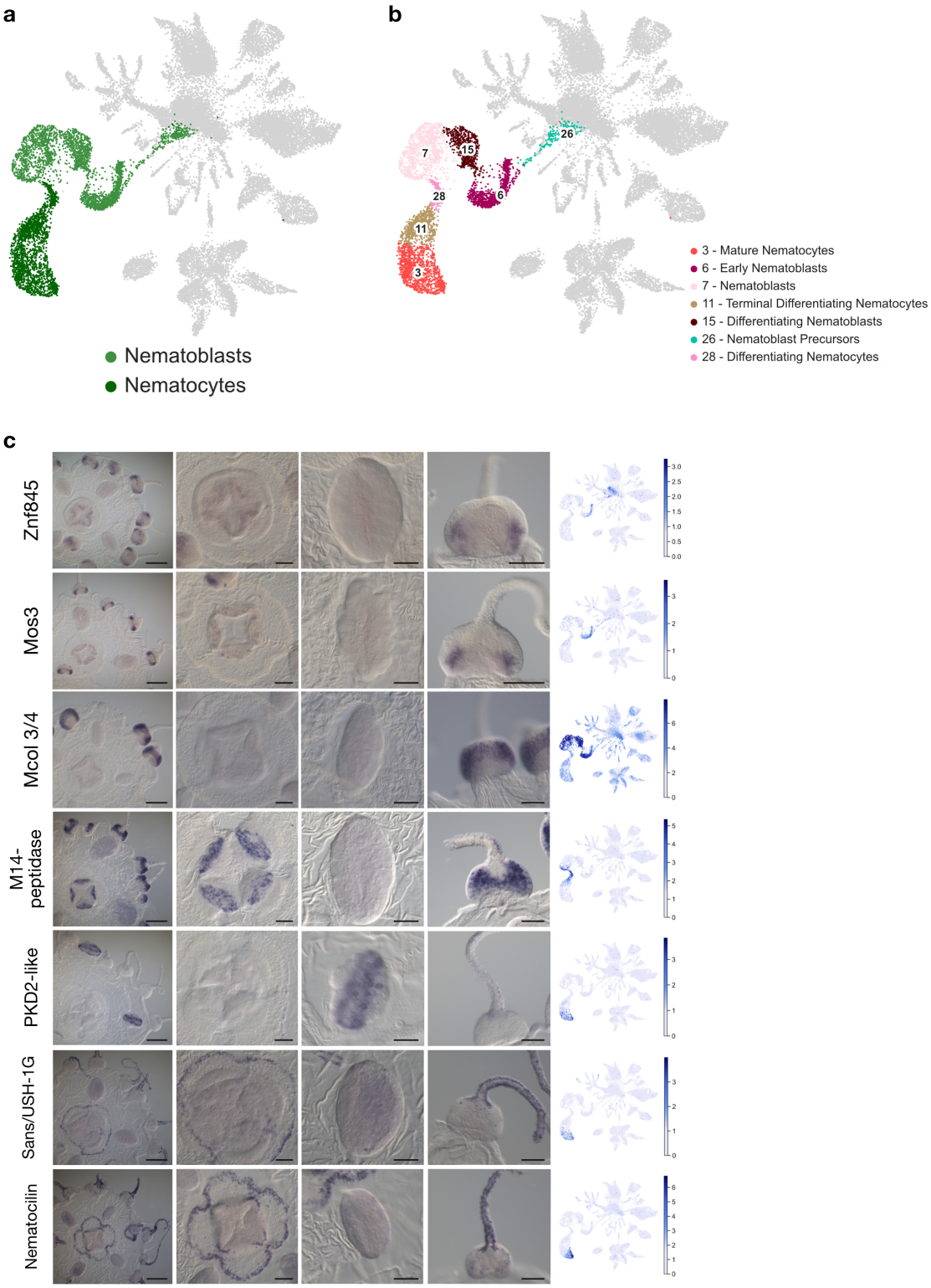


Figure 2.10 - Validation of nematocytes

(a) 2D-UMAP plot coloured by nematocyte class annotation. (b) 2D-UMAP plot coloured by nematocyte cell type annotations. (c) *In situ* hybridization of nematocytes marker genes of 10-14 days-old medusa edited from Chari et al., 2021, Supplementary Figures. Images from left to right: whole medusa, manubrium, gonad, tentacle bulb. Right column: 2D-UMAP plots representing the expression for each of the marker genes. Scale bars in the whole medusa images represent 200 µm; in the manubrium, gonad and tentacle bulb images they represent 100 µm.

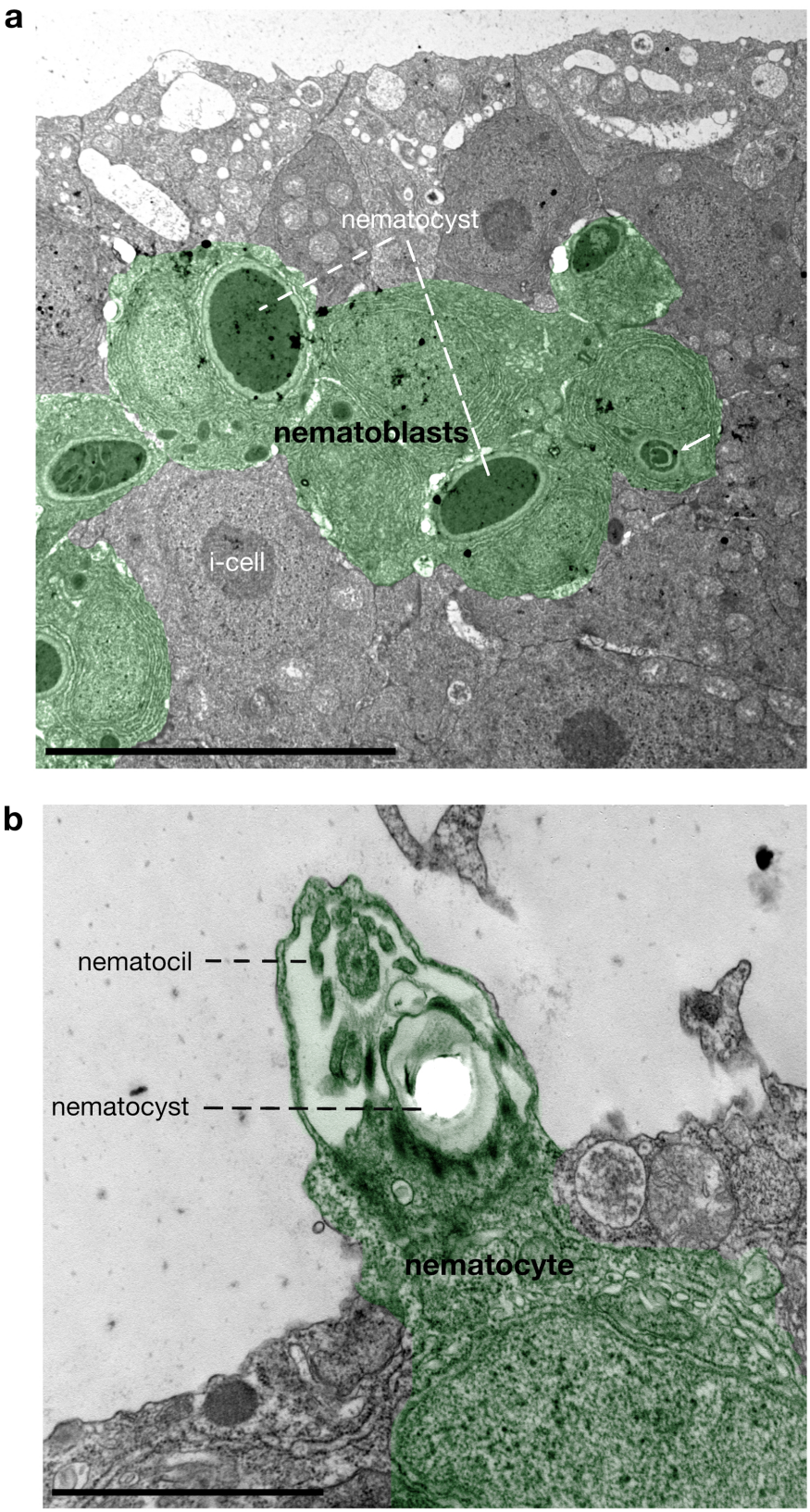


Figure 2.11 - Developmental stages of nematocytes
(a) Micrograph of a section through the epidermis of the tentacle bulb of the medusa of *Clytia.* Nematoblasts (green) show high content of Golgi and nematocysts (or stinging capsules) at different stages of development. The arrow indicates a nematocyst in its initial phase of formation. (b) Micrograph

of a section through the tentacle of the medusa. The mature nematocyte (green) exhibits a mature nematocysts and a typical nematocil on the apical side. Scale bar in (a) represents 10 µm. Scale bar in (b) represents 2µm.

2.2.1.7 - Neural Cells

In our integrated medusa atlas we identified seven neural cell types expressing genes of the ELAV family, widely used as neural markers (Nakanishi et al., 2012). These included one cluster of putative neural precursors and six putative differentiated neural cells enriched in various neuropeptide precursors (Fig. 2.12 a and b).

During the analysis of the initial medusa atlas, we validated several putative neural subpopulations of *Clytia* medusa based on reclustering analysis on isolated neural cells (Chari et al., 2021). Sequence analysis of the gene lists obtained from reclustered neural cells facilitated the prediction of novel neuropeptide precursors (Chari et al., 2021). We analysed the expression of these along with previously characterised neuropeptide precursor genes (Takeda et al., 2018) and other markers by *in situ* hybridization. These experiments allowed a high diversity of expression domains in the nerve rings, in the manubrium and in the tentacular system to be uncovered (Chari et al., 2021). Pseudotime analysis of reclustered neural cells indicated the expression of neuropeptide precursors towards the end of the trajectory consistent with their function in differentiated subpopulations (Chari et al., 2021). In detail, cells expressing the RFamide precursor Pp5 are scattered throughout the manubrium and the nerve rings and are organised in two lines along the axis of the tentacles on the oral (neurogenic) side (Fig. 2.12c). The precursor gene *Pp25* shows partial co-expression with *Pp5* (Fig. 2.12c, atlas plot). Its expression pattern is similar to *Pp5* but highly enriched in the nerve rings and less detectable in the tentacles (Fig. 2.12c). *Pp17* enriched cells are specifically localised in the tentacles in the region opposite to mature nematocytes, on the same side as *Pp5* and *Pp25* expressing cells (Fig. 2.12c). *Pp7* expression could be detected in putative neural cells associated with the statocysts (balance organs) positioned around the rim of the bell (Fig. 2.12c and 2.13b). *Pp26* and *Pp1* show expression in patches of cells scattered throughout the nerve rings and the tentacles (Fig. 2.12c). *Pp11*-expressing cells are scattered on the manubrium and the nerve rings while *Pp20*-enriched cells show a particular organisation in patches on the distal portion of the tentacle bulbs (Fig. 2.12c). ELAV positive cells are located in dense aggregates in the nerve rings and scattered cells on the tentacle bulbs (Fig.2.12c). As described in Chari, we assigned cluster 9 as neural precursors based on the expression of the transcription factors Hlh6, Neurogenin and Sox10, previously described to be involved in neurogenesis.
To further investigate putative neural subtypes I analysed the expression patterns of additionally predicted neuropeptides (Chari et al., 2021). I detected *Pp14*-enriched cells on the oral side of the tentacle bulbs (Fig. 2.12c). *Pp27*-expressing cells are localised on the tentacles showing a similar expression pattern with *Pp17*-expressing cells although expressed in a different cluster (Fig. 2.12c). Finally we detected *Pp9* expression in a small number of

cells scattered in the nerve rings (Fig. 2.12c). These additional patterns contribute to the spatial mapping of putative neural subtypes in the medusa initiated in Chari et al. (2021).
The integration of three medusa datasets allowed me to increase the resolution of the clustering in regard to neural cells, with the detection of three more clusters (Fig. 2.12b). As already mentioned, the constraints associated with the computational correction of the batch effect of the integrated atlas do not allow reclustering analysis (see Methods). However, the increased number of cells certainly represents an advantage for a better representation of the gene expression and potentially allows a more detailed characterisation of putativel neural subtypes in the medusa.

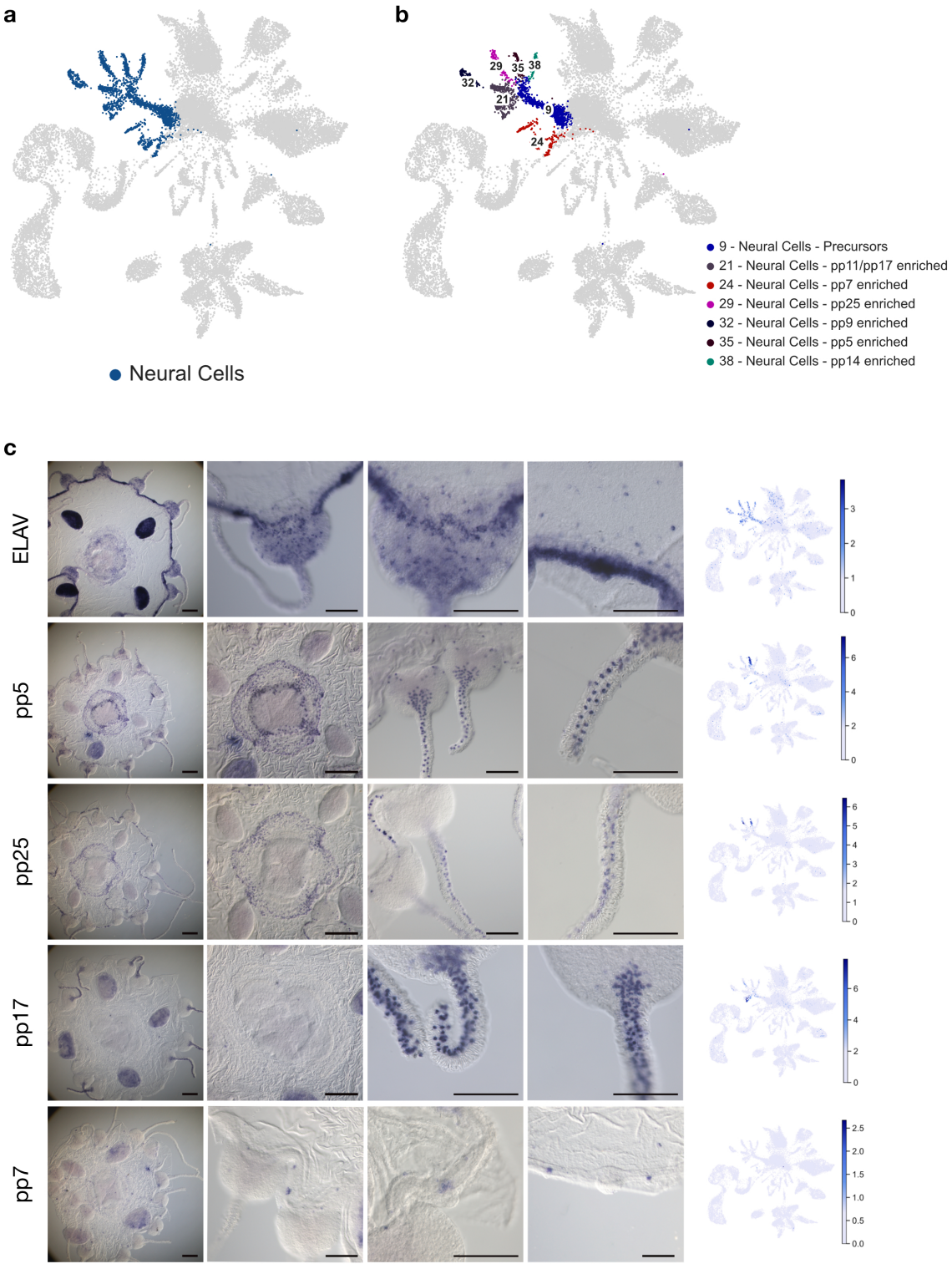


Figure 2.12 - Validation of neural cell types

(a) 2D-UMAP plot coloured by neural cells class annotation. (b) 2D-UMAP plot coloured by neural cells types annotation. (c) *In situ* hybridization of neural cell marker genes of 10-14 days-old medusa edited from Chari et al., 2021, Supplementary Figures. Images from left to right: whole medusa, manubrium, gonad, tentacle bulb, details of tentacles. Right column: 2D-UMAP plots representing the expression for each of the marker genes on the cell atlas. Scale bars in the whole medusa and manubrium images represent 200 µm; in the gonad, tentacle and tentacle bulb images they represent 100 µm.

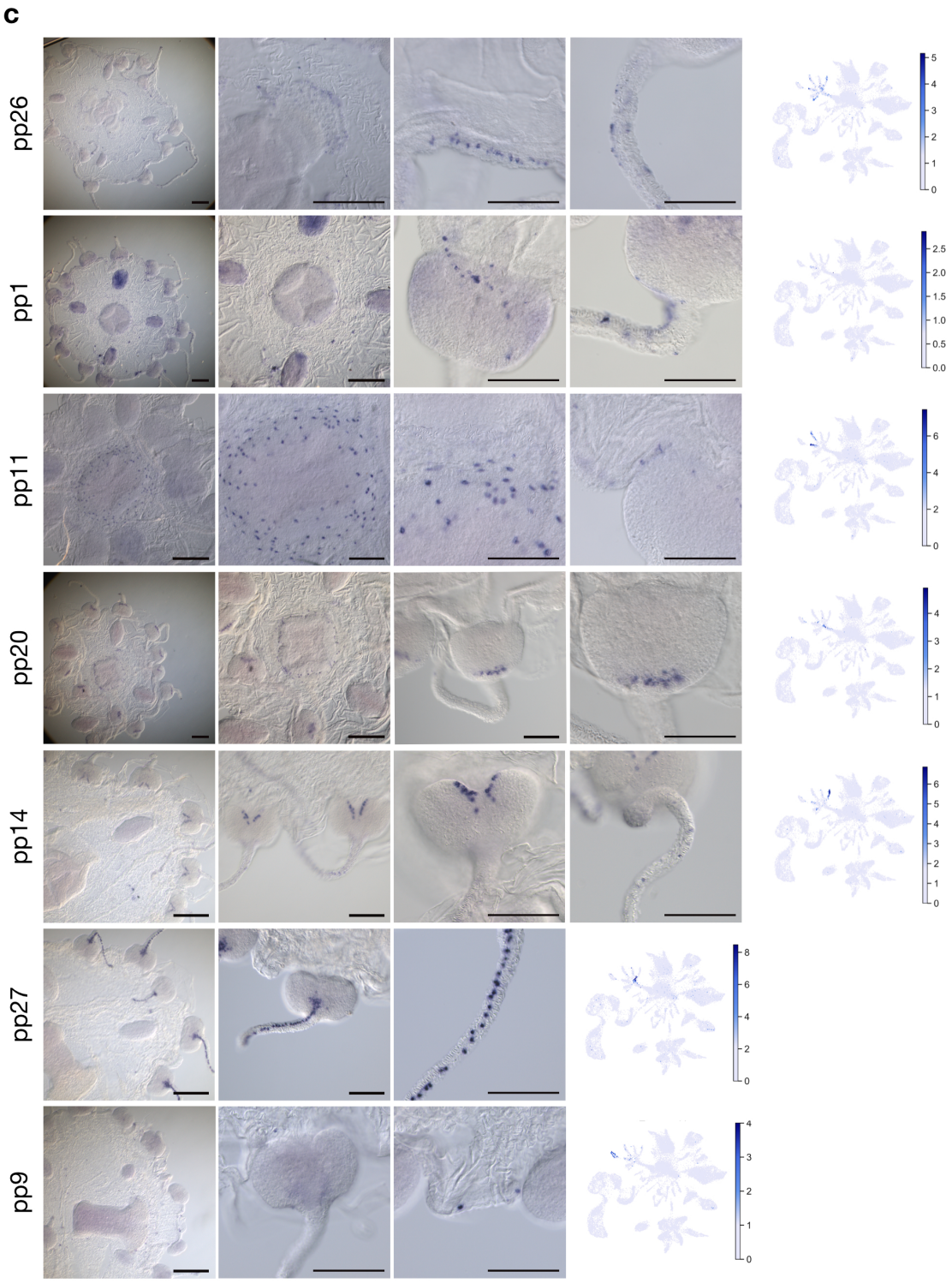


Figure 2.12c - Validation of neural cell types

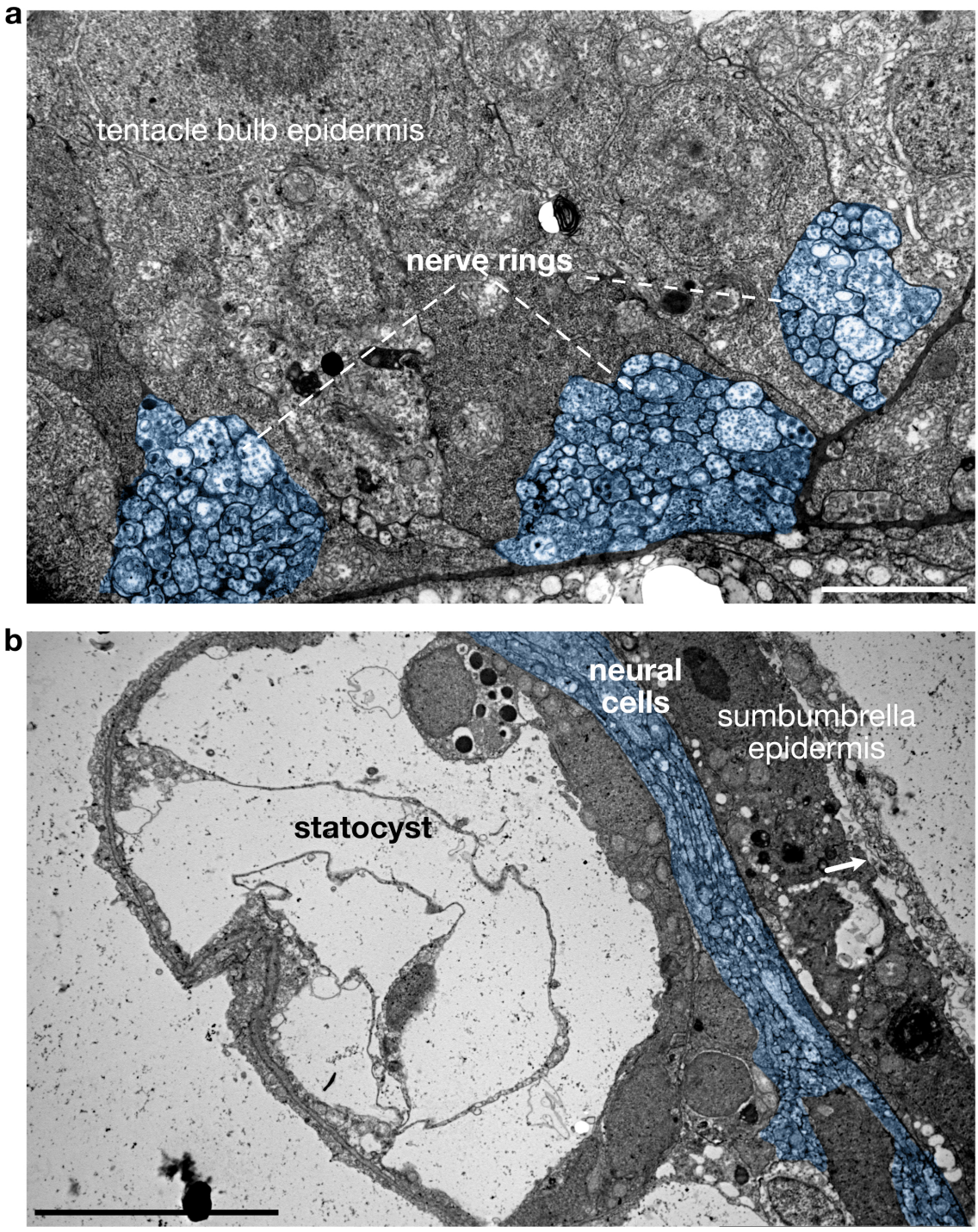


Figure 2.13 - Micrographs of neural cells
(a) Micrograph of a section through the tentacle bulb of the medusa of *Clytia.* Three groups of neural cells (blue) are condensed in the epidermis of the tentacle bulb and show the organisation of the nerve rings. (b) Micrograph of a section through the subumbrella of the medusa. Neural cells (blue) are condensed in a nerve ring (likely the sensory nerve ring) in contact with the sensory organ (statocyst). Scale bar in (a) represents 2µm. Scale bar in (b) represents 10µm.

2.2.1.8 - Gland Digestive Cells

As in our initial medusa atlas (Chari et al., 2021) , I could assign digestive gland cell identity to five clusters obtained from my combined dataset designated as digestive gland cells A-E. These each expressed different combinations of digestive enzymes (Fig. 2.14 a and b).

Previous studies have described at least three types of hydrozoan glandular digestive cells based on morphology and histochemistry. Two of those types are located respectively in the oral region of the hypostome of the polyps and the oral pole of the manubrium. The last type is prevalently found in the digestive trait/stomach region (Bouillon, 1966). Consistent with this observation, we detected the expression of cluster-specific digestive enzymes in particular in the manubrium but also in the gonads, as shown by *in situ* expression patterns of the genes designated as *FibrinogenC-dom1* and *C-type lectin* in gland digestive cells C and D (Fig. 2.14c). Gland digestive cells A specifically express *ShKT-TrypB* and they are detected in aggregates along the proximal part of the manubrium, including the stomach (Fig. 2.14c). Gland digestive cells B are localised at the base of the manubrium and express the gene*Trypsin-Like* as a unique marker (Fig. 2.14c). *ShKT-TrypA* is expressed in both gland digestive cells A and B, although the expression pattern does not recapitulate completely the expression patterns of cluster-specific markers. These cells are localised more orally in the manubrium showing a distinct expression pattern with respect to the other markers (Fig. 2.14c). ShKT-TrypB positive cells potentially represent an additional gland digestive cell type. Alternatively, this gene could mark a subpopulation of gland digestive cells A and B. These two hypotheses indicate that further diversity in this cell class remains to be discovered. Finally, gland digestive cells E express Chitinase, a marker shared with GD-C. This shows a similar expression pattern to the genes *BP10-like* and *Fibulin* at the contact between radial canals, manubrium and gonads, with the additional expression at the distal oral part of the manubrium but absent at the contact between tentacle bulbs and circular canal (Fig. 2.14c and Fig. 2.5c).

The expression profiles enriched in digestive enzymes and the localisation in the manubrium supports the hypothesis that the digestive gland cells are likely involved in the process of primary digestion. The secretion of several enzymes can be advantageous to cope with different types of prey organisms ingested. However, functional studies are needed to understand the exact role of each of those types. Micrographs of morphologically distinct types of digestive gland cells located at the basal side (a), on the “flanks” towards the apical side (b) and at the apical side of the manubrium (c and d) are shown in Figure 2.15.

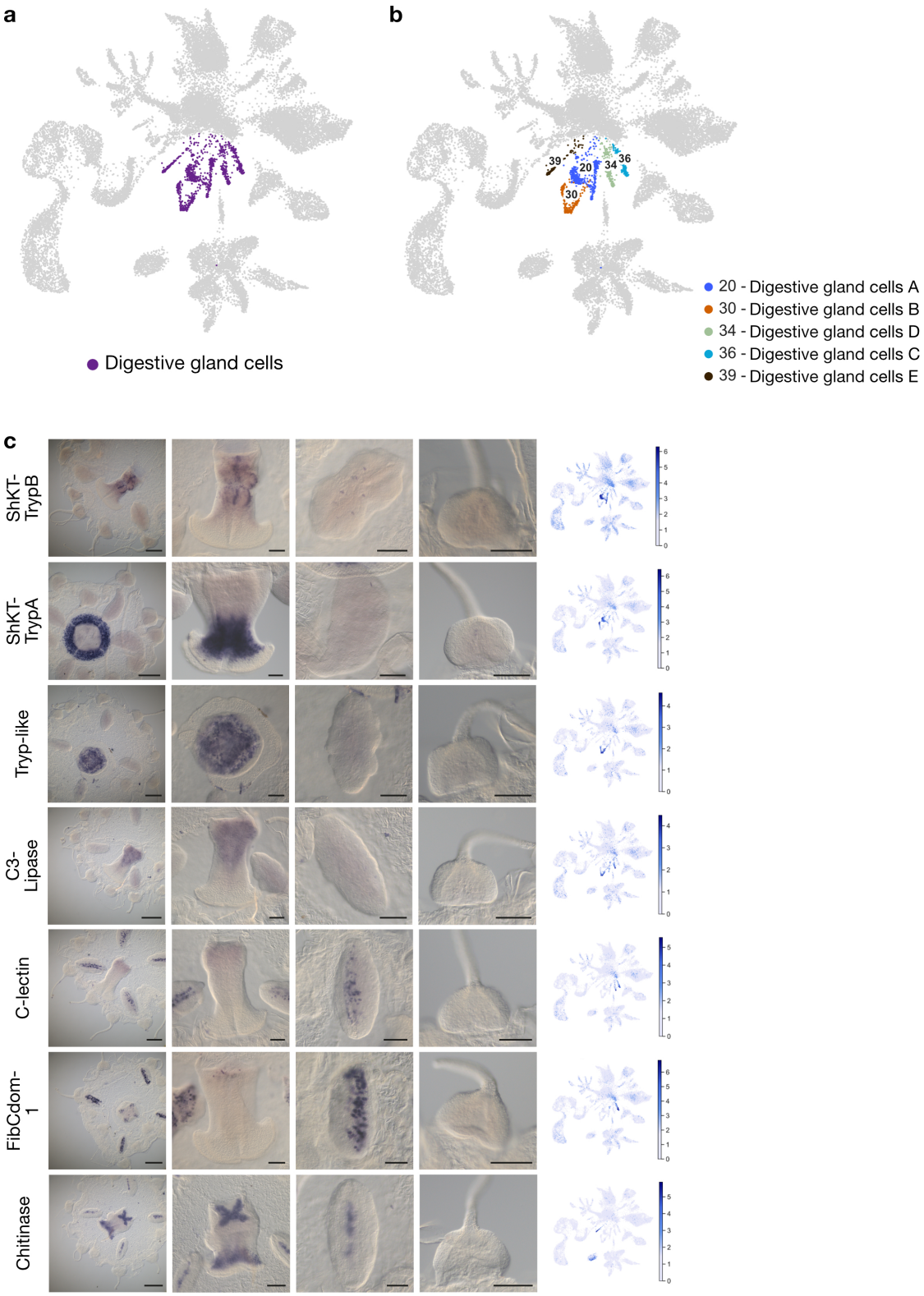


Figure 2.14 - Validation of digestive gland cell types

(a) 2D-UMAP plot coloured by gland digestive cell class annotation. (b) 2D-UMAP plot coloured by digestive gland cell types annotation. (c) *In situ* hybridization of digestive gland cell marker genes of 10-14 days-old medusa edited from Chari et al., 2021, Supplementary Figures. Images from left to right: whole medusa, manubrium, gonad, tentacle bulb, details of tentacles. Right column: 2D-UMAP plots

representing the expression for each of the marker genes on the Cell Atlas. Scale bars in the whole medusa images represent 200 µm; in the gonad, manubrium and tentacle bulb images they represent 100 µm. (d) Diagram of distribution of the digestive gland cell types in the manubrium of the medusa of *Clytia.*

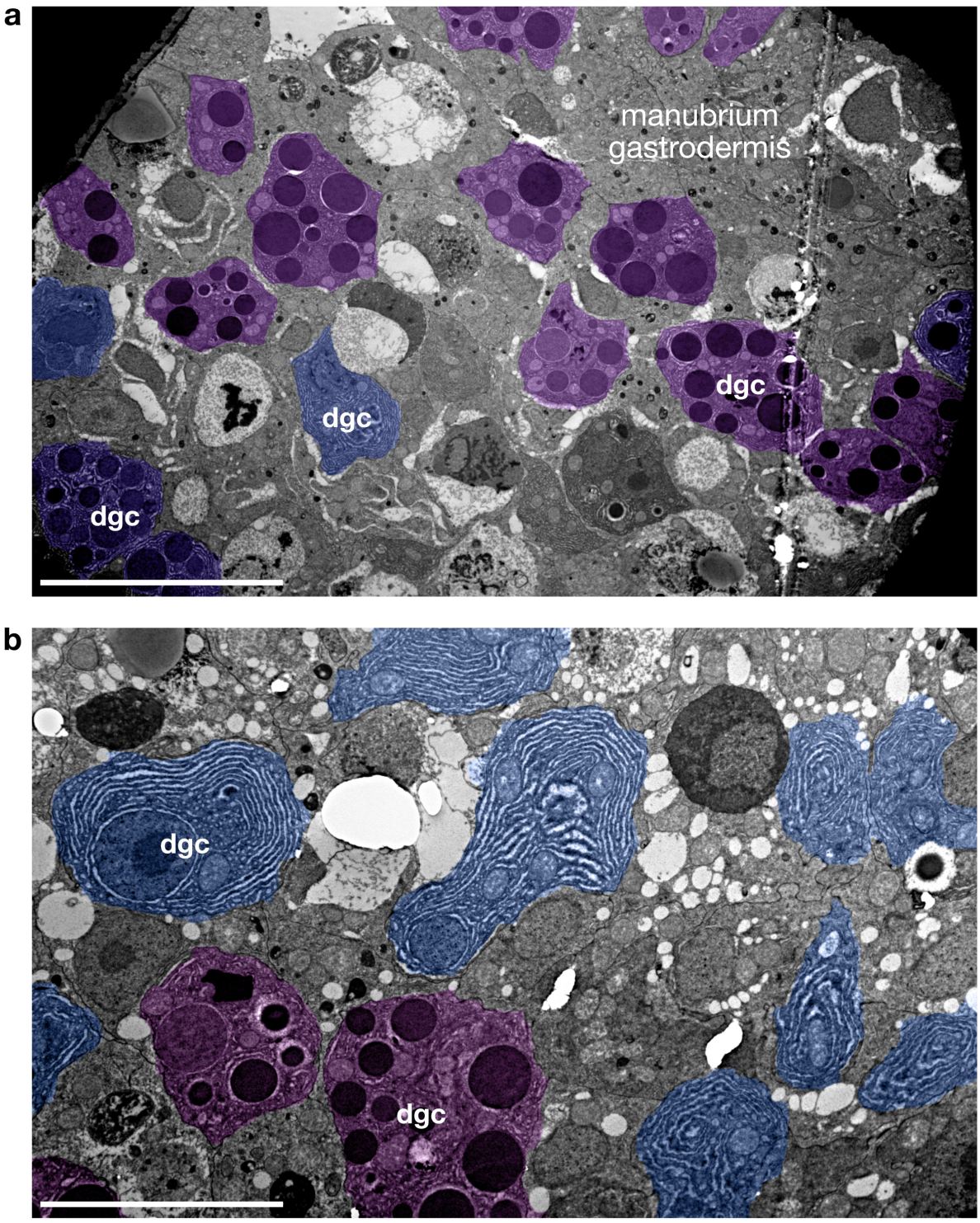


Figure 2.15a-b - Ultrastructure of digestive gland cells
Micrographs of sections through the manubrium of the medusa of *Clytia.* (a) Micrograph of a section through the basal part of the manubrium. Digestive gland cells scattered in the manubrium

gastrodermis. The three morphologically distinct types identified (dgc) are colured in magenta, blue and purple. (b) Micrograph of a section through the flanks of the manubrium toward the apical side, or “lips”. Morphologically similar digestive gland cell types (dgc) are coloured with the colour code in (a), (blue and magenta). Scale bas in (a) represents 20µm. Scale bar in (b) represents 10µm.

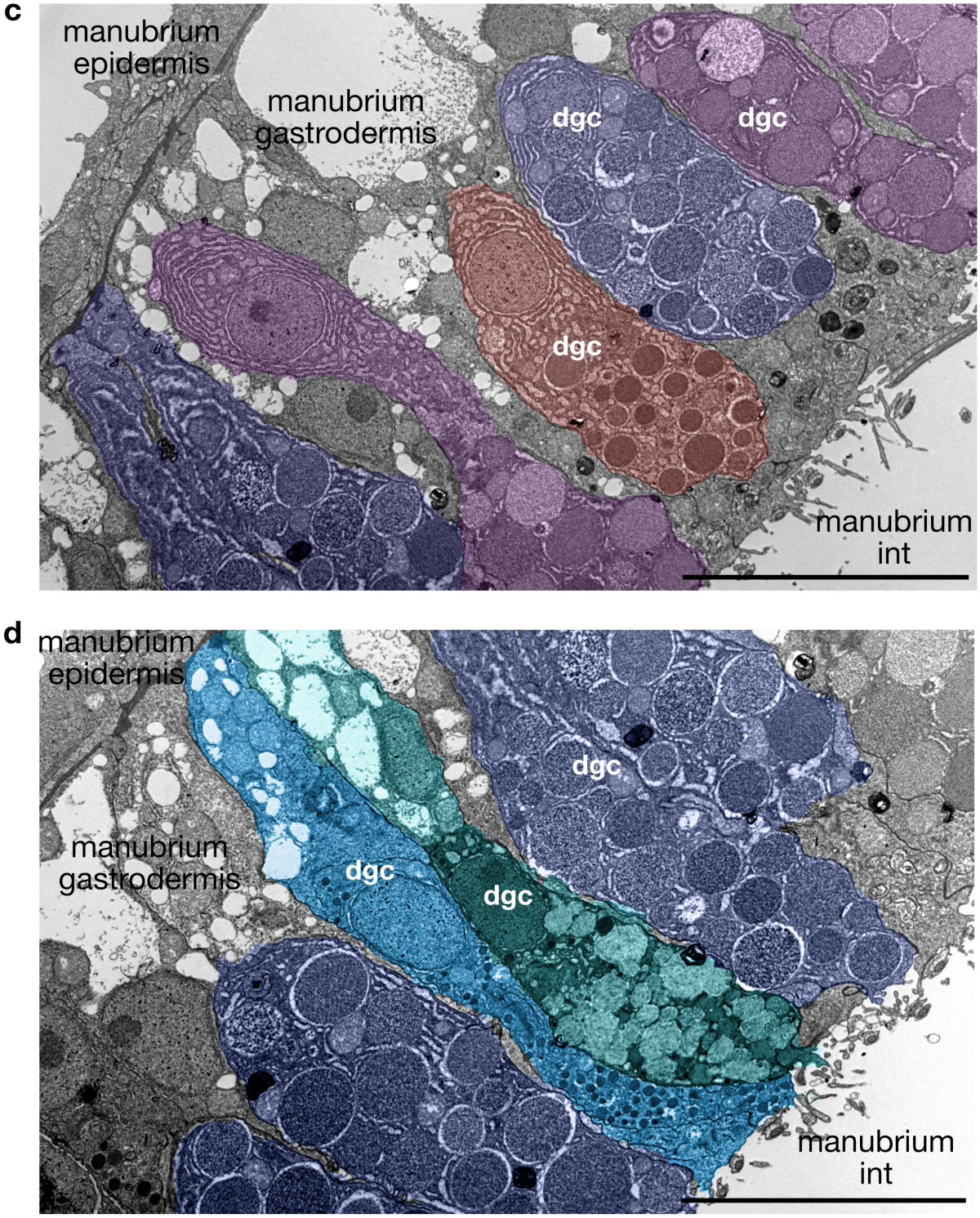


Figure 2.15c-d - Ultrastructure of digestive gland cells
Micrographs of sections through the apical side of the manubrium (“lips”) of the medusa of *Clytia.* (a) Micrograph of a section through the apical side of the manubrium. Digestive gland cells embedded in

the manubrium gastrodermis. The three morphologically distinct types identified (dgc) are colured in pink, lavender and orange. (b) Micrograph of another section through the apical side of the manubrium. Morphologically similar digestive gland cell types (dgc) are coloured with the colour code in (a), (lavender). Two additional morphologically distinct types are coloured in blue and cyan. The digestive gland cells facing the internal side of the manubrium (manubrium int). Scale bars in (a) and (b) represent 10µm.

2.3 - METHODS

2.3.1 - *Clytia* medusae Culture

Clytia female jellyfish (Z4B strain) were cultured as previously described (Lechable et al., 2020) for 10-14 days and fed twice a day.

2.3.2 - *Clytia* medusa dissociation

Dissociation of medusae was performed without the use of digestive enzymes. Medusae were collected between the 10th and the 14th day after hatching when sexual maturation was not yet reached. 2-3 jellyfish were used for each dissociation experiment. Jellyfish were washed in Ca/Mg-free artificial sea water (Table 2.1) three times by serial transfer in medium sized dishes. Cell dissociation was performed by incubating the jellyfish in Ca/Mg-free artificial sea water for ten minutes within a cell strainer and observed with a dissecting microscope. After incubation, excess sea water was removed with a tissue and jellyfish were gently pressed against the 40um mesh using the rubber plunger of a 1ml syringe. The cell strainer was washed five times with 100 µl of Low-Ca artificial sea water (Table 2.2). Single cells were collected in a 2ml tube by pipetting the solution from the bottom of the strainer until 500µl of cell suspension was collected. To estimate cell concentration, 20µl of cell suspension were counted manually at A2 Axio Imager (Zeiss) binoculars with the use of a Neubauer improved counting chamber (Sigma-Aldrich BR717810-1EA) filling the two grids available with 10µl of cell suspension each.
Cell mortality was estimated during previous separate test experiments diluting the cell suspension with 1:1 ErythrosinB 0.5mg/ml solution, which labels dead cells in red.
The mortality data were used to develop the final dissociation protocol with the aim of reducing cell mortality to 10-20% maximum.
Cell dissociation was performed immediately before the encapsulation experiment at the genomics facilities of the IPMC and the EMBL.
Repeated dissociation and cell counting experiments provided enough data to estimate that a young medusa consists of about 80000 cells.

Calcium/Magnesium - Free Artificial Sea Water	
NaCl	31 g
KCl	0.8 g
NaHCO3	0.29 g
Na2SO4	1.6 g
MilliQ H2O	Up to 1 litre

*final pH 8.0

Table 2.1 - Calcium/Magnesium - Free Artificial Sea Water

Low Calcium/Magnesium-Free Artificial Sea Water	
NaCl	26.88 g
KCl	0.74 g
CaCl2	0.16 g
HEPES	2.38 g
MilliQ H2O	Up to 1 litre

*final pH 7.6

Table 2.2 - Low-Calcium/Magnesium - Free Artificial Sea Water

2.3.3 - Cell Sorting, Encapsulation, Preparation of single cell libraries

Encapsulation of dissociated medusa cells was performed once at the UCAGenomiX platform at 'Institut de Pharmacologie Moléculaire et Cellulaire' (IPMC), in Sophia Antipolis, Nice, and once at the 'GeneCore' Genomics facility at the European Molecular Biology Laboratory (EMBL) in Heidelberg. Cells were encapsulated using the 10X Genomics Chromium and cDNA libraries were prepared according to the Chromium Next GEM Single Cell 3′ library preparation protocol v3.1 (https://www.10xgenomics.com). I prepared the cDNA libraries at the EMBL GeneCore while cDNA libraries were prepared by the technician in charge at the IPMC. Sequencing of single cell libraries at EMBL was carried out with a NextSeq 500 Midoutput kit and with NextSeq 500/550 Midouput kit at UCAGenomiX (75 cycles for both). In both cases we loaded 10000 freshly dissociated cells and aimed to recover 6-7000.

2.3.4 - Mapping and Clustering analysis

Demultiplexing and quality check of sequencing output was carried out by the respective genomics facility in both cases.

I received the sequencing files (fastq files) and mapped the reads individually for each sample against the genome reference (Leclère et al., 2019) using STARSolo (Dobin et al., 2013). The reference genome including annotation of mitochondrial genes was generated with STAR according to the STAR manual recommendation (Dobin et al., 2013). I set a cutoff and retained only cells with more than 100 UMIs for the following steps to allow a better exploration of the data. I applied the same workflow to the published "Fed and Starved"' medusa data. The resulting cell x gene count matrices were analysed with Scanpy 1.8.0 (Wolf et al., 2018) and Jupyter Notebook (https://jupyter.org/) .

Preprocessing and quality checks were performed broadly in accordance with the Scanpy tutorial (https://scanpy-tutorials.readthedocs.io/en/latest/pbmc3k.html) but with modifications outlined here. Initially I concatenated the 3 batches ("IPMC", "EMBL" and "FedStarved" medusa gene x count matrices) and removed 6209 low quality cells with less than 500 counts along with cells displaying more than 1% of mitochondrial gene counts, which are assumed to be poor quality cells (Lun et al., 2016). I decided on this threshold after several tests. Additionally I filtered out 4043 genes that were detected in less than 2 cells.
Counts per cell were normalised, log-transformed and highly variable genes were computed retaining genes with a minimum mean value of 0.0125 and a maximum mean value of 3. Counts were scaled and Principal Components (PCs) were computed on highly variable genes. At this step I noticed a large batch effect (Fig.2.16) and tested several batch effect correction methods (BBKNN (Polański et al., 2019), Scanorama (Hie et al., 2019), Seurat (Stuart et al., 2019).

Figure 2.16 - Uncorrected Batches

Finally, by applying the Harmony algorithm (Korsunsky et al., 2019; https://github.com/slowkow/harmonypy) iteratively and re-computing PCs with it, I achieved an homogeneous mixing of the cells from the three batches which converged after 8 iterations with all the batches contributing to all the cell types (Fig. 2.17).

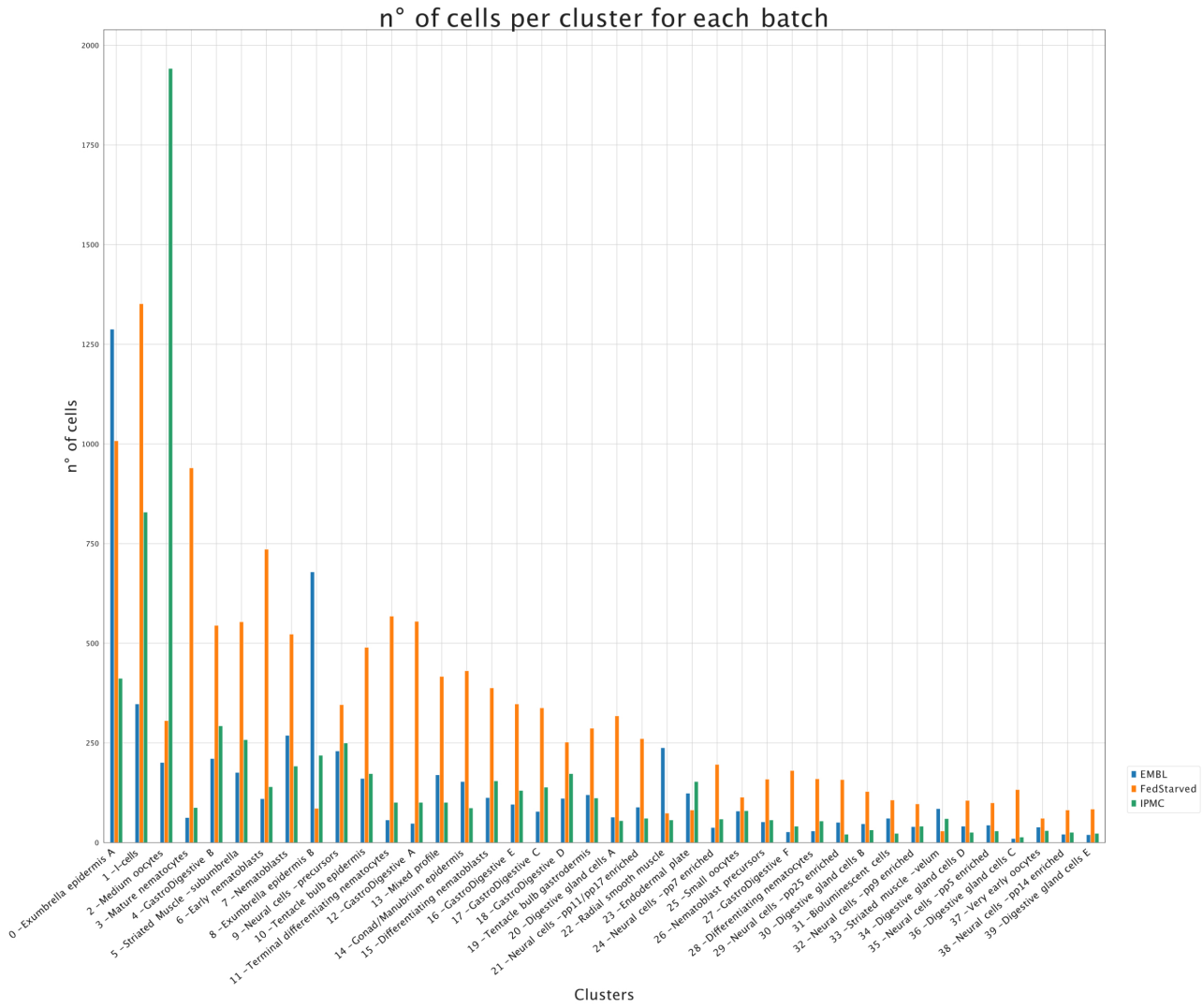


Fig.2.17 - Corrected Batches
Bar plot showing the contribution of the cells belonging to the 3 batches to the cell clusters after applying Harmony batch correction. Number of cells is plotted on the y axis and clusters are plotted on the x axis. Bars are coloured according to the batches (see legend)

The Neighbour Graph was computed with 40 PCs derived from the Harmony algorithm and 50 neighbours. The Neighbour graph was embedded in two dimensions using UMAP (McInnes et al., 2018). Clustering of the Neighbour Graph was performed with the Leiden algorithm (Traag et al., 2019) at resolution = 1.5 which resulted in 40 clusters. I applied the PAGA partition-based graph abstract method which provides a more faithful representation of the connectivities between cells (Chari et al., 2021; Wolf et al., 2019). I then generated the UMAP

embedding initialised with PAGA-graph structure and obtained the final atlas shown in Fig.2.1 and throughout this chapter.

Cell clusters in our integrated medusa dataset were annotated according to the expression of marker genes described in Chari et al., 2021. Batch effect correction performed with Harmony preserved key biological observations already reported in Chari et al., 2021, i.e. it returned largely the same cell types.

2.3.5 - Constraints of batch effect correction analysis

Batch effects can be caused by several factors, for instance, different technology platforms (10X Genomics, MARs-Seq etc), experimental design or sequencing conditions. At present, assessing which factors are causing batch effects in single cell data is still a challenge .

The Harmony method performs batch effect correction by aligning the cells in a low dimensional space and by recomputing the Principal Components iteratively until convergence. The underlying gene expression values remain unaltered (Korsunsky et al., 2019) which represents a disadvantage in the context of differential expression analysis. To overcome this issue I extracted marker genes using the 'FindAllMarker' function from the Seurat workflow (Stuart et al., 2019) in RStudio (https://satijalab.org/seurat/), adding the options 'test.use=LR' and 'latent.vars = "batch"' which allows the generation of a list of marker genes per cluster while taking into account the batches.

```
x <- FindAllMarkers(object = object, test.use = "LR", latent.vars = "batch")
```

According to the documentation (https://satijalab.org/seurat/reference/findallmarkers) this function provides a list of differentially expressed genes for each of the variables stated with the 'latent.vars' option which in our case refers to the 3 batches constituting the integrated medusa dataset (notably "IPMC", "EMBL" and "FedStarved"). Additionally, the option 'test.use=LR' computes differentially expressed genes through a logistic regression model and by comparing it to a null one using a Likelihood Ratio test, which provides a more robust prediction of differentially expressed genes with respect to other tests allowed by this function.

Given the constraints derived by the batch effect correction I did not perform pseudotime analysis to address the development of digestive gland cells given that it couldn't be possible to confidently determine genes expressed in pseudotime. Furthermore, I tested re-clustering of isolated neural cells using raw data by computing batch effect correction after preprocessing. I found that this analysis was unreliable because of the lack of correction of the gene expression values between batches.

2.3.6 - *In situ* Hybridisation

In situ hybridisation was performed in a robot (Intavis AG, Bioanalytical Instruments) as previously described (Chari et al., 2021). Probes were generated by polymerase chain reaction (PCR) from cDNA clones of our expressed sequence tag (EST) collection (Chari et al., 2021; Chevalier et al., 2006) and from cDNA of medusa. Oligonucleotide primers were designed with PrimerBlast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) for Pp9, Pp14 and putative neuropeptide probes and successfully cloned in pGemT-easy vector according to the manufacturer recommendations. Sequences of the oligonucleotide primers and EST identification names are provided in Table A which is included in the annexed electronic documents for this manuscript.
Some of the *in situ* images shown in the figures of the chapter were generated by merging images from several z-planes with the program Helicon Focus (https://www.heliconsoft.com/).

2.3.7 - Confocal microscopy

Confocal microscopy of the gonads of the medusa of *Clytia* was performed as previously described (Chari et al., 2021).

2.3.8 - Transmission Electron Microscopy

Fixation of 10 days old medusae, embedding in Epoxy resin and sectioning for TEM were performed by Sophie Pagnotta at the Plateforme Commune de Microscopie Électronique, using the Osmium/glutaraldehyde fixation method described by Eisenman & Alfert (1982). With Evelyn Houliston, we joined Sophie Pagnotta for the imaging sessions at the Plateforme Commune de Microscopie Électronique, Université Côte d'Azur.

2.4 - DISCUSSION

The integrated medusa cell atlas that I have presented in this chapter is an extension and consolidation of our initial medusa cell atlas published in Chari et al. (2021). The main purpose of performing the integration was to increase resolution of clusters within the cell classes and provide for a more robust characterisation of the cell types. Cell types in this integrated medusa atlas match almost perfectly with the cell types described in Chari et al. (2021) supporting the previous classification. In addition, I increased resolution within the neural cell class and increased the overall number of cells per cluster.

The final integrated medusa cell atlas consists of 40 cell types grouped into seven cell classes as in Chari et al. (2021). Two of the seven classes correspond to epitheliomuscular cells of the epidermis and gastrodermis. The stem-cells/germ-cells class includes oocytes and the population of hydrozoan stem cells or i-cells. The remaining classes are the neural cells, nematocytes, bioluminescent cells and digestive gland cells.

Single cell transcriptome data of the medusa revealed a rich diversity within the epidermis/muscle cell class which includes seven cell types covering basic structural epidermis types and unique smooth and striated muscle types. Unexpected diversity was also detected within the gastrodermis cell class which includes eight cell types. Within those, we could identify two 'structural' types associated with extracellular matrix and mesoglea formation and six previously undescribed types (or states), very likely involved in different digestive functions, which remain to be investigated in detail.

Our atlas supports the presence of the hydrozoan stem cell population, the interstitial cells or i-cells, well described in Hydra (Bode, 1996). In *Hydra* i-cells give rise to four main derivatives: neural cells, nematocytes, gland cells and germ cells (Bode, 1996). We provided developmental transcriptional signatures of neurogenesis and nematogenesis from i-cells by pseudo-time analysis (Chari et al., 2021). In contrast we could not detect convincing developmental transcriptional signatures linking the digestive gland cells to the i-cells (Chari et al., 2021). This is possibly due to a self-renewal mechanism gland cells following differentiation which results in the maintenance of a distinctive transcriptional profile. Additionally, the link between i-cells and gland cells in *Hydra* is corroborated by analysis of the single cell transcriptomics data. Likely, the large availability of those cells played a role in the determination of this connection (Siebert et al., 2019). My integrated medusa atlas provides a larger set of cells allowing to test this hypothesis in *Clytia.* For this purpose, a refinement of the computational correction of the batch effect will be essential.

As in Chari et al. (2021), we assigned six clusters to digestive gland cells based on strong and distinctive transcriptional signatures, including the production of different combinations of multiple digestive enzymes (Chari et al., 2021). We classified seven clusters as neural cells given the expression of ELAV (Nakanishi et al., 2012) and several neuropeptide precursors (Chari et al., 2021). Re-clustering analysis of neural cells in Chari et al. (2021) revealed 14 sub-clusters encompassing one cluster of neural precursors and 13 putative specialised

subtypes. I further explored the possibility that additional neural subpopulations could be present in the medusa by *in situ* hybridisation of additional newly predicted neuropeptide precursors. Although these did indeed show some differences in spatial distributions, in several instances the neuropeptide precursors mapped to the same clusters of our integrated medusa atlas. This suggests that additional neural heterogeneity is still overlooked. Increased scRNAseq data targeted to medusa neural cells, further re-clustering analyses as well as functional studies will be required to fully address the neural subpopulations of the medusa.

Finally, we assigned seven clusters to the *Clytia* nematocytes. We recovered a large number of cells for this class, which allowed the characterisation of two phases of nematogenesis displaying distinct but consecutive transcriptional programs. We assigned four clusters as nematoblasts, covering the first phase of nematogenesis, characterised by the formation of the characteristic nematocyte capsule (Chari et al., 2021). The expression of genes widely associated with this process (Condamine et al., 2019; Denker et al., 2008) argues in favour of the classification. The second and terminal phase of nematogenesis includes three clusters that correspond to the cell migration along the axis of the tentacles and their subsequent deployment in the tentacle ectoderm. These stages are characterised by the formation of the nematocil, a specialised mechanosensory structure associated with nematocyte cilium which triggers the ejection of the capsule upon contact with prey. This is supported by the expression of the specific structural component Nematocilin in the final step of maturation (Balasubramanian et al., 2012) along with homologs of components of the functionally equivalent 'stereocilia' of vertebrate hair cells (Gillespie & Müller, 2009). Expression of these genes underlines the close relationship of this unique cnidarian cell type to mechanosensory neurons (Bosch et al., 2017).

The integration analyses represented a challenge due to a large batch effect. Assessing which is the factor responsible for batch effect is also quite difficult since several variables can contribute to it at the same time. Very likely in our case the major cause is the use of different versions of the 10X genomics chemistry to encapsulate and generate the single-cell cDNA libraries (v2 for "Fed and Starved" dataset, Chari et al., 2021; v3.1 for "IPMC" and "EMBL" datasets). Another factor is probably represented by the sequencing setup and by the experimental conditions (fixed cells for "Fed and Starved" dataset, Chari et al., 2021; freshly dissociated cells for "IPMC" and "EMBL" datasets). After testing several batch effect correction workflows we obtained the best result by applying Harmony (Korsunsky et al., 2019). The Harmony algorithm computes the alignment of the cells in a low dimensional space achieving a joint embedding of similar cells. However, the underlying gene expression matrix remains unaltered which is not suitable for differential expression analysis as well as pseudotime or re-clustering analysis (Korsunsky et al., 2019). We overcame this issue by using an alternative method to extract marker genes taking into account the variability between the batches. Furthermore, we detected a similar content of cell types and marker genes across the three atlases when analysed individually. Finally, the description of the medusa cell types and their marker genes allowed the validation of the integrated medusa dataset by mapping the previously identified genes on the integrated atlas.

The *Clytia* medusa cell type atlas represents a valuable tool to address the development and the evolution of this most complex cnidarian body form at the cell type level. We found larger

cell type diversity in the medusa of *Clytia* compared with the hydrozoan polyp-only *Hydra* and the anthozoan *Nematostella*, concordant with the higher degree of complexity of the medusa. As the single-cell "field" moves forward very rapidly and the number of the available atlases increases accordingly, this medusa cell type atlas can be seen as a useful resource to encourage the classification of cell types in other medusozoan species.
In addition, to the molecular characterisation we initiated a morphological description of the medusa cell types based on their ultrastructure which opens the way for more in-depth studies of the 3D organisation of selected types.

Annex – Whole animal multiplexed single-cell RNA-seq reveals transcriptional shifts across Clytia medusa cell types

SCIENCE ADVANCES | RESEARCH ARTICLE

GENETICS

Whole-animal multiplexed single-cell RNA-seq reveals transcriptional shifts across *Clytia* medusa cell types

Tara Chari[1]†, Brandon Weissbourd[1,2,3]†, Jase Gehring[4]†, Anna Ferraioli[5]†, Lucas Leclère[5], Makenna Herl[6], Fan Gao[7], Sandra Chevalier[5], Richard R. Copley[5]*, Evelyn Houliston[5]*, David J. Anderson[1,2,3]*, Lior Pachter[1,8]*

We present an organism-wide, transcriptomic cell atlas of the hydrozoan medusa *Clytia hemisphaerica* and describe how its component cell types respond to perturbation. Using multiplexed single-cell RNA sequencing, in which individual animals were indexed and pooled from control and perturbation conditions into a single sequencing run, we avoid artifacts from batch effects and are able to discern shifts in cell state in response to organismal perturbations. This work serves as a foundation for future studies of development, function, and regeneration in a genetically tractable jellyfish species. Moreover, we introduce a powerful workflow for high-resolution, whole-animal, multiplexed single-cell genomics that is readily adaptable to other traditional or nontraditional model organisms.

INTRODUCTION

Single-cell RNA sequencing (scRNA-seq) is enabling the survey of complete transcriptomes of thousands to millions of cells (*1*), resulting in the establishment of cell atlases across whole organisms (*2*–*6*), exploration of the diversity of cell types throughout the animal kingdom (*3*, *7*–*9*), and investigation of transcriptomic changes under perturbation (*10*, *11*). However, scRNA-seq studies involving multiple samples can be costly and may be confounded by batch effects resulting from multiple distinct library preparations (*12*, *13*). Recent developments in scRNA-seq multiplexing technology expand the number of samples, individuals, or perturbations that can be incorporated within runs, facilitating well-controlled scRNA-seq experiments (*11*, *14*–*18*). These advances have created an opportunity to explore systems biology of whole organisms at single-cell resolution, merging the concepts of cell atlas surveys with multiplexed single-cell experimentation.

Here, we apply this powerful experimental paradigm to a planktonic model organism. We examine the medusa (free-swimming jellyfish) stage of the hydrozoan *Clytia hemisphaerica*, with dual motivations. First, *Clytia* is a powerful, emerging model system spanning multiple fields, from evolutionary and developmental biology to regeneration and neuroscience (*19*–*24*). While previous work has characterized a number of cell types in the *Clytia* medusa (*21*), a whole-organism atlas of transcriptomic cell types has been lacking. Such an atlas is a critical resource for the *Clytia* community and an important addition to the study of cell types across animal phylogeny.

Second, emerging multiplexing techniques present new opportunities for system-level studies of cell types and their changing states at unprecedented resolution in whole organisms. The *Clytia* medusa offers an appealing platform for pioneering these studies. It is small, transparent, and has simple tissues and organs, stem cell populations actively replenishing many cell types in mature animals, and remarkable regenerative capacity (*19*, *22*, *24*–*27*). Furthermore, the 1-cm-diameter adult medusae used in this study contain on the order of 10^5 cells, making it possible to sample cells comprehensively across a whole animal in a cost-effective manner using current scRNA-seq technology (fig. S1 and tables S1 and S2). In this study, we generate a cell atlas for the *Clytia* medusa while simultaneously performing a whole-organism perturbation study, providing the first medusa single-cell dataset and an examination of changing cell states across the organism. Our work also provides a proof-of-principle for perturbation studies in nontraditional model organisms, using multiplexing technology and a reproducible workflow with lessened reliance on functional annotation, from the experimental implementation to the data processing and analysis.

RESULTS

We compared control versus starved animals, as this strong, naturalistic stimulus was likely to cause notable, interpretable changes in transcription across multiple cell types. Laboratory-raised, young adult, female medusae were split into two groups of five animals, one deprived of food for 4 days, and the second fed daily (see Materials and Methods). We observed numerous phenotypic changes in starved animals, including a marked size reduction reflecting two- to threefold fewer cells (Fig. 1, fig. S2, and see Materials and Methods) (*28*), and a notable reduction in gonad size. Correspondingly, the number of eggs released per day decreased (fig. S3) (*29*).

For scRNA-seq, single-cell suspensions were prepared from each whole medusa and individually labeled with unique ClickTag barcodes (*14*) using a seawater (SW) compatible workflow (see Materials and Methods, tables S2 and S3, Supplementary Methods, and fig. S4). All labeled suspensions were pooled and processed with the 10X Genomics V2.0 workflow and Illumina sequencing, allowing construction of a combined dataset across organisms and treatments,

[1]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA. [2]Tianqiao and Chrissy Chen Institute for Neuroscience, Pasadena, CA 91125, USA. [3]Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125, USA. [4]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA. [5]Sorbonne Université, CNRS, Laboratoire de Biologie du Développement de Villefranche-sur-mer (LBDV), 06230, France. [6]University of New Hampshire School of Law, Concord, NH 03301, USA. [7]Caltech Bioinformatics Resource Center, California Institute of Technology, Pasadena, CA 91125, USA. [8]Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125, USA.
*Corresponding author. Email: copley@obs-vlfr.fr (R.R.C.); houliston@obs-vlfr.fr (E.H.); wuwei@caltech.edu (D.J.A.); lpachter@caltech.edu (L.P.)
†These authors contributed equally to this work.

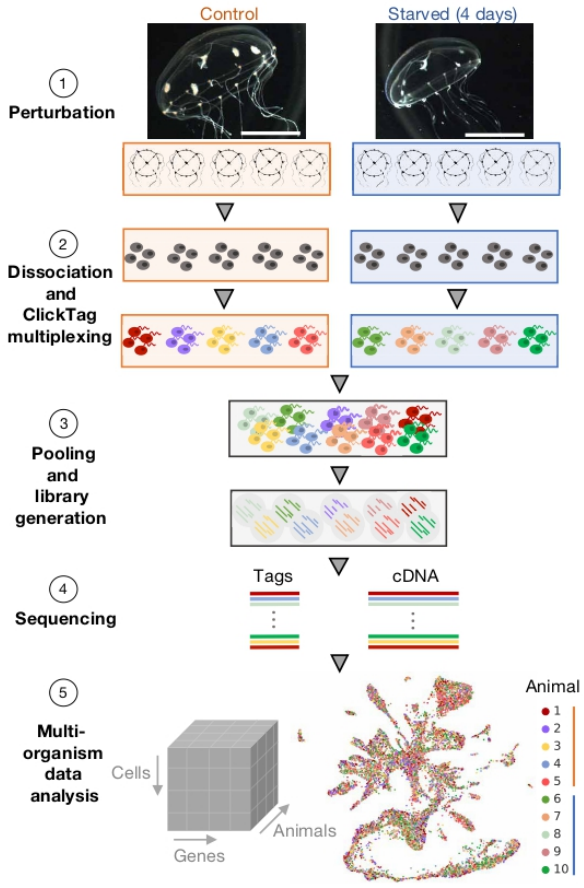


Fig. 1. Overview of whole-organism multiplexed experimentation. Experimental design of the starvation experiment showing (1) images of control versus 4-day starved animals (scale bars, 0.5 cm), (2) dissociation of individual medusa and chemical tagging of cells with ClickTags to enable multiplexed scRNA-seq, (3) pooling of cells and library generation from lysed cells to generate (4) sequencing libraries for the multiplexed cDNA and ClickTag data and create (5) single-cell resolved gene expression count matrices from all animals (see Materials and Methods).

without requiring batch correction (Fig. 1, fig. S4, A to D, and table S1). A total of 13,673 single-cell profiles derived from 10 individuals (5 control and 5 starved) passed quality control, with high concordance in cell type abundance and gene expression among animals in the same treatment condition (see Materials and Methods and fig. S5). From this gene expression matrix, we (i) derived a *Clytia* medusa cell atlas and (ii) generated a high-resolution resource of the transcriptional impact of starvation across all observed cell types.

To validate the cell atlas and assess technical variability across and within multiplexed experiments, we performed an additional, independent round of sequencing from 12 individuals. We found that cell types were highly concordant between experiments and confirmed a reduction of batch effect–induced variability within multiplexed experiments (see below). During this second sequencing run, we took advantage of our multiplexing approach to perform an experiment designed both to search for transcripts with "immediate early gene (IEG)"–like behavior in *Clytia* and to test its sensitivity for detecting more rapid or subtle gene changes than those of the extreme starvation perturbation. For this, we exposed *Clytia* medusae to multiple transient, ionic stimuli and dissociated ~1 hour later. This paradigm allowed us to identify candidate genes with IEG-like properties across many cell types, including neurons (figs. S6 to S8 and table S4) (*30*). IEGs are valuable tools in neuroscience, to identify neurons that are active following a specific stimulus or behavior (*30*). This methodology is thus able to detect transcriptional responses across diverse stimulus-response paradigms (table S5).

A *Clytia* cell atlas

To generate the cell atlas, we clustered the cells using the gene expression matrix, extracting 36 cell types and their corresponding marker genes (see Materials and Methods; Fig. 2, A and B; figs. S9 to S11; and table S5). Each of the cell types was present in each of the individual animals sequenced (fig. S12). We then generated a low-dimensional representation (*31*, *32*) of these cell types (Fig. 2A). We could group the cell types into seven broad classes (Fig. 2A) that correspond to the outer epidermis, the inner gastrodermis, and to likely derivatives of the multipotent interstitial stem cell population (i-cells). I-cells are a specific feature of hydrozoans, and are particularly well characterized in *Hydra*, where they generate neural cells, gland cells, and stinging cells (nematocytes), as well as germ cells (*8*, *20*, *33*). Our dataset was derived from female medusae so it lacks male germ cells, and late stage oocytes are expected to be too large for capture by the dissociation procedure.

The 36 cell types (see Materials and Methods and Fig. 2, B to D) were concordant between the two separate multiplexed experiments (see "Starvation" and "Stimulation" sections in Materials and Methods) and robust to different transcriptome annotations (figs. S6 and S13). For some of them, cell type identity could be assigned on the basis of published information on gene expression in *Clytia* and/or of homologous genes in other animals, while for the others we performed in situ hybridization for selected marker genes (Fig. 2C, figs. S11 and S14, and table S3). Previously known cell types apparent in our data included i-cells (*34*) and nematocytes at successive stages of differentiation (*35*–*37*), as well as oocytes (*38*), gonad epidermis, manubrium epidermis, and bioluminescent cells in the tentacles that each express specific endogenous green fluorescent proteins (GFPs) (*39*).

In situ hybridization for a selection of diagnostic muscle cell type genes allowed us to describe cell types making up the smooth and striated muscles, for instance, distinguishing the striated muscle cells lining the bell (subumbrella) and velum (Fig. 2, C and D, and fig. S14) (*23*, *27*). Within known cell types, clustering revealed an unappreciated degree of cell heterogeneity, yielding novel subtypes. For example, eight cell types could be distinguished within the gastrodermis, six of which were designated gastro-digestive (GD A-F) on the basis of a largely shared set of marker genes (Fig. 2B), including enzymes associated with intracellular digestion, such as CathepsinL (*40*). Unlike most other clusters, the GD clusters differ primarily in their relative levels of gene expression, rather than by unique marker genes. They thus likely represent variations on a similar digestive-absorptive epithelial cell type with different functional specializations, distributed across the main digestive compartments of the gastrodermis—the manubrium, gonad, and tentacle bulb—and the gastrovascular canals that link them (figs. S11 and S14). Comparison of gene modules

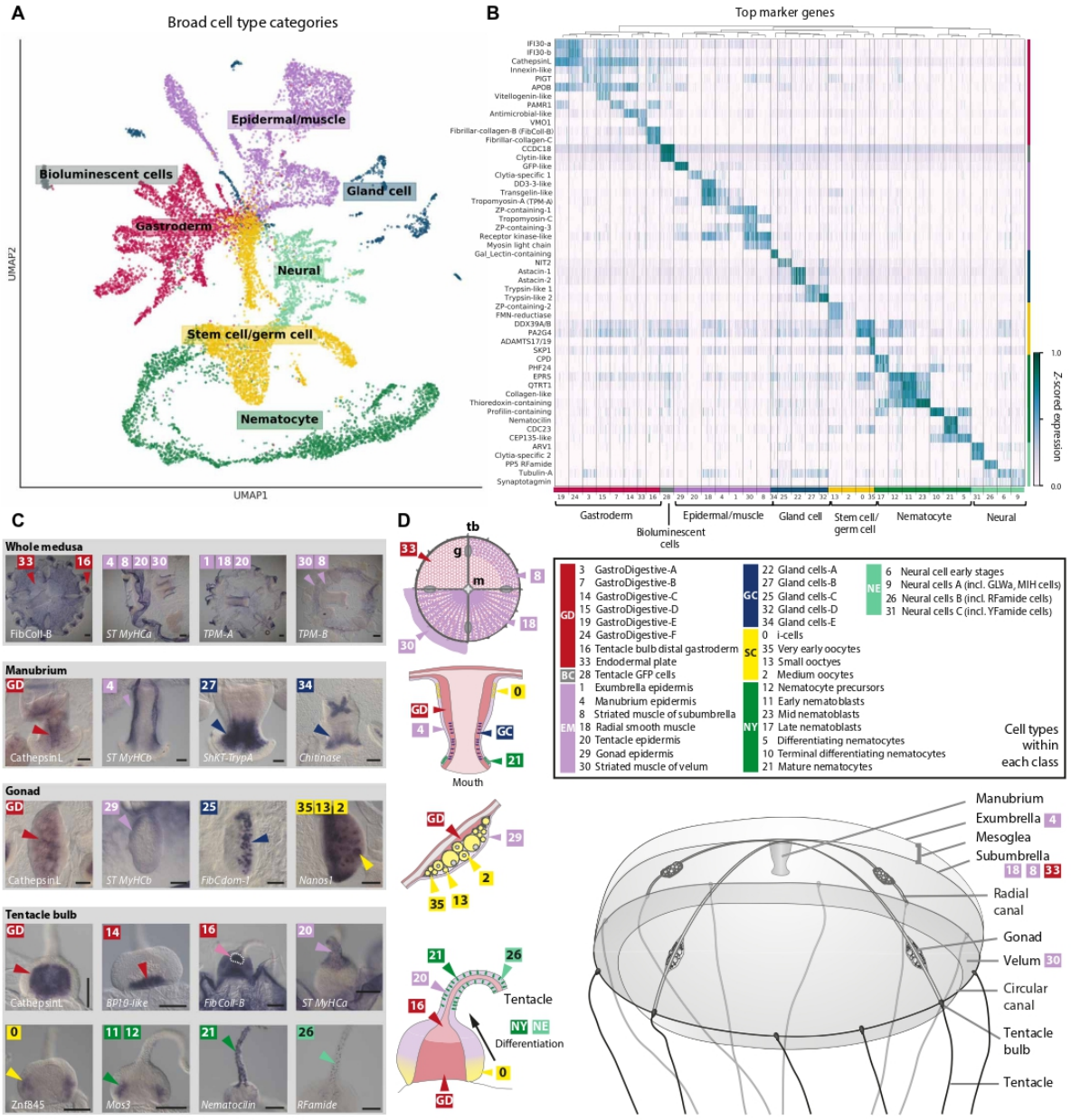


Fig. 2. The *Clytia* Medusa Cell Atlas. (**A**) Two-dimensional UMAP embedding of cells labeled by seven cell type classes. Class colors are retained in (**B**) to (**D**). (B) Heatmap of top marker genes from the sequencing data with 36 Louvain clusters comprising the seven cell type classes. (C) In situ hybridization patterns for a selection of cluster marker genes providing spatial location on the animal (comprehensive set in fig. S14). The label GD denotes general markers for GastroDigestive cell types. Scale bars, 100 μm. (D) Schematics of *Clytia* medusa, manubrium, gonad, and tentacle bulb showing the main cell types. Abbreviations of cell class names: GD, GastroDigestive; BC, bioluminescent cells; EM, epidermal/muscle; GC, gland cells; SC, stem cell/germ cell; NY, nematocytes; NE, neural cells.

discriminating these gastrodermal clusters (fig. S15) indicates that GD-B may have a particular role in transforming growth factor–β signaling, likely involving the ligands BMP2/4 and BMP5/8. GD-D is enriched for a module associated with cell-cell junctions, while GD-F shows relative depletion, suggesting poorer integration into the gastrodermal epithelium (fig. S15) and possible involvement in GD cell mobilization during starvation and regeneration (*19*). GD-C cells, localized closest to the endodermal plate, are enriched for transcripts associated with extracellular matrix and mesoglea (jelly) production (Fig. 2, C and D, and figs. S11 and S14). Expression of these and other genes implicated in mesoglea production, such as fibrillar collagens, is also a characteristic of endodermal plate cells (cluster 33) and proximal tentacle-bulb endoderm cells (cluster 16).

Digestive gland cells fell into five types expressing different mixtures of enzymes for extracellular digestion. These showed overlapping distributions in the mouth and stomach regions of the manubrium. Two subtypes of gland cells (type C and E) were also present within the gonad gastroderm. Four broad clusters corresponding to neural cells each appeared to represent mixed populations and could be subdivided by further analyses to define 14 likely subpopulations of neurons (see below). Seven major clusters could be assigned identities as nematocytes at different developmental stages, comprising two groups with highly distinct transcriptional signatures. Four of these we designate "nematoblasts" on the basis of high levels of transcripts related to formation of the nematocyst (stinging capsule) (*35*, *36*, *41*). The other three, designated as differentiating and mature "nematocytes," show no enrichment of these nematocyst transcripts but strongly express highly conserved proteins of the actin-rich "stereovilli" of vertebrate hair cells, including Whirlin, Harmonin, and Sans/USH-IG. This is consistent with observations of similar actin-based protrusions surrounding a central cilium in many of the mechanosensory cell types described in other cnidarian species (*42*). Related but more elaborate actin structures are associated with the cnidocil of mature nematocytes, but it had not previously been known to share functional hair-cell components (*42*, *43*). Nematocilin, a hydrozoan-specific component of the nematocil (ciliary trigger for nematocyte discharge) (*44*), is also expressed in these clusters (fig. S14, table S5, and see below). In situ hybridizations revealed marker expression in morphologically distinguishable nematocytes, notably including two lines along the oral face of each tentacle (Fig. 3E and fig. S14), a notable arrangement overlooked in previous studies.

A remarkable feature of the *Clytia* medusa is that it constantly generates many cell types, notably neural cells and nematocytes from prominent i-cell pools in the tentacle bulb epidermis (*36*) and at other sites (*34*). Within our dataset, we thus expected to be able to capture dynamic information relating to the development of i-cell–derived cell types, similar to that extracted from Hydra polyp single-cell transcriptome data (*8*). As in *Hydra*, our cell atlas revealed clear connections between the neuronal and nematocyte populations and the i-cell population (Fig. 2A and figs. S11 and S14) (*3*, *8*), likely corresponding to differentiation trajectories (*35*, *36*). In contrast, we found no clear developmental connection between i-cells and gland cells and little to no expression of markers of the common neuronal-gland cell precursors identified in *Hydra* (*8*) (fig. S16). In *Hydra*, gland cells are generated not only from the i-cell lineage but also by processes of self-renewal and position-dependent transdifferentiation (*8*). In the *Clytia* medusa, digestive gland cells show widespread distribution across distinct regions of the manubrium and gonad compartments of the gastrovascular system, spatially separated from i-cell populations positioned proximally in both these organs (fig. S14). It is possible that these alternative pathways may dominate over direct differentiation from the i-cell lineage in this system.

To address the developmental relationships between the different neural and nematocyte clusters and identify developmental markers, we assigned pseudo-time values to the cells and ranked genes in each trajectory (see Materials and Methods and Fig. 3A). This revealed trajectories consistent with these cell types both deriving from i-cells (Fig. 3, A and B) (*8*). Examination of the nematocyte trajectory revealed the early expression of genes previously not associated with this process, including *Znf845* and *Mos3* (Fig. 3, C and D) (*45*). Nematocyst-related genes—such as minicollagens, polyglutamate synthases, *Dkk3*, and *NOWA*—were then expressed during a first major phase of nematogenesis, consistent with previous reports (Fig. 3, C and D; corresponding expression domains in Fig. 3E and fig. S14) (*35*–*37*, *41*). The trajectory analysis confirmed continuity between the "nematoblast" clusters and the distinct and underappreciated nematocyte differentiation phase, characterized by expression of putative nematocil structural proteins and nematocilin expression at the end of the trajectory (Fig. 3, C and D, and figs. S11 and S14) (see above). The two phases of nematogenesis were linked by the expression of rare specific marker genes for cluster 17 (e.g., M14 peptidase in Fig. 3E and fig. S14). Consistent with this linking of the nematoblast and differentiation phases revealed in trajectory analysis, in situ markers showed distinct expression territories in the tentacle bulb and tentacle, respectively (Fig. 3E). Furthermore, we found that markers of both phases and their respective orthologs, including the "hair cell" gene set, were appropriately distributed among transcriptomes derived from dissected *Clytia* bulb and tentacle regions (*35*) and between developing and mature nematocyte scRNA-seq clusters in *Hydra* (*8*).

Cnidarian nervous systems represent both valuable points of phylogenetic comparison and tractable platforms for systems neuroscience (*3*, *8*, *20*). However, the molecular heterogeneity of neural cell types and their developmental progression remains largely unexplored, particularly in the more complex medusa forms. We therefore extracted genes expressed during neural development that included those encoding *bHLH*, *Sox*, and other transcription factors with potential roles in neurogenesis or fate specification and numerous other genes of interest in neuronal development, such as cell adhesion molecules (Fig. 4, A to C, and table S5) (*36*, *46*–*48*). In mature neurons, neuropeptides are thought to be the dominant neurotransmitters in cnidarians (*49*, *50*) but are challenging to identify because of rapid sequence evolution (*51*, *52*). In conjunction with sequence-based analysis, we were able to identify 10 new likely neuropeptides on the basis of their inclusion as marker genes for the four basic neural clusters (6, 9, 26, and 31 in Fig. 2, B and D), increasing the number of predicted *Clytia* neuropeptides to 21 (table S3). Our pseudo-time ranking revealed that many of these predicted neuropeptides mark the later stages of neural cluster trajectories, likely defining distinct, mature neural subpopulations (Fig. 4D).

We extracted and reclustered the neural supergroup ("Neural;" Fig. 2, A and B) to characterize neural subtypes. This distinguished 14 subpopulations of neurons and a progenitor population, expressing cell cycle and conserved neurodevelopmental genes including the *bHLH* transcription factor Neurogenin (subcluster 0; Fig. 4D). Notably, the neuronal subpopulations show combinatorial neuropeptide precursor expression, often with a distinct and identifying

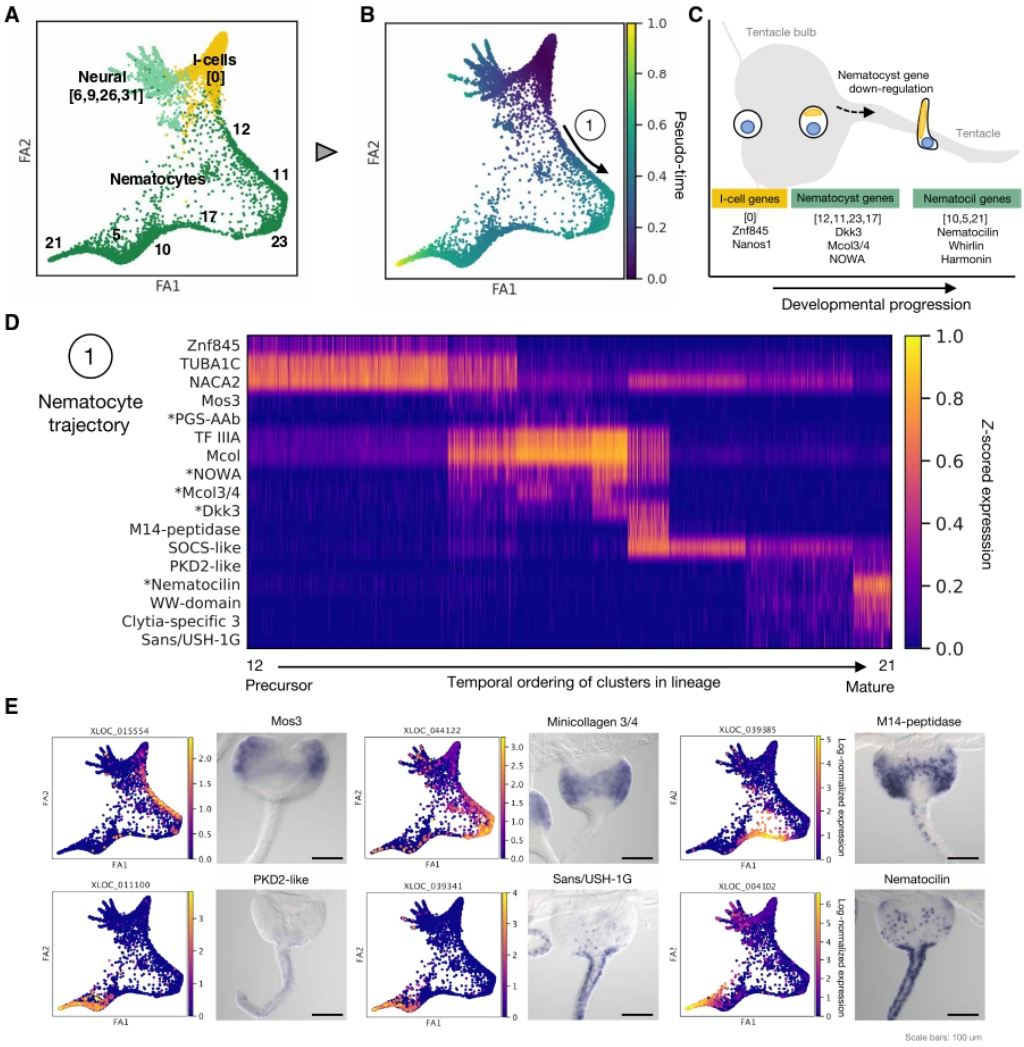


Fig. 3. Developmental and hierarchical analysis of nematocytes. (**A**) Subset of i-cell, neural, and nematocyte populations for trajectory analysis with respective clusters denoted. (**B**) Pseudo-time trajectories for differentiating nematocytes. (**C**) Diagram indicating the two main phases of nematogenesis with the corresponding "nematocyst" and "nematocil" stages and representative genes. (**D**) Heatmap of expression of dynamic genes in nematocyte development across the maturing populations. * denotes published nematogenesis genes. (**E**) Trajectory embeddings for selected nematogenesis markers and corresponding in situs. Scale bars, 100 μm.

neuropeptide (Fig. 4D). Expression of putative neuropeptide processing enzymes was detected across the subpopulations and in the nematocytes and gland cells (fig. S17). The UMAP (uniform manifold approximation and projection) expression embedding for neuropeptide precursors suggest that further complexity remains to be discovered within these subpopulations; for instance, subpopulation 4 includes cells either expressing pp1 or pp11. In situ hybridization for a set of neuropeptide precursors indicated that some were broadly distributed across the animal, while others had highly specific spatial locations, suggestive of distinct functions in regulating physiology and behaviors relating to swimming, feeding, and orientation (Fig. 4E and fig. S18). For example, $pp5^+$ (GRFamide precursor) cells were widely detected across the tentacles, bulbs, nerve rings, subumbrella, and mouth, while $pp11^+$ (GLWamide precursor) cells were detected predominantly in the manubrium and nerve ring. $Pp7^+$ cells were located both around the rim of the mouth and in patches of the nerve ring, close to the statocysts (vestibular organs). Subpopulations of neurons also occupied different expression domains within the tentacles: pp25, which generates distinct RFamide family neuropeptides (fig. S18 and table S3), distinguishes a subpopulation of the $pp5^+$ neurons positioned on the aboral side of the tentacle, pp17 labels distinct cells in the same region, and pp20 labels a small group of cells at the base of each tentacle (Fig. 4, E and F, and fig. S18).

Unlike cells using neuropeptides, the transcripts for which are directly assayable in scRNA-seq data, cells using classical chemical neurotransmitters are identified by enzymatic or transporter proxies.

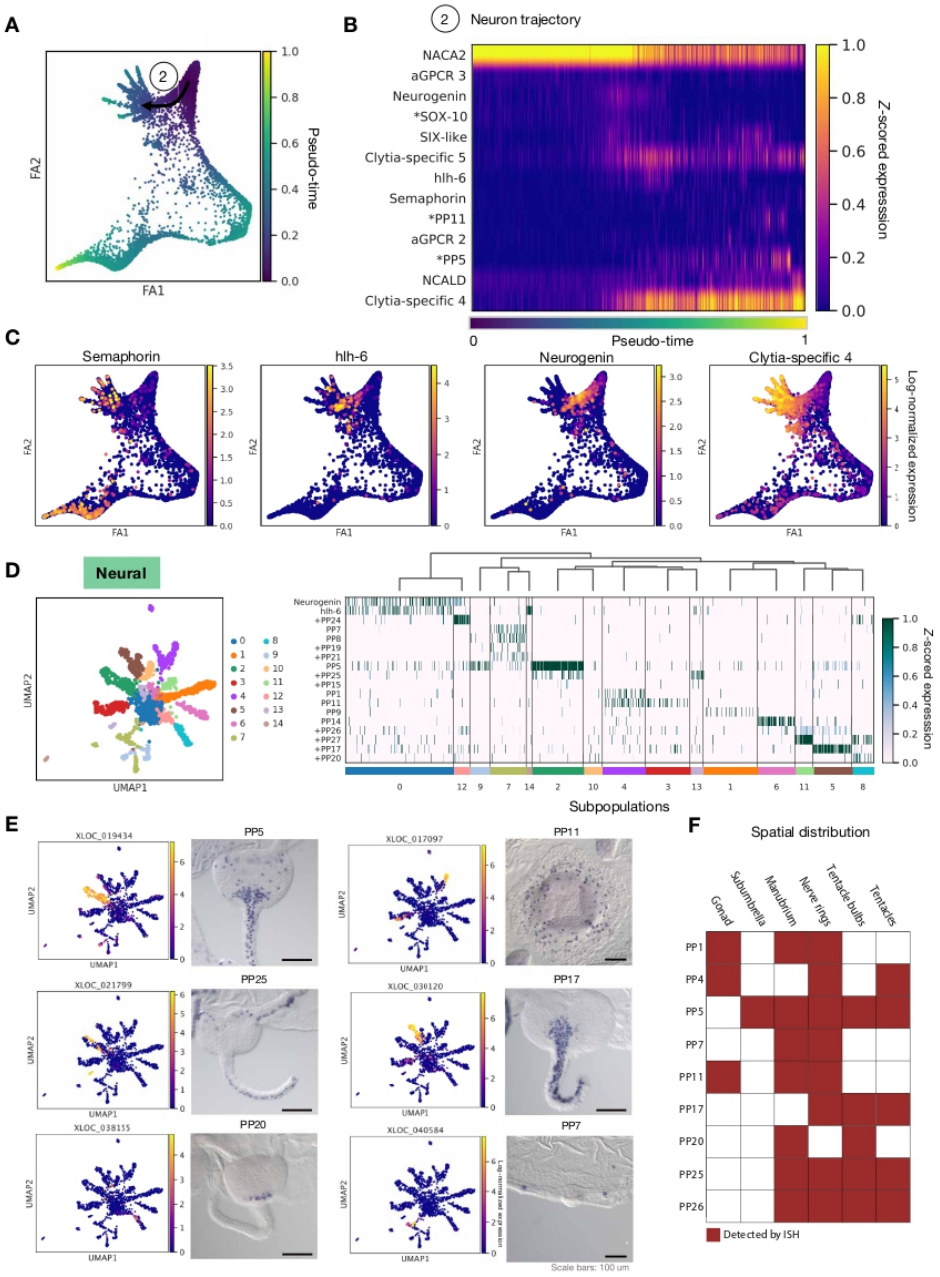


Fig. 4. Developmental and hierarchical analysis of neural cell types. (**A**) Pseudo-time trajectories for differentiating neurons. (**B**) Heatmap of expression of dynamic genes in neural development over pseudo-time. * denotes previously reported genes. (**C**) Embeddings show gene expression of markers for neural development across the neurons. (**D**) Diagram of the 14 neural subpopulations and progenitor cells along with a heatmap of neuropeptide markers and the *bHLH* transcription factors Neurogenin and Hlh6 across these subpopulations. Hlh6, a member of the Atonal superfamily, corresponds to *Nematostella* Ath-like (*82*). + denotes previously unpublished neuropeptide candidates. XLOC identifiers and accession numbers in table S3. (**E**) Embedding for neural subpopulation neuropeptide marker expression and corresponding in situs. Scale bars, 100 μm. (**F**) Summary of subpopulation marker expression between medusa territories.

Cells using glutamate as a neurotransmitter are usually inferred via the presence of vesicular glutamate transporter markers (vGluts; in human SLC17A6/7/8) (*53*). Of the closest *Clytia* homologs of human vGluts, one was expressed in neurons and non-neural cell types, and several were marker genes for nematocytes. However, in common with nearly all cnidarian genes annotated in silico as vesicular glutamate transporters, *Clytia* sequences lack an arginine residue, conserved in all bilaterian vGluts, and recently shown to be required for vGlut function (*54*). Furthermore, we detected no neuronal expression of the closest homologs of glutamate decarboxylase (GAD), a marker for gamma-aminobutyric acid-ergic (GABAergic) neurons in Bilateria. GAD homologs were instead detected in some gastrodermal subtypes. There were a number of other interesting neuron subtype-specific genes potentially involved in "chemical" neurotransmission, for example, a possible nitric oxide synthase, a choline (SLC5A7-like) transporter, a gene encoding a taurine dioxygenase–like domain, and a member of the Slc6 family of transporters (table S5).

Cell state shifts in response to starvation across the cell atlas

To assess the transcriptional impact of starvation, we mapped individual cells to their corresponding control or starved labels. As there are around 60% fewer cells in a starved animal (fig. S2), we first asked whether there were significantly different numbers of cells per cluster between control and starved conditions. We found that only one cluster had a significant difference (cluster 11, early nematoblasts in fig. S5B), suggesting a nearly uniform reduction across cell types in the starved condition. In contrast, the distribution of cells from control versus starved animals across the atlas embedding showed marked shifts in the local density of cells from the two conditions within most clusters (Fig. 5A).

Given the cell type resolution of the atlas, as determined operationally by clustering, we then asked how marked the transcriptional changes incurred by perturbation were in comparison to the transcriptional differences defining the cell types, i.e., are the perturbation-induced changes encompassed within these cell type designations or are they larger in magnitude. We thus compared distances between control and starved cells within clusters to the distances between clusters. As a metric, we used the L_1 distance (see Materials and Methods), the sum of the absolute differences between centroid coordinates in principal components analysis (PCA)–reduced space. We found that the L_1 distances between control and starved cells within a cell type, versus between cell types (regardless of condition), formed nearly nonoverlapping distributions (see Materials and Methods and Fig. 5B). This suggests that, overall, in *Clytia*, the transcriptional responses to starvation are defined by cell state shifts, and their cell type repertoire is well represented by the original clusters. However, the impact of starvation was variable across cell types, as reflected by the range of internal (state) distances (Fig. 5B). Starvation produced the largest perturbations in cells of the gastrovascular system, causing control-versus-starved distances large enough to overlap with the smallest inter-type distance, i.e., that between the stem cells and nematocyte precursors (Fig. 5B). This distinction between state shifts and type was also clearly visible in the lack of overlap between the distributions of inter- and intracluster distances within the second multiplexed experiment (fig. S6E). Although classification and distinction of cell state and type is a complex task (*55*), this analysis, based on relative distance in transcriptional space, provides a quantitative basis for delineation of type/state effects that may be useful in other contexts. We additionally validated the ability of this

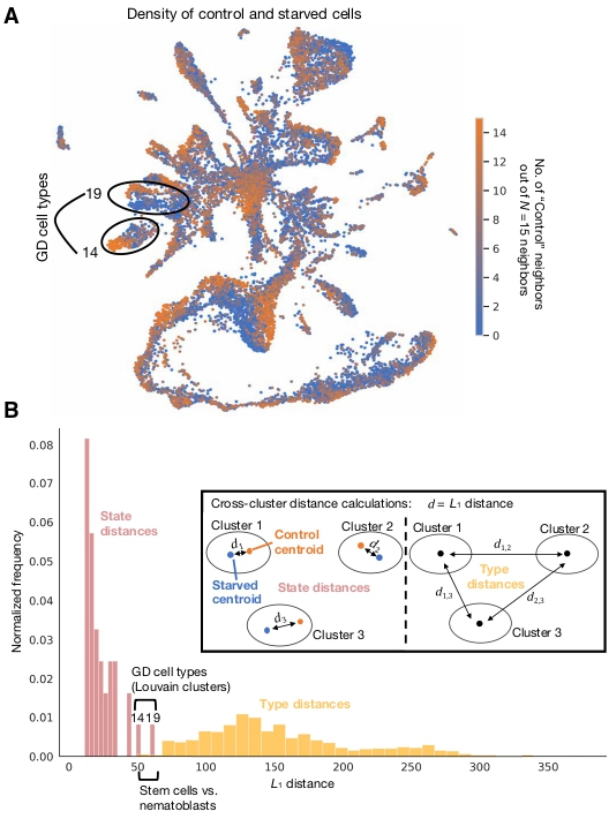


Fig. 5. Starvation-induced shifts in cell types. (**A**) Density map of control and starved cells across cell types (see cell types across conditions in fig. S5B), with each cell colored by the number of its 15 nearest neighbors that are control cells to highlight the clusters most affected by starvation, i.e., most clearly separated from their control neighbors. Clusters 14 and 19, with particularly dense and separated control/starved regions are circled. (**B**) Histogram of L_1 distances between centroids of control and starved cells within cell types versus pairwise L_1 distances between centroids of the 36 cell types. Clusters 14 and 19, with largest internal distances, are highlighted, and clusters with small pairwise distances are also noted for comparison. Inset illustrates inter- and intra-L_1 distance calculations.

method to recapitulate the magnitude of state shifts in response to graded stimuli and state-versus-type distinctions, on two other published, multiperturbation datasets (see Materials and Methods; fig. S19).

To characterize gene-level responses underlying these starvation-induced shifts, we then asked whether responses are shared or unique across the cell types and compared the extent of the responses, in terms of gene quantity and expression level, across the atlas. For each cell type, we collected genes that were differentially expressed under starvation ("perturbed genes"; Fig. 6A) (see Materials and Methods). For a high-level view of the general functions and processes affected by starvation and their cell type specificity, we clustered perturbed genes into apparent "gene modules" (*56*) by their patterns of coexpression across cells (see Materials and Methods and Fig. 6A). We assigned putative functions to these gene modules through Gene Ontology (GO) term enrichment, giving a global view of affected processes (fig. S20), and examined the distribution of cell

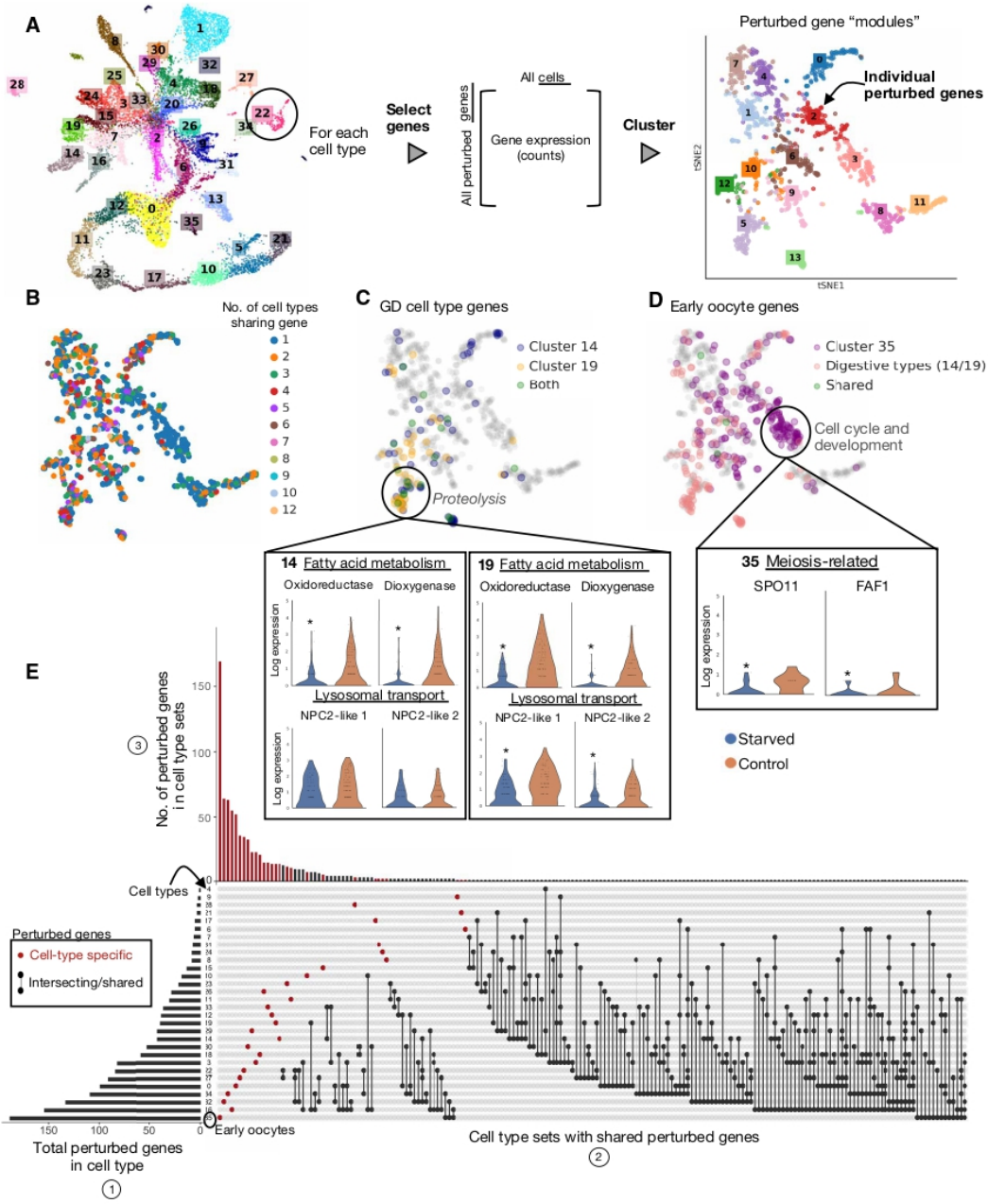


Fig. 6. Perturbation responses in the oocyte and digestive cell types. (**A**) Workflow for extracting perturbed genes per cell type and clustering on genes to extract coexpressed groups or modules of genes (fig. S20). (**B**) Each gene in module embedding colored by number of cell types in which the gene was differentially expressed under starvation. (**C**) Visualization of locations of perturbed genes (among the embedded modules) of the two highly perturbed GD cell types (clusters 14 and 19) and (**D**) early oocytes (cluster 35). Violin plots showing expression profiles for several perturbed (differentially expressed) genes in "functional" categories of interest. *$P < 0.05$ from nonparametric Wilcoxon test. Horizontal lines show quartiles and width of violins denote density of points in that region. (**E**) UpSet plot visualization for intersecting sets of perturbed genes. The barplot in (1) shows the number of genes that are differentially expressed under starvation (perturbed genes) for each cell type. This is the cardinality or size of each cell type's set of perturbed genes. (2) The connected dots represent intersections (overlaps) between the sets of each cell type's perturbed genes. Genes with starved differential expression in only one cell type are denoted by the red, unconnected dots. (3) The top bar plot is the number of perturbed genes within these sets of shared perturbed genes, i.e., the cardinality of the intersections in (2).

types across modules by asking in how many cell types is a given gene a perturbed gene (Fig. 6B). We found that certain gene modules were broadly shared across cell types, while others were almost entirely cell type–specific (fig. S20). Notable examples include gene module 5, which is enriched in proteolytic genes (Fig. 6C and fig. S20) and has shared expression across multiple GD cell types (Fig. 6C). Notably, there is also divergent gene expression between GD types (Fig. 6C). In comparison, gene module 3 is largely composed of early oocyte gene expression (~70%) and is enriched in cell cycle and developmental genes, which are commonly enriched in growing oocytes (Fig. 6D). Changes in expression of these genes likely reflect the processes of oocyte phagocytosis activated in the gonads of starving animals (see below). Overall, these modules thus provide an overview of which processes affected by starvation are shared across cell types and reveal divergent expression potentially reflecting different motility (*19*), locations, or as-yet-undescribed functional differences between the cell types (fig. S21).

To examine how individual perturbed genes are distributed across cell types, we visualized, for each cell type, how many perturbed genes it had, and how many of these genes are unique versus shared with other cell types (Fig. 6E). We found a large number of perturbed genes (~72%) were cell type specific (Fig. 6E). For the most perturbed cell types, we examined whether the state shifts that we had observed were due to changes in a large or small number of genes, and how highly these genes were expressed. Consistent with the marked shrinkage of the gonads during starvation treatment (Fig. 1), early oocytes contained the highest number of perturbed genes, which were spread across many gene modules (Fig. 6D). In contrast, the GD cell types had fewer perturbed genes that were expressed at higher levels and localized to more specific gene modules (Fig. 6C), highlighting the diverse logic used by cell types under starvation (fig. S22).

In accordance with these distinct responses in GD cells and oocytes, comparison of the cellular organization of gonads from control and starved medusae revealed major reorganization of both the gastrodermis and the oocyte populations (Fig. 7). Most notably, the population of midsized, growing oocytes, which progress daily through vitellogenesis under conditions of normal feeding (*29*), was largely depleted following starvation, leaving a majority of previtellogenic oocytes (Fig. 7A). A sparse population of large oocytes in starved gonads likely results from growth of a minor subpopulation of oocytes fueled by recycling of somatic tissue and oocytes (disintegration and phagocytosis of smaller oocytes visible in Fig. 7A, asterisks). Consistently, GD cells in many parts of the gonad lost their regular epithelial organization and, despite the absence of any external food supply, showed evidence of active phagocytosis involving variably sized vesicles (arrows in Fig. 7A). Changes in organization and activity of the gonad gastrodermis were also evident from in situ hybridization images for the GD cell marker CathepsinL, while reduced expression was confirmed for a protease (*ShKT-TrypA*) expressed in gland cell types A and B positioned within the manubrium gastroderm, which is down-regulated during the starvation treatment (Fig. 7B). Shifts between gonad gastrodermis organization and transcriptional profiles induced by starvation thus accompany activation of tissue autodigestion programs and likely the mobilization of GD cells [termed MGD for mobilizing gastro digestive cells (*19*)] from the gonad through the gastrovascular canal system, which has been observed both under conditions of starvation and during regeneration of the feeding organ (*19*).

DISCUSSION

The *Clytia* medusa single-cell atlas presented here is an important addition to the growing number of single-cell atlases across the animal tree of life. It is available for easy interrogation by the community through the UCSC Genome Browser (see Data and materials availability). This provides the first cell-level transcriptomic characterization of a pelagic medusa stage, the most complex of the life cycle forms within the large and diverse phylum Cnidaria. Reflecting this complexity, we found greater cell type diversity in the *Clytia* medusa than in its polyp-only hydrozoan cousin *Hydra* (*8*). The outer, epidermal body layer could be subdivided into seven clusters encompassing all of the described *Clytia* muscle types, including two types of fast-contracting striated swimming muscle (*23*, *27*). Rich diversity was also uncovered in the inner gastroderm layer, which is elaborated in the medusa into distinct digestive compartments (mouth, stomach, gonad, and tentacle bulb) and generates the thick mesoglea (jelly) characteristic of the medusa form. Of the eight gastroderm cell clusters, four could be mapped to distinct sites by marker gene in situ hybridization including three likely involved in mesoglea/extracellular matrix production and thus in modulating the medusa structure. Only two of the four clusters belong to the "GastroDigestive" cluster set, characterized by a largely shared transcriptomic profile (Fig. 2B and table S5), with the other GD clusters distributed more uniformly across the digestive compartments (fig. S14), suggesting a dominant role in intracellular digestion but with functional specializations that remain to be fully understood. Our starvation experiment analyses revealed that these clusters were maintained operationally as distinct "cell types" rather than "cell states" between the two extreme conditions tested, but we cannot rule out that responses to other environmental or physiological perturbations may reveal plasticity between these clusters; for instance, transdifferentiation between muscle and nerve cell types is well documented in hydrozoan medusae [overview in (*27*)].

In addition to the epithelial cell types of the epidermis and gastroderm, our single-cell atlas confirms the presence of an i-cell population in *Clytia* providing a similar set of somatic cell types to that described in *Hydra*, as well as the germ cells (*8*, *34*). In these medusa data, we do not find strong evidence for direct progression from i-cells to gland cells or for the shared neural-gland cell progenitors described in (*8*). In contrast, our pseudo-time analyses provide transcriptional signatures of the progressive stages of nematogenesis and neurogenesis from i-cells that will guide future studies of their developmental regulation. The large representation of nematogenic stages in this *Clytia* medusa scRNA dataset allowed us to link two distinct phases of nematocyte formation with extremely different transcriptional profiles. The initial phase covering nematocyst formation has been the focus of many studies (*35*–*37*, *41*), but the terminal phase has been largely overlooked in previous transcriptomics studies, likely due to the relatively low mRNA content (*8*, *35*) and the extremely abrupt degradation of nematocyst-related mRNAs before the terminal phase (*37*). We uncovered 14 mature neuronal subtypes in *Clytia*, which is similar to the number reported in *Hydra* and *Nematostella* (*3*, *8*). It is likely that further heterogeneity exists within these 14 subpopulations. Spatial expression analysis of neuropeptides that contributed to the signatures of one or more subpopulations revealed a wide variety of neuronal populations either associated with specific anatomical structures—such as the tentacles, nerve rings, and manubrium—or distributed across the medusa (Fig. 4, D and E, and fig. S18). How molecular cell type maps to

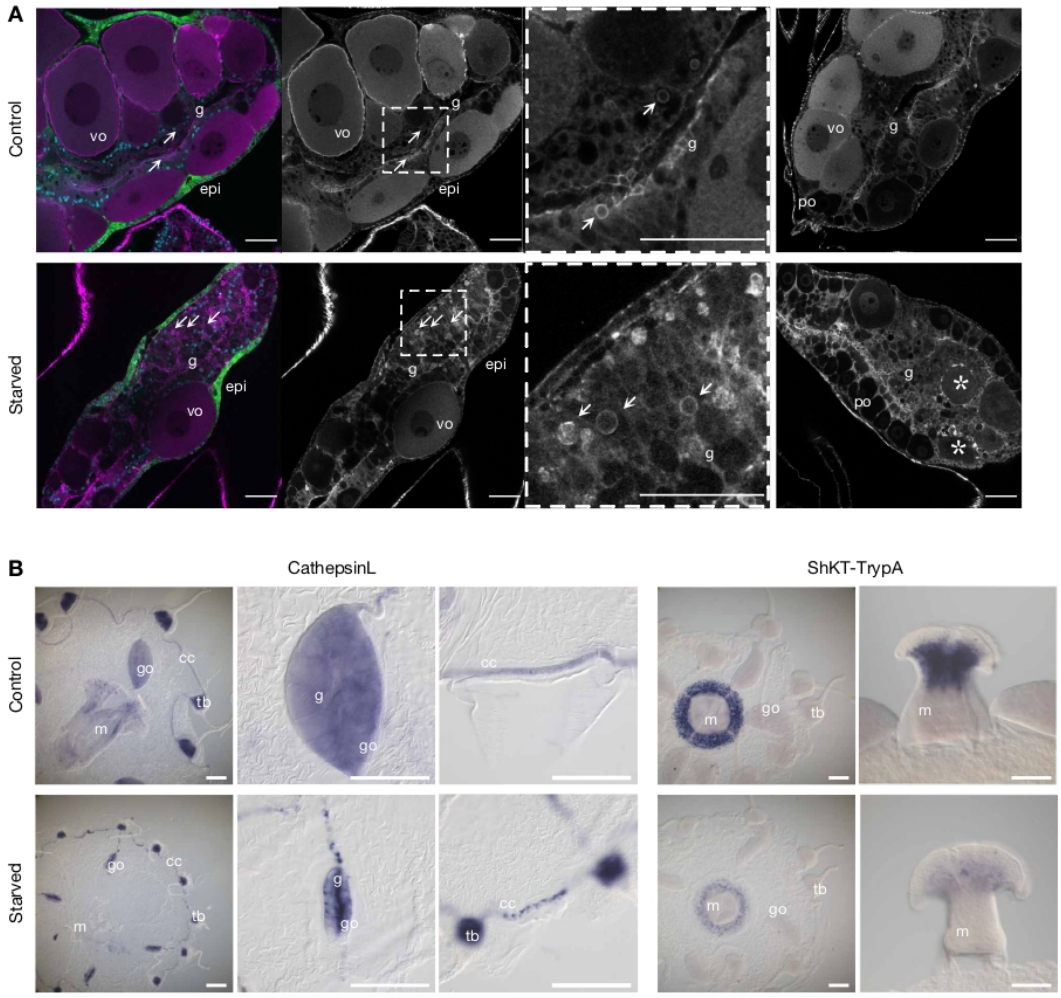


Fig. 7. Perturbations to gastroderm cell types in response to starvation. (**A**) Confocal sections through gonads from control and starved medusae with cell morphology revealed by phalloidin staining of cell boundaries (magenta/gray). The first panel of each row shows costaining of nuclei with Hoechst (blue) and endogenous GFP4 (green) in the outer epidermis (epi). Vitellogenic oocytes (vo) are largely absent during starvation, leaving mostly previtellogenic oocytes (po). The gastroderm (g) is heavily reorganized, with evidence of active phagocytosis (vesicles arrowed) and disintegrating oocytes (asterisks). The third panel in each row is a higher magnification of the boxed area in the second panel, and the fourth panel shows a second example gonad for each condition. Scale bars, 50 μm. (**B**) In situ hybridization visualization of example cell-types (left) and genes (right) affected by starvation. The general GD cell marker CathepsinL confirms extensive reorganization of the digestive gastroderm in starved animals, especially its collapse into central regions of the strongly reduced, oocyte-depleted, gonad, and fragmentation in the circular canal; expression of the gland cell A/B marker "ShKT and trypsin domain protein A" (ShKT-TrypA) in the manubrium consistent with its strong down-regulated following starvation. Gonads, go; manubrium, m; tentacle bulbs, tb; gastroderm, g; circular canal, cc. Scale bars, 200 μm.

function both within and across body parts, the roles of these peptides as primary transmitters and/or neuromodulators, and the uses, if any, of classical, small-molecule neurotransmission remain unknown. Moving forward, with this cell atlas as the foundation, the ability to perform whole-organism, multiplexed scRNA-seq, in combination with emerging genetic tools and advantageous life history traits, makes *Clytia* a powerful, tractable platform for high-resolution systems biology.

This work further serves as a case study in using multiplexed single-cell transcriptomics to assess cellular responses to whole-organism perturbations and provides a guide for deployment in other organisms. We anticipate that whole animal multiplexed scRNA-seq (WHAM-seq) will benefit researchers studying various biologies, from developing embryos to organoids to nonmodel organisms. The techniques for multiplexed experimentation that underlie this study are also well suited to large-scale perturbation studies (such as temperature, pH, or other environmental disturbances) in other marine organisms given the SW compatible workflow. Although the inclusion of multiple animals and conditions may currently limit the detection of very rare cell populations (fig. S1), as sequencing costs drop and cell throughput in scRNA-seq grows, WHAM-seq should become tractable for larger, more complex systems. The lack of library-induced batch effects

demonstrates how large-scale experiments can be conducted without introduction (or minimizing introduction) of confounding factors from multiple experiments, which can be highly nonlinear and difficult to account for (*13*). The second perturbation dataset also demonstrates both how batch effect variability is reduced within multiplexed experiments (see Materials and Methods; fig. S11) and the utility of this multiplexed approach in discerning cell type–specific activity markers (see Materials and Methods; fig. S11).

The fully reproducible framework we have presented, which includes code that can be run on a laptop or for free in the cloud (see "Code availability" in Acknowledgements), will further assist in extending this expression-based analysis to other organisms. By relying on expression, our strategy reduces the reliance on prior gene functional annotation using specificity of expression to identify genes of interest, allowing for targeted annotation. This includes determination of strong diagnostic markers for cell type definition, cell type–specific and shared transcriptional responses to starvation, and "modules" of coexpressed genes underlying these responses. The extent of these expression-based changes additionally highlights areas of the organism's biology that are strongly or uniquely affected by a perturbation. By applying simple and interpretable quantitative analyses to the various cell type–specific perturbation responses, we revealed the large-scale down-regulation of gene expression in two GastroDigestive cell types and severe disruption of oocyte development under starvation. Together, this approach markedly lowers the barriers for working with nontraditional models and affords opportunities to match uniquely suited organisms to specific questions. Moving forward, the combination of scRNA-seq and other sequencing-based genomics techniques with multiplexing and annotation-agnostic analyses could foster comprehensive high-resolution molecular studies of diverse organisms and their responses to numerous environmental perturbations.

MATERIALS AND METHODS

All animal experimentation performed did not require Institutional Animal Care and Use Committee approval.

Animal culture and experimental setup

Starvation

Culture of the *Clytia* life cycle was carried out as previously described (*57*), with some modifications to the tank design. The system used in this study to culture polyps uses zebrafish tanks (Pentair), with polyp slides held in glass slide racks (Fisher, catalog no. 02-912-615). Medusae used in this experiment were raised in 4-liter beakers with a circular current generated by stirring with a constant speed of 5 rpm DC motor (Uxcell), attached to the lid of a multiwell tissue culture plate. Artificial SW for culture and experiments was made using Red Sea Salts (Bulk Reef Supply, catalog no. 207077) diluted into building deionized (DI) water to 36 parts per thousand. Experiments used ~1-cm female medusa of the Z4B strain.

For the experiment, baby medusae were collected overnight and then cultured together until they reached ~1 cm (about 2 weeks). Animals were fed once per day using 2- to 4-day old brine shrimp. Before the experiment, animals were split into two beakers. Feeding continued as before for one beaker, while the other was starved for 4 days. The "control" group was not fed on the day of the experiment. The 4-day time point was chosen as animals show strong phenotypic changes (Fig. 1), but it is far from their survival limit following starvation, as *Clytia* medusae can survive for more than 3 weeks with no food.

Stimulation

Rearing of the medusae before the experiment was performed as described above. On the day before the experiment, each *Clytia* medusa (3 to 5 weeks old) was placed in a separate container (~150 ml of SW), which was covered with foil and moved to the experimentation area to acclimate overnight in the dark. The morning following overnight acclimation, lights were turned on for ~2.5 hours to allow for spawning to complete (*58*). Each animal was then given repeated bouts of stimulation over a period of 30 min, with each stimulus administered every 2 min. One hundred microliters of each stimulant [150 mM KCl, DI water, or SW as a control] was gently added just below (or just above for KCl) each medusa by pipette (fig. S7A). Stimuli were chosen on the basis of their ability to reliably induce crumpling behavior, a protective response in which the bell is drawn in toward the mouth using the radial muscle (*59*). Animals were dissociated 30 min following the last stimulation.

Single-cell suspension and multiplexing

For the starvation experiment, animals were washed with hypertonic phosphate-buffered saline [PBS; 500 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4 (pH 7.4)] by serial transfer from SW through three successive containers each with 150 ml of hypertonic PBS to prepare cells for fixation and to avoid precipitation of SW salts in methanol. The animals were then thoroughly homogenized with a dounce. After homogenization, cells were collected by centrifugation at 500*g* for 5 min and resuspended in 100 µl of hypertonic PBS. Cells were then fixed by addition of 400 µl of ice-cold methanol and stored at −80°C until sample indexing and library preparation. Each sample was labeled according to the ClickTag labeling procedure described previously (*14*). ClickTags used for each animal are outlined in table S2. Each sample was labeled with two distinct and unique ClickTags, and samples were then pooled after addition of the "blocking" oligo. Control and starved samples were counted on a Countess, both to estimate cell numbers per animal and to determine concentrations for 10× loadings. A total of 200,000 cells/ml were counted for the starved sample and 1,000,000 cells/ml for the control. Around 120 to 150,000 cells of the starved and 200,000 cells of the control were then pooled. A loading of 20,000 pooled cells was then used as input into two lanes of the 10X Chromium Controller with v2 chemistry. Sample tag libraries were separated and processed after an SPRI size-selection step as previously described (*14*). cDNA samples were run on two lanes of HiSeq 4000 (two HiSeq 3000/4000 SBS 300 cycle kits), and tag libraries were run on two lanes of MiSeq (using MiSeq v3 150 cycle kits).

For cell counting (fig. S2), the starvation experiment was repeated with four animals per condition with animals dissociated and resuspended in 500 µl of SW, and then cells were counted on a hemocytometer (InCyto DHC-B02) with a 10× objective. Two 16-square grids were counted per sample.

The same protocol was followed for fixing and labeling cells from the stimulation experiment. Each animal in this case was assigned one unique ClickTag and one ClickTag per condition (table S2). After the separation of the cDNA and ClickTag samples, ClickTags were added at a 3% final concentration to the cDNA samples sequenced on the HiSeq, in addition to separate ClickTag sequencing on the MiSeq. The full protocol is described in Supplementary Methods.

In situ hybridization

Colorimetric in situ hybridization (Figs. 2C, 3E, and 4E) was performed as previously (*60*) with minor modifications. Briefly, 2-week-old medusae (Z4B strain) were relaxed in 0.4 mM menthol in SW, and tentacles were trimmed before fixation in a prechilled solution of 3.7% formaldehyde and 0.2% glutaraldehyde in PBS on ice for 40 min. Specimens were then washed thoroughly with PBST (PBS + 0.1% Tween 20), dehydrated in methanol stepwise, and stored in 100% methanol at −20°C. Hybridization (at 62°C for 72 hours) and washing steps were performed with a robot (Intavis AG, Bioanalytical Instruments) using 20× saline-sodium citrate pH adjusted to 4.7 throughout. Acetylation steps using 0.1 M triethanolamine in PBST (2× 5 min) and then 0.25% acetic anhydride in 0.1 M triethanolamine (2× 5 min) followed by PBST washes (3× 10 min) were included before prehybridization to reduce probe nonspecific binding. Incubation with 1:2000 anti-DIG (digoxigenin)AP in 1× blocking solution was performed for 3 hours before washing and the nitroblue tetrazolium–5-bromo-4-chloro-3-indolyl-phosphate (NBT-BCIPP) color reaction at pH 9.5. Following postfixation, washing, and equilibration of samples in 50% glycerol/PBS, images were acquired using a Zeiss Axio Imager A2.

Probes were generated by polymerase chain reaction (PCR) from cDNA clones corresponding to our expressed sequence tag collection (*61*) or from medusa cDNA; the Elav probe was synthesized as a gBlock by Integrated DNA Technologies (details in table S3). For probes against the Elav, the T3 polymerase recognition site (AATTAACCCTCACTAAAGGG) was added to the 3′-end of the PCR product, or gBlock, respectively. Products were TOPO cloned (Thermo Fisher Scientific, catalog no. K280020) and sequence verified. All probes were labeled with DIG RNA labeling mix (Sigma-Aldrich, 11277073910) and purified with ProbeQuant G-50 Micro Columns (GE Healthcare Life Sciences, catalog no. 28-9034-08).

Confocal microscopy

Visualization of cell morphology within the gonads of control and starved young adult female medusae (Z4B strain) by confocal microscopy was performed as previously (*58*). Fixation used 4% EM-grade paraformaldehyde in 0.1 M Hepes (pH 6.9)/50 mM EGTA/10 mM $MgSO_4$/80 mM maltose/0.2% Triton X-100 for 2 hours at room temperature. Specimens were washed 3× for 15 min with PBS/0.02% Triton and 3× for 5 min in PBS. Cell boundaries and nuclei staining were performed by overnight incubation in 1:50 rhodamine-phalloidin (1 mg/ml; Molecular Probes) and 1:5000 Hoechst 33258 (1 mg/ml stock; Sigma-Aldrich) in PBS. Samples were washed 3× for 15 min with PBS/0.02% Triton and 3× for 5 min with PBS and equilibrated in 50% PBS/Citifluor (Citifluor AF1) before imaging using a Leica SP5 confocal microscope. Control medusae were fixed 24 hours after the last feeding, while starved ones were fixed 4 days after the last feeding.

Generation of reference transcriptome

Assignments to PANTHER database entries (version 11) were made using the "pantherScore2.0.pl" script available from the database website (*62*). Human-*Clytia* orthologs were assigned using the OMA (orthologous matrix) program (*63*) as described in (*21*) and taken from the pairwise human/*Clytia* orthologs output, rather than the orthologous groups.

To ensure highly sensitive transcriptome alignment, a new transcriptome assembly for *C. hemisphaerica* was generated from bulk RNA-seq data produced from *Clytia* medusae (organisms at the same life stage as in the single-cell experiments) (www.ncbi.nlm.nih.gov/sra/ERX2868482%5Baccn%5D). We used the Trinity (*64*) de novo assembler, with default parameters, to generate a transcriptome (http://dx.doi.org/10.22002/D1.1824) and the Cufflinks Cuffcompare utility (*55*) to merge the Trinity assembled transcripts with any XLOC annotations from the MARIMBA v.1 (created on 30 May 2016; http://dx.doi.org/10.22002/D1.1830) transcriptome assembly (*21*). Then, with the CD-HIT (clustering method), assembled sequences with at least 95% were clustered and only one representative sequence was kept for each cluster. With the published MARIMBA v.1 genome sequence as a reference (created on 30 May 2016; http://dx.doi.org/10.22002/D1.1828) (*21*), the GMAP (genomic mapping and alignment program) aligner converted the collapsed Trinity fasta records to gff3 coordinate file (http://dx.doi.org/10.22002/D1.1824). Most differentially expressed genes found in scRNA-seq data from this study were previously identified and annotated (fig. S13). This annotation was used for the preprocessing and quantification of scRNA-seq described below. Protein sequences were then obtained by running TransDecoder (*65*) with default settings for the Trinity transcriptome (http://dx.doi.org/10.22002/D1.1827).

Preprocessing and clustering of sequencing data

Initial cell ranger demultiplexing for ClickTags

Initial demultiplexing of ClickTag libraries was done using output from the 10X Cell Ranger pipeline using Cell Ranger 3.0 *count* and *aggr* functions with the MiSeq ClickTag fastqs as input () and combining the counts from the two lanes using the denoted sample IDs. A ClickTag count matrix (cell-by-ClickTag) was generated by counting ClickTag barcodes that had high sequence similarity to the designed sequences using the Python fuzzywuzzy package to identify targets within Levenshtein distance 1. We additionally quantified gene expression with the kallisto-bustools workflow (*66*), which reproduced concordant results described below.

Initial cell ranger demultiplexing and clustering for cDNA

Initial processing of starvation cDNA libraries was performed with the 10X Cell Ranger pipeline using Cell Ranger 3.0 *count* and *aggr* functions to align and quantify the HiSeq reads and combining the counts from the two lanes using the denoted sample IDs. This was followed by filtering cells for the high-quality cells chosen during ClickTag analysis, in addition to filtering cells by thresholding the rank-unique molecular identifier (UMI) versus cell barcodes plot. Values were log1p-normalized, mean-centered, and scaled for downstream dimensionality reduction and visualization using Scanpy (*67*).

We then conducted Louvain clustering (*68*) on the data mapped to a lower dimensional space by applying PCA to the expression data filtered for highly variable genes, initially using Scanpy's filter_genes_dispersion on only the log-normalized data. This resulted in the identification of 36 clusters (Fig. 2B and figs. S9 to S11), which we also refer to as cell types. The marker genes were selected by analyzing the top 100 markers extracted by Scanpy's rank_genes_groups using default settings (*P* values adjusted with the Benjamini-Hochberg method for multiple testing). The clusters were annotated and validated with marker genes previously identified in the literature and manually categorized into the seven classes in Fig. 2A based on the marker gene patterns and their functional annotations.

Kallisto bustools for demultiplexing and clustering: Standardization of workflow

To integrate and update the analysis using a platform with streamlined ClickTag demultiplexing and count matrix generation workflows,

we used the kITE demultiplexing protocol, which is based on the kallisto-bustools workflow and is described in the ClickTag demultiplexing protocol (*14*). Briefly, MiSeq reads are aligned to possible tag sequences (Hamming distance 1 away from designed oligo sequence whitelist) by building a kallisto index and pseudo-aligning reads to this index. Counts for these sequences were then collapsed into counts for their respective ClickTags, creating a cell-by-tag count matrix. We used Louvain clustering of cell barcodes based on the observed ClickTags to filter for clearly delineated cells (clusters strongly marked by the individual's two corresponding tags) and exclude sample doublets (*14*). We also followed similar preprocessing to standard cell-by-gene workflows using the inflection point in rank-UMI versus cell barcodes knee plots to filter cell barcodes based on their tag UMI counts. We do find, similar to findings in the original ClickTag multiplexing publication (*14*), that the number of ClickTags per fed cell is higher than for the starved counterparts, possibly supporting the previous observation of ClickTag number per cell increasing with cell size (fig. S4E, shown for oocytes).

For the stimulation experiment, we concatenated sequencing data from the MiSeq and HiSeq as input to the previously described kallisto-kITE workflow. With the same clustering procedure, we selected cell barcodes in clusters with strong overlapping expression of both individual and condition ClickTags.

To standardize the cDNA analysis workflow, we reprocessed the starvation data using the kallisto-bustools workflow to generate gene count matrices for each lane, which were then concatenated. Cells were also filtered on the basis of the ClickTag analysis. Values were log1p-normalized, mean-centered and scaled, and filtered for highly variable genes using the same procedure described above for downstream dimensionality reduction and visualization (e.g., PCA) using Scanpy. We found that with the kallisto-processed data (with the same Cell Ranger clustering applied to the cells), the top 100 markers for each of the 36 clusters determined with Scanpy's rank_genes_groups function (table S5) (using the nonparametric Wilcoxon test) overlap with markers in the Cell Ranger expression data, verifying that the cluster labels were concordant (fig. S10). To ensure that we additionally detected low expression marker genes, we extracted low expression markers in table S5, which consist of genes with at least 10 counts over all cells and with 90% of those counts deriving from the same cell type.

We used the PAGA partition-based graph abstraction method (*31*) to generate an underlying graph representation of the connectivity between cells (determining connectivity by the number of inter-edges between cell groups compared to the number of interedges under random assignment). We then generated a two-dimensional UMAP (*32*) embedding initialized with the PAGA graph structure for cell atlas visualization (Fig. 2A) using Scanpy. For Fig. 2B, 100 cells were randomly subsampled from each cell type to generate the heatmap.

The stimulation cDNA data were processed with the same kallisto-bustools workflow and commands as the starvation experiment data. We initially used Louvain clustering to also filter low UMI count clusters that were then removed from downstream analysis.

Distance-based comparative analysis of clusters

We first used L_1 distances between starved and control cells (within each cell type) to assess how comprehensive our cell type designations were (Fig. 5). Centroids for a given cell type were calculated for starved and control cells separately in PCA-reduced space (60PC coordinates for each cell as opposed to the raw gene expression matrix). The centroid vectors are represented as c_s and c_c for the starved and control cells, respectively. The L_1 distance (d) between them was calculated as the sum of absolute difference between the centroid coordinates

$$d = \sum_{i=0}^{60} | c_{s_i} - c_{c_i} | \quad (1)$$

These intracluster distances were then compared to the pairwise L_1 distances between the cell types, with centroids calculated for all cells in a given type (c_1 and c_2 for cell types 1 and 2) in the same manner as described above. The L_1 distances were then calculated for all possible pairs of these cell type centroids using (Eq. 1). The distributions of the inter- and intracluster distances are shown in (Fig. 5B). We chose to use the L_1 distance metric as it tends to better retain relative distances in high dimensions, particularly in comparison to the commonly used Euclidean distance or other higher L-norms (*69*, *70*).

We validated this method of transcriptional distance measurement on the multiplexed perturbation of mouse NSCs in (*14*) and the multiple immunomodulatory drug treatments across heterogeneous cell populations in human peripheral blood mononuclear cells (PBMCs) in (*71*). For (*14*), we calculated centroid distances between perturbed and control populations, in PCA space (*14*), as well as pairwise distances between individual perturbed and control cells in (*14*) to demonstrate the ability of the L_1 distance to recapitulate the graded response to a perturbant, with larger distances representing greater perturbant impact (with respect to the control cells) (fig. S19, A and B). We additionally compared all pairwise L_1 distances between perturbed and control cells within the monocyte population or "T cell" PBMC populations in (*71*) using the oNMF space. We then compared these measurements against the magnitude of cell type distances (between control monocyte and T cell populations) to highlight "state"-versus-"type"-level transcriptional differences (fig. S19C). To then extract the most perturbed monocyte populations, we calculated centroid distances between the centroid of each perturbation condition and the CTRL1 (control) condition to rank populations by perturbation "distance" (fig. S19D).

We then used the stimulation dataset to assess the validity of the clusters/cell types generated with the starvation data. We created a joint representation of the two datasets by using a concatenated cell-by-gene matrix including only the genes highly variable in both the control and starved datasets and used 70% of the starvation data to train a k-nearest neighbor (KNN) classifier (using sklearn's KNNClassifier with $k = 15$) to assign cluster labels to the remaining starvation cells and to the stimulation dataset. This showed that the stimulation cells' labels from their neighbors in starvation data were assigned at the same accuracy as the test starvation data, meaning that clusters from the starvation data are applicable to the stimulation dataset and capture the main features of the stimulation dataset to the same extent (fig. S6A).

We also examined batch effects in the stimulation experiment. As with the starvation experiment, the L_1 metric was used to visualize the magnitude of batch effect within the multiplexed experiments compared to between experiments (fig. S8). We used the merged representation (used for the KNN assignment) between both experiments to find the average pairwise distances between cell types of control condition individuals within the starvation experiment, within the stimulation experiment, and across both experiments. We found that cell type distances between organisms were reduced within multiplexed experiments compared to distances across experiments. The merged atlas was also used for determination of strong in situ

markers, mainly for gastrodermal cell types, since these highly related types share many marker genes.

RNA-seq analysis and clustering with MARIMBA annotation

To make our dataset more easily searchable with the MARIMBA v.1 (created on 05/30/2016) transcriptome annotation (*21*), we generated single-cell gene count matrices with respect to transcript sequences distributed via the MARIMBA website (http://dx.doi.org/10.22002/D1.1830). The gene count matrices were generated using the same kallisto-bustools workflow. We compared the application of the previous clustering/cell type assignment (including the overlap in differentially expressed genes delineating these clusters) to validate the quantification derived from the Trinity/Cuffcompare assembled transcriptome (fig. S13).

We also produced a notebook for visualization of gene expression in this dataset to facilitate its use in future studies. This establishes a code base for rapidly and transparently processing and comparing single-cell datasets with future transcriptome annotations.

Neural analysis

We clustered all cells within the broad class of Neural with Louvain clustering to obtain distinct subpopulations (labeled in Fig. 2A). Markers were determined with Scanpy's rank_genes_groups function (using the Wilcoxon test) for each subpopulation (table S5).

Marker genes from neural clusters, with predicted signal peptides (*72*), were screened for candidate neuropeptide cleavages sites (regular expression G[KR][KRED]). In cases where a sequence had more than one match to this motif, the six residues immediately N-terminal to the motif were inspected for similarity to each other—when similarity was present, the protein was considered a neuropeptide candidate. Predicted sequences can be found in table S3.

Pseudo-time analysis

We selected cells from cell types of interest (i-cells, nematoblasts, nematocytes, and neural cells) and used diffusion maps (*73*) to create a reduced dimension representation of cells, along with Scanpy's dpt function that uses geodesic distance along the graph of cells (in the determined "diffusion component" space) to estimate pseudotime. We then computed a PAGA-based embedding to visualize the cells in the context of the different trajectories with a ForceAtlas2 layout. To determine which genes constituted the important features in a given pseudo-time trajectory, we implemented a method based on the random forest method used in the dynverse R package for extracting "important" genes (*74*). A random forest regression model implemented with sklearn's random:forest_regressor was used to identify genes that were good predictors of the generated pseudo-time values (grouped into quantiles). This was run for each of the two inferred trajectories separately (stem cells to nematocytes, and stem cells to neurons). The training set consisted of 80% of the genes' expression data, within which 80% was used for optimizing the model, and 20% were used to evaluate the mean squared error (MSE). Both trajectory models had an R^2(coefficient of determination) of 0.85 or greater. The remaining 20% of data from the full dataset were used to calculate the gene-wise permutation importance scores, providing a ranking of each feature (gene) in terms of its contribution to the model's predictive capabilities using sklearn's permutation_importance. In the ranking, positive scores indicate importance, with the importance $i_j$for gene j being the difference between the original score s (MSE) calculated from the original model, and the average score across $K = 5$ random permutations of the feature columns in the validation dataset

$$i_j = s - \frac{1}{K}\sum_{k=1}^{K} s_{k,j}$$

Genes with nonzero (positive) scores in the permutation test were retained and ranked (table S5).

To assess the relationship of the gland cells to the neural populations, in the investigation of a neural-gland progenitor population, we applied the URD tree construction pipeline (*75*) to create a pseudotime–biased dendrogram of the neural and gland cell types beginning at the defined i-cell population. This included the i-cell, gland cell, and neuron clusters. We additionally plotted expression along the generated tree of *Myb* and *Myc3* orthologs corresponding to the neural-gland progenitor markers in (*8*) and other progenitor/developmental markers from (*8*) for the gland and neural cell types (fig. S16).

Perturbation response analysis

Extracting DE (perturbed) genes

We used a likelihood ratio test (LRT) with negative binomial–based models of each gene's expression in DESeq2 (*76*, *77*), with the reduced model not including the condition label (control or starved) and single cells treated as individual replicates. There are other approaches that can be used (*78*, *79*); we found that this method yielded many biologically relevant genes, which we could validate via the in situ experiments. All nonzero-expression genes were used for analysis. From the LRT, we obtained P values (corrected with the Benjamini-Hochberg multiple testing correction method across the number of genes tested) to determine whether the gene's expression was significantly affected by the condition and thus the perturbation. Genes with $\alpha < 0.05$ and $|\log_2 FC| > 1$ were selected as significant. We used the parameters sfType = "poscounts", minmu = 1×10^6, and minReplicatesForReplace = Inf in the DESeq2 model. Clusters with greater than 100 cells in each condition were subsampled evenly from fed and starved cells to reduce the effect of uneven cluster sizes on differential expression analysis (fig. S5). Clusters with less than 10 cells in any condition were not used for this analysis. To create the UpSet Plot (*80*) in Fig. 6E, P values from the LRT analysis were corrected for multiple testing (Bonferroni correction with n = number of clusters, since the test is assessing the intersection of genes across all cell types). All genes are included in table S5.

We additionally include low expression DE genes detected by pseudo-bulk analysis in table S5. We applied the same methodology as above, treating each animal as a replicate (five control and five starved replicates) and summing the counts per gene for each organism. For the DESeq2 model, we used the default parameters, as we were simulating bulk results.

Quantitative PCR validation for up-regulated, stimulation-responsive genes

To further validate the gene candidates from the stimulation multiplexed experiment, which displayed immediate early gene-like expression under KCl or DI stimulation, we chose broadly expressed candidates, i.e., DE genes in multiple cell types, for whole-animal, bulk quantitative PCR (qPCR). Here, we looked for up-regulation of these genes as compared to a "housekeeping" collagen marker, XLOC_008048. Two organisms were housed together in 150 ml of SW, with four animals total per each condition (SW, KCl, and DI)

treated as described previously in Materials and Methods. The two animals were homogenized together in 1 ml of TRIzol on ice with a syringe. A total of 0.2 ml of chloroform were added per 1 ml of TRIzol and shaken for 15 s. After spinning at 12,000*g* for 15 min at 4°C, the top phase is removed and, in a new tube, mixed with an equal volume of 60% EtOH. The RNEasy Mini Kit (QIAGEN, catalog no. 74004) was then used for bulk RNA extraction from the sample. Probes for the candidates and housekeeping gene were produced as previously described, and the LightCycler 480 SYBR Green I Master Mix (2X) was used for 20 µl of qPCR reactions with 5 µl of RNA template per reaction, using the standard cycling procedure with a 62°C annealing temperature. For each RNA sample ($n = 6$), three qPCR replicate reactions were done, resulting in a total of six replicates for each perturbation condition. Fold-change ($\log_2$FC) values presented in fig. S7E were calculated by first normalizing C_t values for each replicate to the average housekeeping gene C_t value in the same condition ($\Delta C_{tNorm} = C_{tRep} - C_{tHouse\_Avg}$). The $\log_2$FC was then calculated as $\log(2^{\Delta\Delta CtNorm})$ where $\Delta\Delta C_{tNorm} = \Delta C_{tNorm} - avg.(\Delta C_{tNorm:SW})$, is the difference of each normalized C_t from the average normalized C_t for the SW (control) condition. All primers for qPCR can be found in table S3.

De novo perturbed gene clusters

To cluster genes affected by perturbation (table S5) and to obtain information on their coexpression and possible functional similarity, we transposed the cell-by-gene expression matrix (obtaining a gene-by-cell matrix) for only the aggregated perturbed genes with padj > 0.05 (adjusted across genes and clusters) from the DeSeq2 analysis. We used Louvain clustering on the gene expression matrix, identifying both coexpressed genes and cell type–specific genes, similar to Monocle's procedure for detecting gene modules (*56*). We then used the aggregated information from the modules to determine putative functions/response types. The topGO weight algorithm (*81*) was used to determine GO terms that were significantly enriched in each gene module compared to GO terms in all other groups, with significance threshold $\alpha < 0.05$. The *P* values were also adjusted for multiple testing over the number of different gene modules using Bonferroni correction, and only significant GO terms were used to label the response types among the modules (fig. S20 and table S5). This analysis was replicated to determine gene modules distinguishing the GD cell types by clustering groups of coexpressed marker genes across the subtypes.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abh1683

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. B. Hwang, J. H. Lee, D. Bang, Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp. Mol. Med.* **50**, 1–14 (2018).
2. H. M. Vergara, P. Y. Bertucci, P. Hantz, M. A. Tosches, K. Achim, P. Vopalensky, D. Arendt, Whole-organism cellular gene-expression atlas reveals conserved cell types in the ventral nerve cord of *Platynereis dumerilii*. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 5878–5885 (2017).
3. A. Sebé-Pedrós, B. Saudemont, E. Chomsky, F. Plessier, M.-P. Mailhé, J. Renno, Y. Loe-Mie, A. Lifshitz, Z. Mukamel, S. Schmutz, S. Novault, P. R. H. Steinmetz, F. Spitz, A. Tanay, H. Marlow, Cnidarian cell type diversity and regulation revealed by whole-organism single-cell RNA-seq. *Cell* **173**, 1520–1534.e20 (2018).
4. J. Cao, J. S. Packer, V. Ramani, D. A. Cusanovich, C. Huynh, R. Daza, X. Qiu, C. Lee, S. N. Furlan, F. J. Steemers, A. Adey, R. H. Waterston, C. Trapnell, J. Shendure, Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017).
5. X. Han, R. Wang, Y. Zhou, L. Fei, H. Sun, S. Lai, A. Saadatpour, Z. Zhou, H. Chen, F. Ye, D. Huang, Y. Xu, W. Huang, M. Jiang, X. Jiang, J. Mao, Y. Chen, C. Lu, J. Xie, Q. Fang, Y. Wang, R. Yue, T. Li, H. Huang, S. H. Orkin, G.-C. Yuan, M. Chen, G. Guo, Mapping the mouse cell atlas by microwell-seq. *Cell* **172**, 1091–1107.e17 (2018).
6. D. E. Wagner, C. Weinreb, Z. M. Collins, J. A. Briggs, S. G. Megason, A. M. Klein, Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 (2018).
7. C. T. Fincher, O. Wurtzel, T. de Hoog, K. M. Kravarik, P. W. Reddien, Cell type transcriptome atlas for the planarian Schmidtea mediterranea. *Science* **360**, eaaq1736 (2018).
8. S. Siebert, J. A. Farrell, J. F. Cazet, Y. Abeykoon, A. S. Primack, C. E. Schnitzler, C. E. Juliano, Stem cell differentiation trajectories in *Hydra* resolved at single-cell resolution. *Science* **365**, eaav9314 (2019).
9. T. Gerber, P. Murawala, D. Knapp, W. Masselink, M. Schuez, S. Hermann, M. Gac-Santel, S. Nowoshilow, J. Kageyama, S. Khattak, J. D. Currie, J. G. Camp, E. M. Tanaka, B. Treutlein, Single-cell analysis uncovers convergence of cell identities during axolotl limb regeneration. *Science* **362**, eaaq0681 (2018).
10. A. Dixit, O. Parnas, B. Li, J. Chen, C. P. Fulco, L. Jerby-Arnon, N. D. Marjanovic, D. Dionne, T. Burks, R. Raychowdhury, B. Adamson, T. M. Norman, E. S. Lander, J. S. Weissman, N. Friedman, A. Regev, Perturb-seq: Dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853–1866.e17 (2016).
11. J. M. McFarland, B. R. Paolella, A. Warren, K. Geiger-Schuller, T. Shibue, M. Rothberg, O. Kuksenko, W. N. Colgan, A. Jones, E. Chambers, D. Dionne, S. Bender, B. M. Wolpin, M. Ghandi, I. Tirosh, O. Rozenblatt-Rosen, J. A. Roth, T. R. Golub, A. Regev, A. J. Aguirre, F. Vazquez, A. Tsherniak, Multiplexed single-cell transcriptional response profiling to define cancer vulnerabilities and therapeutic mechanism of action. *Nat. Commun.* **11**, 4296 (2020).
12. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
13. H. T. N. Tran, K. S. Ang, M. Chevrier, X. Zhang, N. Y. S. Lee, M. Goh, J. Chen, A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biol.* **21**, 12 (2020).
14. J. Gehring, J. Hwee Park, S. Chen, M. Thomson, L. Pachter, Highly multiplexed single-cell RNA-seq by DNA oligonucleotide tagging of cellular proteins. *Nat. Biotechnol.* **38**, 35–38 (2020).
15. C. S. McGinnis, D. M. Patterson, J. Winkler, D. N. Conrad, M. Y. Hein, V. Srivastava, J. L. Hu, L. M. Murrow, J. S. Weissman, Z. Werb, E. D. Chow, Z. J. Gartner, MULTI-seq: Sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat. Methods* **16**, 619–626 (2019).
16. C. Guo, W. Kong, K. Kamimoto, G. C. Rivera-Gonzalez, X. Yang, Y. Kirita, S. A. Morris, CellTag Indexing: Genetic barcode-based sample multiplexing for single-cell genomics. *Genome Biol.* **20**, 90 (2019).
17. S. R. Srivatsan, J. L. McFaline-Figueroa, V. Ramani, L. Saunders, J. Cao, J. Packer, H. A. Pliner, D. L. Jackson, R. M. Daza, L. Christiansen, F. Zhang, F. Steemers, J. Shendure, C. Trapnell, Massively multiplex chemical transcriptomics at single-cell resolution. *Science* **367**, 45–51 (2020).
18. M. Stoeckius, C. Hafemeister, W. Stephenson, B. Houck-Loomis, P. K. Chattopadhyay, H. Swerdlow, R. Satija, P. Smibert, Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
19. C. Sinigaglia, S. Peron, J. Eichelbrenner, S. Chevalier, J. Steger, C. Barreau, E. Houliston, L. Leclère, Pattern regulation in a regenerating jellyfish. *eLife* **9**, e54868 (2020).
20. T. C. G. Bosch, A. Klimovich, T. Domazet-Lošo, S. Gründer, T. W. Holstein, G. Jékely, D. J. Miller, A. P. Murillo-Rincon, F. Rentzsch, G. S. Richards, K. Schröder, U. Technau, R. Yuste, Back to the basics: Cnidarians start to fire. *Trends Neurosci.* **40**, 92–105 (2017).
21. L. Leclère, C. Horin, S. Chevalier, P. Lapébie, P. Dru, S. Peron, M. Jager, T. Condamine, K. Pottin, S. Romano, J. Steger, C. Sinigaglia, C. Barreau, G. Quiroga Artigas, A. Ruggiero, C. Fourrage, J. E. M. Kraus, J. Poulain, J.-M. Aury, P. Wincker, E. Quéinnec, U. Technau, M. Manuel, T. Momose, E. Houliston, R. R. Copley, The genome of the jellyfish *Clytia hemisphaerica* and the evolution of the cnidarian life-cycle. *Nat. Ecol. Evol.* **3**, 801–810 (2019).
22. L. Leclère, R. R. Copley, T. Momose, E. Houliston, Hydrozoan insights in animal development and evolution. *Curr. Opin. Genet. Dev.* **39**, 157–167 (2016).
23. P. R. H. Steinmetz, J. E. M. Kraus, C. Larroux, J. U. Hammel, A. Amon-Hassenzahl, E. Houliston, G. Wörheide, M. Nickel, B. M. Degnan, U. Technau, Independent evolution of striated muscles in cnidarians and bilaterians. *Nature* **487**, 231–234 (2012).
24. Z. Kamran, K. Zellner, H. Kyriazes, C. M. Kraus, J.-B. Reynier, J. E. Malamy, In vivo imaging of epithelial wound healing in the cnidarian *Clytia hemisphaerica* demonstrates early evolution of purse string and cell crawling closure mechanisms. *BMC Dev. Biol.* **17**, 17 (2017).
25. B. Galliot, V. Schmid, Cnidarians as a model system for understanding evolution and regeneration. *Int. J. Dev. Biol.* **46**, 39–48 (2002).

26. A. Amiel, P. Chang, T. Momose, E. Houliston, Clytia hemisphaerica: A cnidarian model for studying oogenesis, in *Oogenesis: The Universal Process* (John Wiley & Sons, 2010), pp. 81–102.
27. L. Leclère, E. Röttinger, Diversity of cnidarian muscles: Function, anatomy, development and regeneration. *Front. Cell Dev. Biol.* **4**, 157 (2016).
28. S. Fujita, E. Kuranaga, Y.-I. Nakajima, Cell proliferation controls body size growth, tentacle morphogenesis, and regeneration in hydrozoan jellyfish *Cladonema pacificum*. *PeerJ.* **7**, e7579 (2019).
29. A. Amiel, E. Houliston, Three distinct RNA localization mechanisms contribute to oocyte polarity establishment in the cnidarian Clytia hemisphaerica. *Dev. Biol.* **327**, 191–203 (2009).
30. M. Sheng, M. E. Greenberg, The regulation and function of c-fos and other immediate early genes in the nervous system. *Neuron* **4**, 477–485 (1990).
31. F. A. Wolf, F. K. Hamey, M. Plass, J. Solana, J. S. Dahlin, B. Göttgens, N. Rajewsky, L. Simon, F. J. Theis, PAGA: Graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 59 (2019).
32. L. McInnes, J. Healy, J. Melville, UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv [stat.ML]* (2018); http://arxiv.org/abs/1802.03426.
33. G. Hemmrich, K. Khalturin, A.-M. Boehm, M. Puchert, F. Anton-Erxleben, J. Wittlieb, U. C. Klostermeier, P. Rosenstiel, H.-H. Oberg, T. Domazet-Loso, T. Sugimoto, H. Niwa, T. C. G. Bosch, Molecular signatures of the three stem cell lineages in *Hydra* and the emergence of stem cell function at the base of multicellularity. *Mol. Biol. Evol.* **29**, 3267–3280 (2012).
34. L. Leclère, M. Jager, C. Barreau, P. Chang, H. Le Guyader, M. Manuel, E. Houliston, Maternally localized germ plasm mRNAs and germ cell/stem cell formation in the cnidarian Clytia. *Dev. Biol.* **364**, 236–248 (2012).
35. T. Condamine, M. Jager, L. Leclère, C. Blugeon, S. Lemoine, R. R. Copley, M. Manuel, Molecular characterisation of a cellular conveyor belt in *Clytia* medusae. *Dev. Biol.* **456**, 212–225 (2019).
36. E. Denker, M. Manuel, L. Leclère, H. Le Guyader, N. Rabet, Ordered progression of nematogenesis from stem cells through differentiation stages in the tentacle bulb of *Clytia hemisphaerica* (Hydrozoa, Cnidaria). *Dev. Biol.* **315**, 99–113 (2008).
37. K. Sunagar, Y. Y. Columbus-Shenkar, A. Fridrich, N. Gutkovich, R. Aharoni, Y. Moran, Cell type-specific expression profiling unravels the development and evolution of stinging cells in sea anemone. *BMC Biol.* **16**, 108 (2018).
38. N. Takeda, Y. Kon, G. Quiroga Artigas, P. Lapébie, C. Barreau, O. Koizumi, T. Kishimoto, K. Tachibana, E. Houliston, R. Deguchi, Identification of jellyfish neuropeptides that act directly as oocyte maturation-inducing hormones. *Development* **145**, dev156786 (2018).
39. C. Fourrage, K. Swann, J. R. Gonzalez Garcia, A. K. Campbell, E. Houliston, An endogenous green fluorescent protein–photoprotein pair in *Clytia hemisphaerica* eggs shows co-targeting to mitochondria and efficient bioluminescence energy transfer. *Open Biol.* **4**, 130206 (2014).
40. P. R. H. Steinmetz, A non-bilaterian perspective on the development and evolution of animal digestive systems. *Cell Tissue Res.* **377**, 321–339 (2019).
41. E. Denker, E. Bapteste, H. Le Guyader, M. Manuel, N. Rabet, Horizontal gene transfer and the evolution of cnidarian stinging cells. *Curr. Biol.* **18**, R858–R859 (2008).
42. L. A. Bezares-Calderón, J. Berger, G. Jékely, Diversity of cilia-based mechanosensory systems and their functions in marine animal behaviour. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **375**, 20190376 (2020).
43. D. R. McPherson, Sensory hair cells: An introduction to structure and physiology. *Integr. Comp. Biol.* **58**, 282–300 (2018).
44. J. S. Hwang, Y. Takaku, J. Chapman, K. Ikeo, C. N. David, T. Gojobori, Cilium evolution: Identification of a novel protein, nematocilin, in the mechanosensory cilium of *Hydra* nematocytes. *Mol. Biol. Evol.* **25**, 2009–2017 (2008).
45. P. Lapébie, A. Ruggiero, C. Barreau, S. Chevalier, P. Chang, P. Dru, E. Houliston, T. Momose, Differential responses to Wnt and PCP disruption predict expression and developmental function of conserved and novel genes in a cnidarian. *PLOS Genet.* **10**, e1004590 (2014).
46. M. E. Massari, C. Murre, Helix-loop-helix proteins: Regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* **20**, 429–440 (2000).
47. P. Kurki, M. Vanderlaan, F. Dolbeare, J. Gray, E. M. Tan, Expression of proliferating cell nuclear antigen (PCNA)/cyclin during the cell cycle. *Exp. Cell Res.* **166**, 209–219 (1986).
48. J. R. Sanes, S. L. Zipursky, Synaptic specificity, recognition molecules, and assembly of neural circuits. *Cell* **181**, 536–556 (2020).
49. G. Jékely, The chemical brain hypothesis for the origin of nervous systems. *Philos. Trans. R. Soc. B* **376**, 20190761 (2021).
50. C. J. P. Grimmelikhuijzen, F. Hauser, Mini-review: The evolution of neuropeptide signaling. *Regul. Pept.* **177** Suppl, S6–S9 (2012).
51. G. Jékely, Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8702–8707 (2013).
52. S. K. D. Nielsen, T. L. Koch, F. Hauser, A. Garm, C. J. P. Grimmelikhuijzen, De novo transcriptome assembly of the cubomedusa Tripedalia cystophora, including the analysis of a set of genes involved in peptidergic neurotransmission. *BMC Genomics* **20**, 175 (2019).
53. E. Serrano-Saiz, R. J. Poole, T. Felton, F. Zhang, E. D. De La Cruz, O. Hobert, Modular control of glutamatergic neuronal identity in C. elegans by distinct homeodomain proteins. *Cell* **155**, 659–673 (2013).
54. F. Li, J. Eriksen, J. Finer-Moore, R. Chang, P. Nguyen, A. Bowen, A. Myasnikov, Z. Yu, D. Bulkley, Y. Cheng, R. H. Edwards, R. M. Stroud, Ion transport and regulation in a synaptic vesicle glutamate transporter. *Science* **368**, 893–897 (2020).
55. C. Trapnell, B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. van Baren, S. L. Salzberg, B. J. Wold, L. Pachter, Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
56. C. Trapnell, D. Cacchiarelli, J. Grimsby, P. Pokharel, S. Li, M. Morse, N. J. Lennon, K. J. Livak, T. S. Mikkelsen, J. L. Rinn, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
57. M. Lechable, A. Jan, A. Duchene, J. Uveira, B. Weissbourd, L. Gissat, S. Collet, L. Gilletta, S. Chevalier, L. Leclère, S. Peron, C. Barreau, R. Lasbleiz, E. Houliston, T. Momose, An improved whole life cycle culture protocol for the hydrozoan genetic model Clytia hemisphaerica. *Biol. Open* **9**, bio051268 (2020).
58. G. Quiroga Artigas, P. Lapébie, L. Leclère, N. Takeda, R. Deguchi, G. Jékely, T. Momose, E. Houliston, A gonad-expressed opsin mediates light-induced spawning in the jellyfish Clytia. *eLife* **7**, e29555 (2018).
59. L. H. Hyman, Observations and experiments on the physiology of medusae. *Biol. Bull.* **79**, 282–296 (1940).
60. C. Sinigaglia, D. Thiel, A. Hejnol, E. Houliston, L. Leclère, A safer, urea-based in situ hybridization method improves detection of gene expression in diverse animal species. *Dev. Biol.* **434**, 15–23 (2018).
61. S. Chevalier, A. Martin, L. Leclère, A. Amiel, E. Houliston, Polarised expression of FoxB and FoxQ2 genes during development of the hydrozoan Clytia hemisphaerica. *Dev. Genes Evol.* **216**, 709–720 (2006).
62. H. Mi, A. Muruganujan, X. Huang, D. Ebert, C. Mills, X. Guo, P. D. Thomas, Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat. Protoc.* **14**, 703–721 (2019).
63. A. M. Altenhoff, J. Levy, M. Zarowiecki, B. Tomiczek, A. Warwick Vesztrocy, D. A. Dalquen, S. Müller, M. J. Telford, N. M. Glover, D. Dylus, C. Dessimoz, OMA standalone: Orthology inference among public and custom genomes and transcriptomes. *Genome Res.* **29**, 1152–1163 (2019).
64. M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, A. Regev, Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
65. B. J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. LeDuc, N. Friedman, A. Regev, De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).
66. P. Melsted, A. Sina Booeshaghi, L. Liu, F. Gao, L. Lu, K. H. Min, E. Beltrame, K. E. Hjorleifsson, J. Gehring, L. Pachter, Modular, efficient and constant-memory single-cell RNA-seq preprocessing. *Nat. Biotechnol.* **39**, 813–818 (2021).
67. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
68. P. De Meo, E. Ferrara, G. Fiumara, A. Provetti, Generalized Louvain method for community detection in large networks, in *Proceedings of the 2011 11th International Conference on Intelligent Systems Design and Applications* (2011), 88–93.
69. V. Ntranos, G. M. Kamath, J. M. Zhang, L. Pachter, D. N. Tse, Fast and accurate single-cell RNA-seq analysis by clustering of transcript-compatibility counts. *Genome Biol.* **17**, 112 (2016).
70. C. C. Aggarwal, A. Hinneburg, D. A. Keim, On the surprising behavior of distance metrics in high dimensional space. *Database Theory — ICDT 2001*, 420–434 (1973).
71. S. Chen, P. Rivaud, J. H. Park, T. Tsou, E. Charles, J. R. Haliburton, F. Pichiorri, M. Thomson, Dissecting heterogeneous cell populations across drug and disease conditions with PopAlign. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 28784–28794 (2020).
72. J. J. A. Armenteros, K. D. Tsirigos, C. K. Sønderby, T. N. Petersen, O. Winther, S. Brunak, G. von Heijne, H. Nielsen, SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).
73. L. Haghverdi, M. Büttner, F. A. Wolf, F. Buettner, F. J. Theis, Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* **13**, 845–848 (2016).
74. W. Saelens, R. Cannoodt, H. Todorov, Y. Saeys, A comparison of single-cell trajectory inference methods. *Nat. Biotechnol.* **37**, 547–554 (2019).
75. J. A. Farrell, Y. Wang, S. J. Riesenfeld, K. Shekhar, A. Regev, A. F. Schier, Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360**, eaar3131 (2018).

76. M. Love, S. Anders, W. Huber, Differential analysis of count data—The DESeq2 package. *Genome Biol.* **15**, 550 (2014).
77. T. Wang, B. Li, C. E. Nelson, S. Nabavi, Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data. *BMC Bioinformatics* **20**, 40 (2019).
78. H. L. Crowell, C. Soneson, P.-L. Germain, D. Calini, L. Collin, C. Raposo, D. Malhotra, M. D. Robinson, muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nat. Commun.* **11**, 6077 (2020).
79. M. D. Luecken, F. J. Theis, Current best practices in single-cell RNA-seq analysis: A tutorial. *Mol. Syst. Biol.* **15**, e8746 (2019).
80. A. Lex, N. Gehlenborg, H. Strobelt, R. Vuillemot, H. Pfister, UpSet: Visualization of intersecting sets. *IEEE Trans. Vis. Comput. Graph.* **20**, 1983–1992 (2014).
81. A. Alexa, J. Rahnenführer, Gene set enrichment analysis with topGO. *Bioconduct. Improv.* **27**, 1–26 (2009).
82. G. S. Richards, F. Rentzsch, Regulation of Nematostella neural progenitors by SoxB, Notch and bHLH genes. *Development* **142**, 3332–3342 (2015).

Acknowledgments: We thank X. Da and X. Wang for technical assistance, T. Momose for assistance with the single-cell experimentation, the Caltech Single-Cell Profiling and Engineering Center for the use of their single-cell and sequencing tools, and the Caltech Bioinformatics Resource Center for transcriptome assembly and annotation analysis. We additionally thank the Caltech Center for Evolutionary Science for the bioinformatics resources to create a local UCSC Genome Browser. We thank A. S. Booeshaghi for help with kallisto, bustools, and the kITE demultiplexing of the ClickTag reads and for rescuing the stimulation experiment sequencing data. We thank J. Malamy for helping to establish *Clytia* work at Caltech. We thank S. Peron for initial characterization of some of the cell type marker genes, P. Lapébie for identification of novel neuropeptide sequences, M. Jager for valuable advice on the in situ protocol, and J. R. Mateu for providing pp11 probe. **Funding:** J.G., M.H., and L.P. were supported in part by a seed grant from the Chen Institute at the California Institute of Technology. T.C., J.G., and L.P. were supported in part by NIH U19MH114830 and NIH RF1AG062324A. We thank the Marine Resources Centre (CRBM and PIV imaging platform) of Institut de la Mer de Villefranche (IMEV), supported by EMBRC-France. The French state funds of EMBRC-France are managed by the ANR within the investments of the Future program. L.L. was supported by the Agence Nationale de la Recherche (ANR-19-CE13-0003). A.F., R.R.C., and E.H. were supported by the H2020/Marie Skłodowska-Curie ITN "EvoCell" Grant agreement no. 766053. B.W. was supported in part by a Howard Hughes Medical Institute Fellowship of the Life Sciences Research Foundation and by NIH K99NS119749. This work was in part supported by the Whitman Center of the Marine Biological Laboratory in Woods Hole, MA and a visiting grant from EMBRC-France. D.J.A. is an Investigator of the Howard Hughes Medical Institute.
Author contributions: Conceived of the experiments: T.C., B.W., J.G., R.R.C., E.H., D.J.A., and L.P. Developed cell dissociation, fixation, and labeling procedures compatible with the 10X Genomics platform: J.G. and M.H. Performed the single-cell experiments: T.C., B.W., and J.G. Performed the in situ hybridization and other microscopy experiments: B.W., A.F., L.L., and S.C. Performed whole-organism qPCR: T.C. Performed bioinformatics analysis including assembly and annotation of the transcriptome: F.G. and R.R.C. Wrote scripts for processing the data and code for the analysis: T.C. and J.G. Developed the Google Colab notebooks: T.C. Analyzed and interpreted the data: T.C., B.W., J.G., A.F., L.L., R.R.C., E.H., D.J.A., and L.P. Writing and editing the manuscript: T.C., B.W., J.G., A.F., L.L., R.R.C., E.H., D.J.A., and L.P.
Competing interests: The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. All raw sequencing and processed data files used for analysis are available from CaltechData, (https://data.caltech.edu/search?page=1&size=25&ln=en&q=clytia), with links additionally provided via the notebooks in the code repository. The sequencing read alignments are available at http://evolution.caltech.edu/genomebrowser/cgi-bin/hgTracks?db=hub_135_clyHem1&lastVirtModeType=default&lastVirtModeExtraState=&virtModeType=default&virtMode=0&nonVirtPosition=&position=scaffold_1%3A1%2D30003&hgsid=3413_LS7OcP5N7VA2rApGOfk8iaX2kVFR, and an interactive browser for gene expression visualization (http://131.215.78.40/cb), is publicly hosted on a UCSC Genome Browser by the Caltech Bioinformatics Resource Center. The softwares used are as follows: Cell Ranger 3.0.1, Trinity-v2.8.4, Cufflinks v2.2.1, kallisto v0.46.2, bustools v0.40.0, anndata 0.7.5, louvain 0.7.0, rpy2 3.4.2, scanpy 1.6.0, biopython 1.78, pysam 0.16.0.1, fuzzywuzzy 0.18.0, numpy 0.19.5, pandas 1.1.5, matplotlib 3.2.2, sklearn 0.0, scipy 1.4.1, seaborn 0.11.1, requests 2.23.0, tqdm 4.41.1, multiprocess 0.70.11.1, DESeq2 1.3.0, topGO 2.42.0, and UpSet 1.4.0. Code availability: All the codes used to perform the analyses and generate the results and figures are available in Google Colab notebooks archived with Zenodo at https://zenodo.org/record/5519756#.YUonytNKgUE and directly available at https://github.com/pachterlab/CWGFLHGCCHAP_2021. The notebooks, which include the complete preprocessing of the raw data and a walkthrough of the code, provide a transparent implementation of the methods and can be run for free in the Google cloud.

Submitted 20 February 2021
Accepted 6 October 2021
Published 26 November 2021
10.1126/sciadv.abh1683

CHAPTER 3 -
Cell types of the planula larva of *Clytia*

Overview of the chapter

The planula is the characteristic larval form of many species belonging to the phylum Cnidaria. Even though for many cnidarian species the general anatomy/morphology of the planula along with some cell types have been described, there is still a gap to be filled regarding cell type transcriptional programs. Here I introduce the cell types atlas of the planula larva of *Clytia* that I obtained by dissociating 2-day-old planulae and processing the single cells via single-cell RNA-seq technology to generate the data. At this stage of development, cell diversity should include both differentiated and differentiating cells.

Dissociated planula cells showed a particular fragility. For this reason I developed a protocol that includes immediate fixation of single cells and a sorting step prior to encapsulation. The final dataset is the result of the integration of 6 libraries and consists of 5347 cells which could be computationally combined in 22 clusters corresponding to 20 transcriptional signatures that could be assigned to cell types and two mixed profile clusters of putative technical artefacts. Reclustering analysis of the 20 cell types yielded a definitive atlas including 4370 cells grouped in 19 cell clusters.

Following in situ hybridization analysis of known and novel genes at three planula developmental stages I could assign cell identities and combine the 19 clusters in 8 broad cell classes. These correspond to the two cnidarian epithelial tissue layers, the epidermis and the gastrodermis, the hydrozoan stem cells (i-cells), the nematocytes (stinging cells), neural cells, aboral neurosecretory cells and distinct population of secretory cells, mucus cells and putative excretory cells (PEC).

This *Clytia* planula cell type atlas represents the first cell atlas of an hydrozoan larva and provides characterisation of previously undescribed cell populations as well as further information on already known cell types. Additional cell diversity might be overlooked in our data considering the overall low number of cells, in particular concerning the neural populations. Therefore, in the future, it would be worth increasing the number of cells by integrating this dataset with newly produced single cell libraries of whole larvae or specifically targeted cells.

3.1 - INTRODUCTION

The ultimate aim of my PhD project is to compare cell type transcriptional profiles across life stages of *Clytia hemisphaerica* in order to understand better which are the cell types that are shared throughout the life cycle and how the different life forms diverged during evolution. Concretely, I focused on the adult medusa and the larval forms of the life cycle of *Clytia* to establish the comparison at cellular level. Previous analysis of the genome and the stage specific transcriptome of *Clytia* revealed a substantially low proportion of planula specific transcription factors compared to the polyp and medusa stages, suggesting a lower complexity in terms of cell types (Leclère et al., 2019). A number of different evolutionary scenarios could have led to this situation, as addressed in more detail in Chapter 4 and in the General Discussion. To distinguish between them, transcriptional profiles at the cellular level are required.

In Chapter 2 I introduced the medusa cell type atlas which represents a first step addressing these aims. In this Chapter, I present the work I carried out to generate a cell type atlas of the planula of *Clytia*. The planula exhibits a basic body organisation along an oral-aboral axis and two main tissue layers, the outer epidermis and the inner gastrodermis. The planula forms in about 24 hours after fertilisation. And after three days in laboratory conditions it undergoes a drastic metamorphosis to generate the primary polyp. The development and the organisation of the planula is reviewed in detail in Chapter 1 along with an overview of the knowledge concerning the cell types prior to this study.

The planula cell atlas provides the first characterization of transcriptional profiles at cellular level of a hydrozoan planula larva. To generate the data I used dissociated cells from whole 2-days old *Clytia* planulae (51-52 hpf at 17°C, see Methods) by initially employing a similar dissociation protocol used for the medusa (see Chapter 2). The goal was to generate data starting from living cells as I did for the medusa. Although the dissociations appeared very promising from observation of dissociated cells through the microscope, the encapsulation experiments were unsuccessful (see Troubleshooting section). For this reason I introduced a fixation step in the protocol, immediately after dissociation and stored the cells at -80°C until the day of the encapsulation. Finally, I generated data in collaboration with the Sebé-Pedros lab and integrated data from six libraries obtained with two fixation methods (see Methods). The resulting atlas consists of 4370 cells that could be grouped in 19 transcriptional profiles and eight broad cell classes.

The cell atlas of the planula of *Clytia* provides molecular characterization of known cell populations and of previously undescribed larval cell types.

3.2 - RESULTS

3.2.1 - A Clytia Planula Cell Atlas

As starting material for the generation of the planula cell type atlas I chose larvae at two days of development (50-52 hpf at 17°C, see Methods). High levels of transcriptional activity are expected at this stage therefore we predicted to detect differentiating as well as differentiated cells.

I initially carried out cell dissociation in Calcium/Magnesium free seawater as described for the medusa in Chapter 2, with the aim of generating data from freshly dissociated cells. Encapsulation of freshly dissociated cells from whole animals was successful for the *Clytia* medusa (see Chapter 2) and other species (Sebé-Pedrós, Chomsky, et al., 2018; Sebé-Pedrós, Saudemont, et al., 2018).
The dissociation of two-day-old planula initially appeared very promising but the encapsulation experiments were unsuccessful. We attribute the poor results to particular fragility of the cells and/or excess debris (see Troubleshooting section). For this reason I adapted the protocol by adding a fixation step immediately after dissociation. Batches of cells were fixed using either 80% Methanol or ACME solution (Chari et al., 2021; García-Castro et al., 2021). The fixed cells could be conserved at -80°C until the day of the encapsulation (see Methods). A FAC sorting step prior to encapsulation was performed at CRG in Barcelona in collaboration with Marta Iglesias and Arnau Sebé-Pedros, and the data were generated from the subsequent encapsulation experiment (see Methods).

The initial planula cell type atlas from this experiment includes 5347 cells resulting from the integration of six single cell libraries (Fig. 3.1a and b). Using the STAR 2.7.0 (Dobin et al., 2013) with exploitation of the STARSolo algorithm and Scanpy (Wolf et al., 2018) pipelines I obtained 22 cell clusters (see Methods). Cells belonging to each of the batches contribute to each of the clusters (Fig. 3.1c).

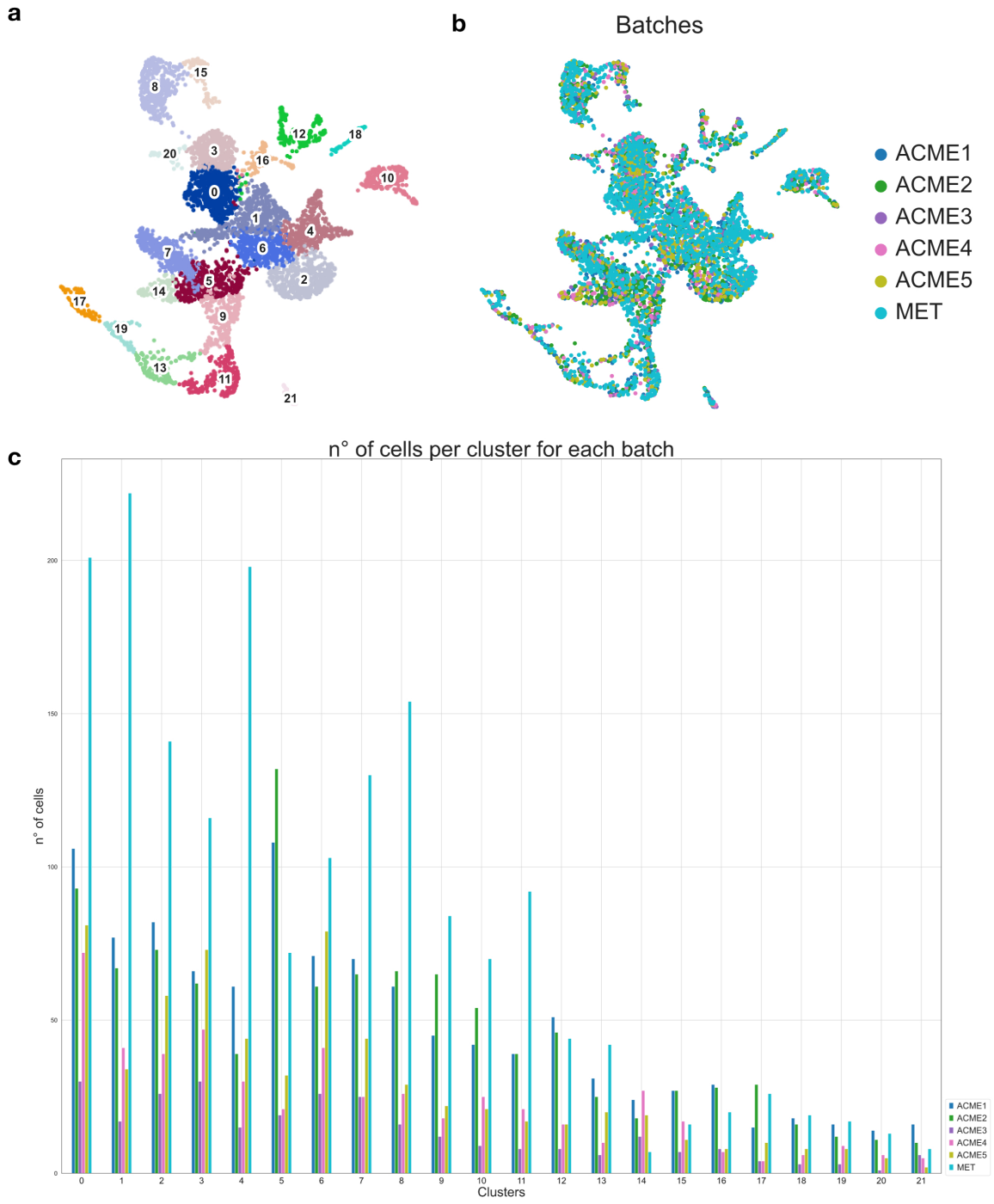


Figure 3.1 - Generation of the cell atlas

(a) 2D-UMAP plot embedding showing the Leiden algorithm annotation of the clusters. (b) UMAP plot embedding of the cell atlas showing cells belonging to the six libraries. Libraries/batches are labelled according to the technique used to prepare the samples: 80% Methanol fixed cells (MET) and ACME fixed cells (ACME). (c) Bar plot showing the contribution of the cells belonging to the six libraries/batches to the Leiden clusters. Number of cells is plotted on the y axis and clusters are plotted on the x axis. Bars are coloured according to the legend in (b).

To address the identity of each cell cluster I first ranked genes expressed in each of the 22 groups using a Wilcoxon non parametric test (Mann & Whitney, 1947; see Methods) and obtained a marker gene list of genes per cluster. I then assigned tentative identities to these clusters on the basis of known genes from the list, including markers from the medusa atlas. Additionally, I selected known and novel candidate genes and carried out *in situ* hybridization to determine the distribution of each cell type at successive stages of planula formation (24 hpf; 48 hpf; 72 hpf). Criteria for selecting candidates included expression levels high enough to be mapped onto the atlas and specificity to the clusters. A heatmap of candidate genes and additional diagnostic genes across cell types is shown in Figure 3.2b. The table including candidate genes information is provided as annex (Annex Table B).

I assigned identities to 20 cell clusters (Fig. 3.2a). I was unable to assign a clear identity to two of the clusters, 0 and 3, as they were characterised by a mixed transcriptome profile. This corresponds to a combination of all the other profiles, notably genes predominantly expressed at high levels in the epidermis cell types and in secretory cells such as neural cells, mucous cells, PEC and aboral neurosecretory cells (Fig. 3.2b). Consequently it was not possible to define statistically significant cluster specific marker genes. We consider that these clusters probably arise as an artefact during sample preparation as the mixed profile resembles the data obtained by previous unsuccessful experiments (see Troubleshooting section).

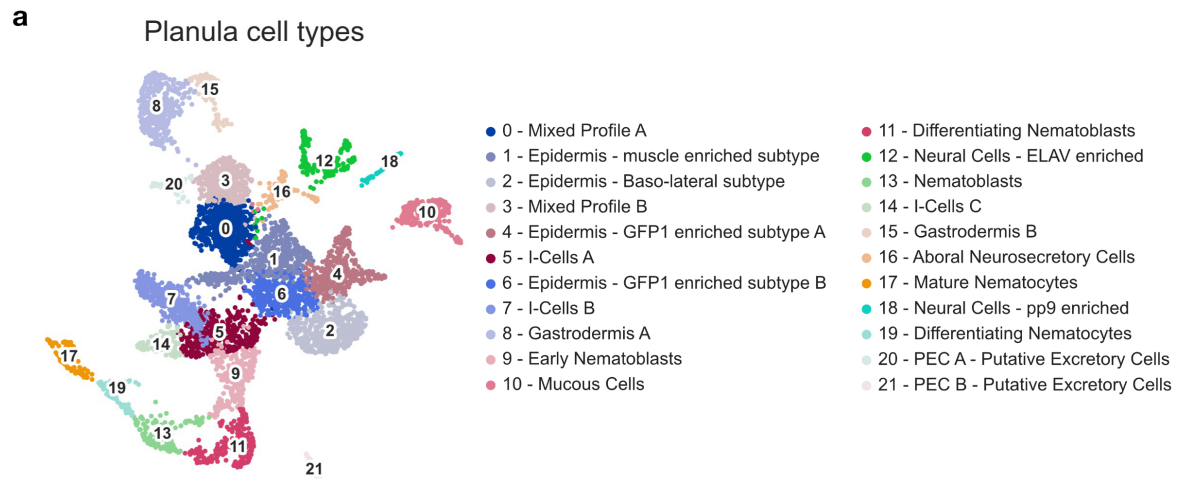


Figure 3.2 - Annotated cell atlas and candidate genes
(a) 2D-UMAP plot embedding with annotated Leiden clusters according to cell type identities following the analysis of in situ patterns of candidate genes. (b) Heatmap of candidate genes used for in situ hybridisation with additional diagnostic markers (following page). Genes are plotted on the y axis and cells are plotted on the x axis grouped by the annotation.

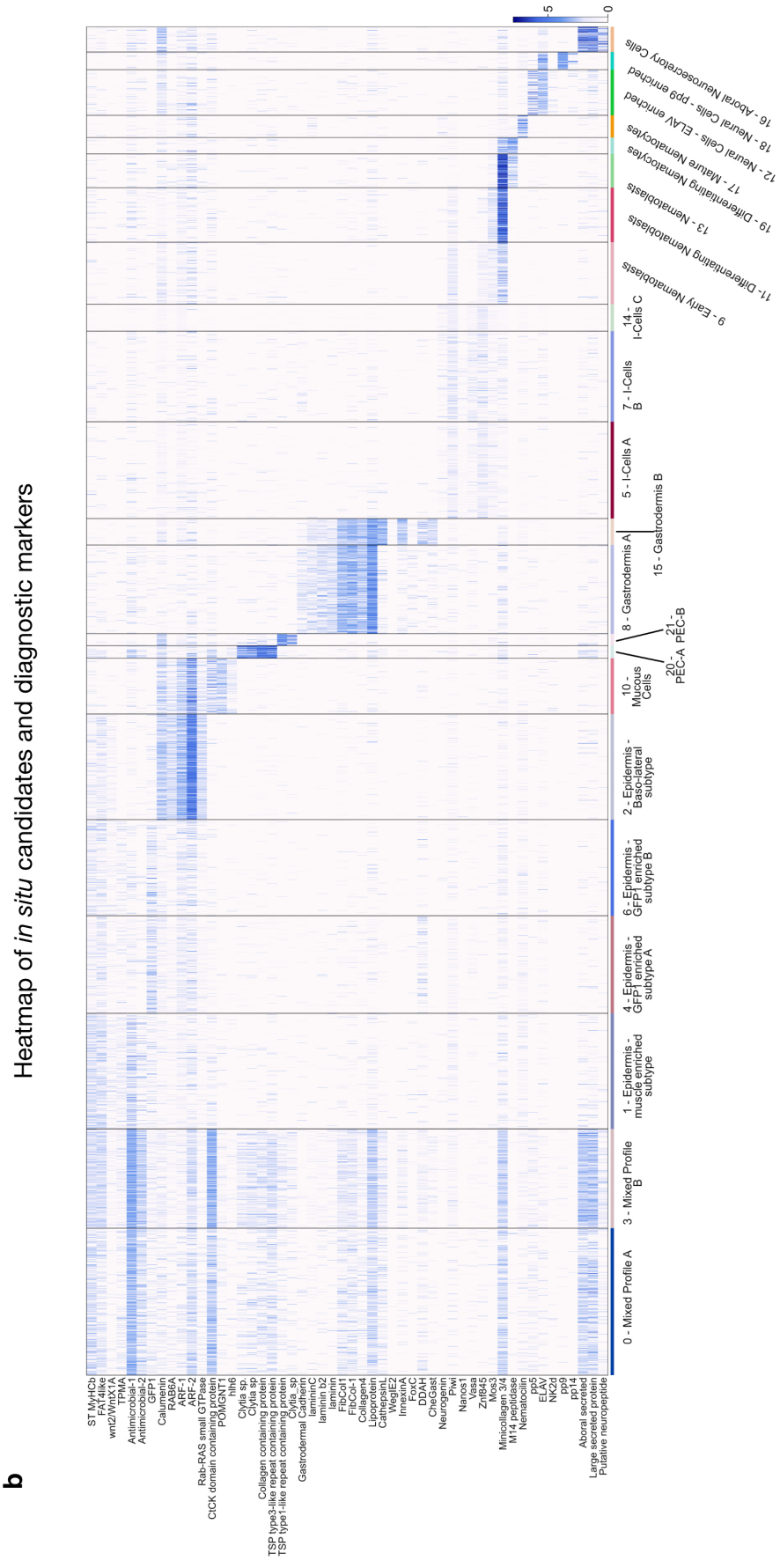


Figure 3.2b - Annotated cell atlas and candidate genes

I derived a second atlas by removing these mixed population ‘cells’ of cluster 0 and 3 (Fig. 3.3a). I transferred the annotation of the initial atlas to the second atlas (Fig. 3.3b) and recomputed the heatmap of candidate genes (Fig. 3.3c). The other cell clusters showed a rearrangement on the embedded UMAP projection graph but could be re-assigned to the same cell types according to the expression of marker genes. The final planula atlas introduced in Figure 3.3 is discussed in the section below and throughout the chapter.

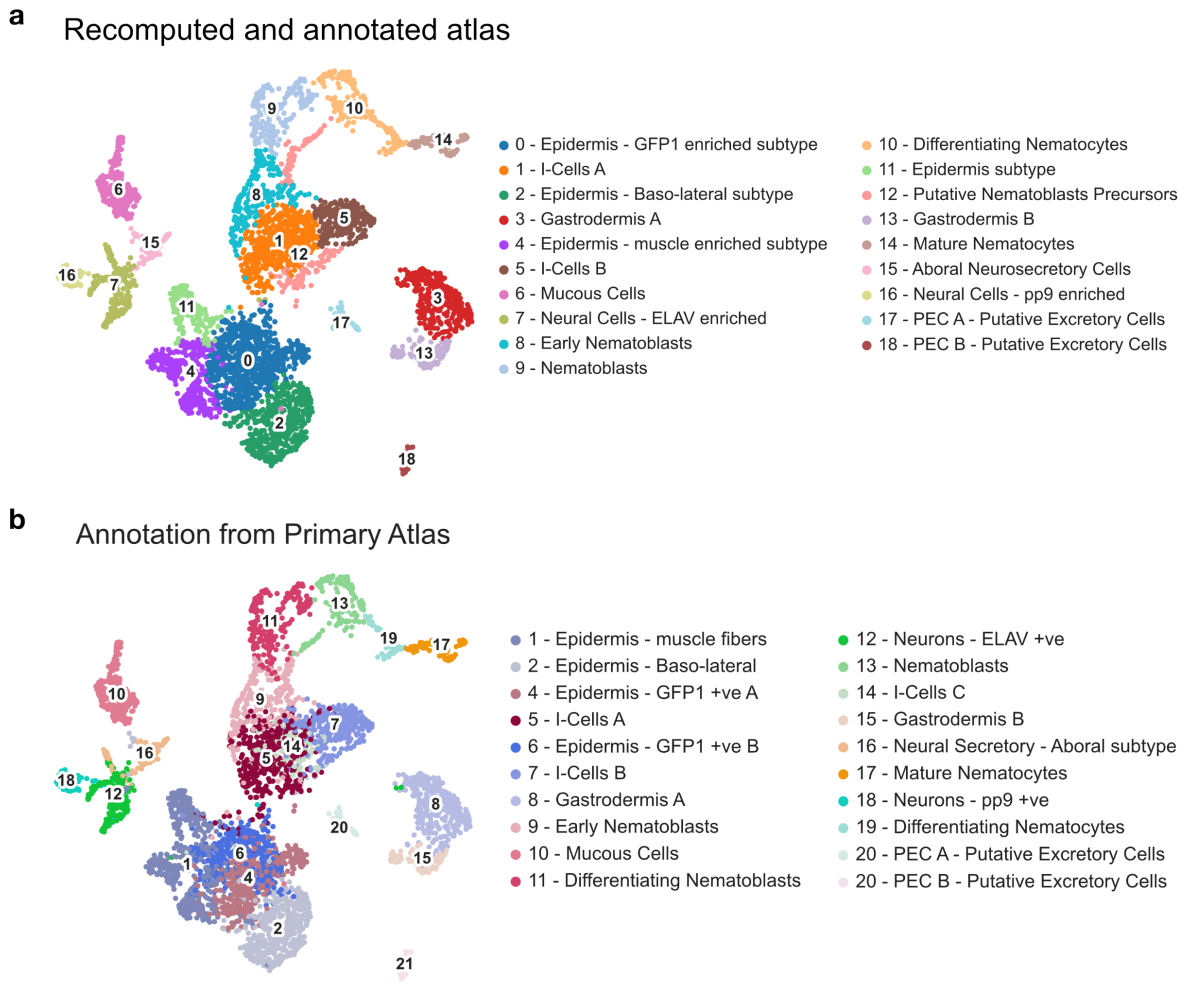


Figure 3.3 - Recomputed and annotated cell atlas and candidate genes
(a) Recomputed atlas following the exclusion of the ‘mixed profile” clusters from the primary atlas, with original annotation of the cell types. (b) Planula cell atlas with cell type annotation transferred from primary atlas. (c) Heatmap of candidate genes used for in situ hybridization with additional diagnostic markers plotted on the recomputed atlas with annotation of cell type (on the following page).

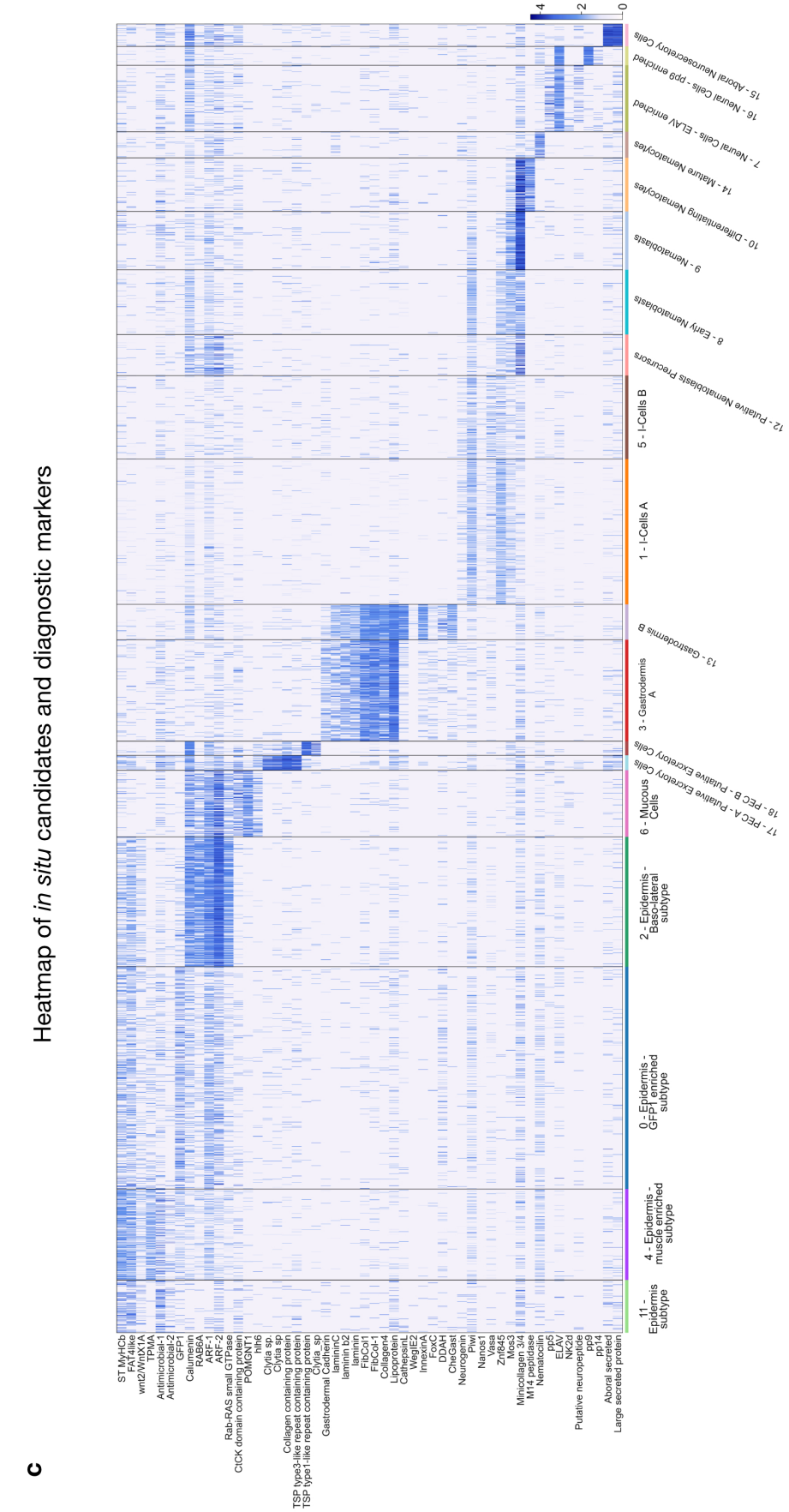


Figure 3.3c - Recomputed and annotated cell atlas and candidate genes

The final planula cell atlas consists of 4370 cells grouped into 19 distinct transcriptional profiles (Fig. 3.3a). Similarly to the medusa atlas, I grouped the 19 cell clusters into 8 broad cell classes: epidermis, gastrodermis, i-cells, nematocytes (differentiating and mature), neural cells, aboral neurosecretory cells, mucous cells and putative excretory cells (Fig 3.4b).

The diagram in Figure 3.4a represents the distribution of cell classes in the two-day-old planula. The heatmap of candidate genes and additional diagnostic genes grouped by cell classes is shown in Figure 3.4c

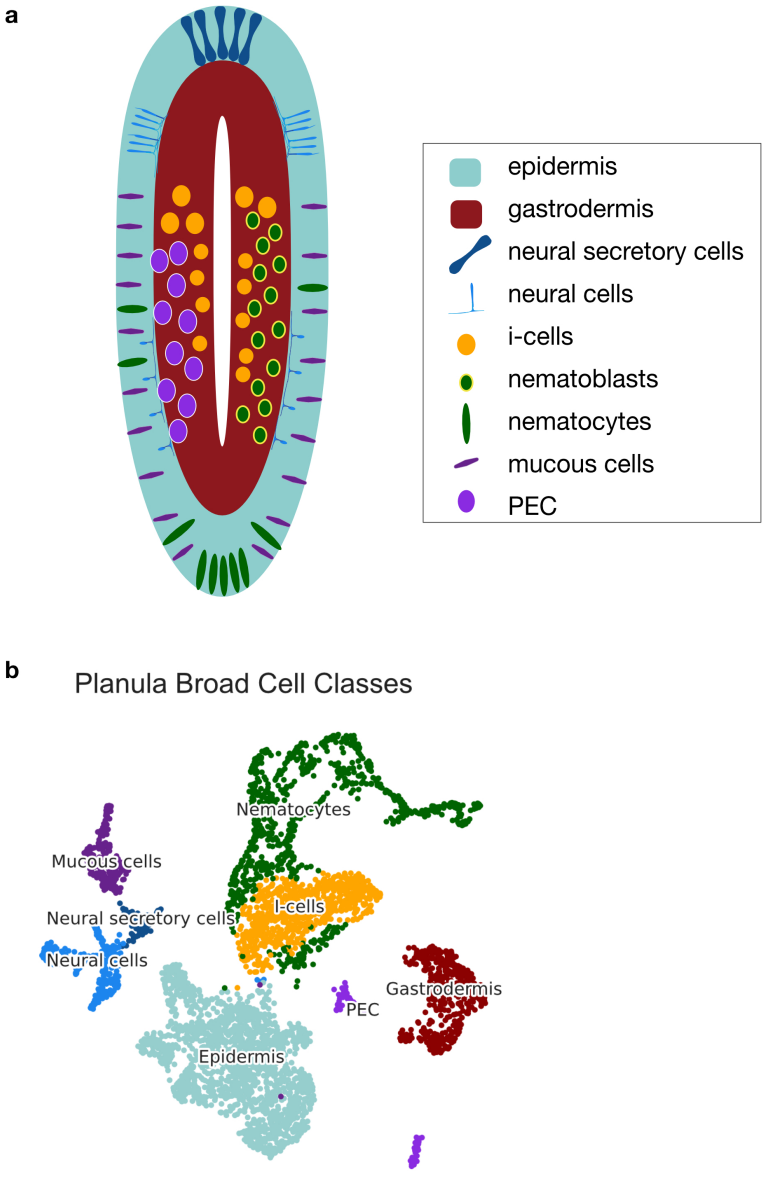


Figure 3.4 - Cell classes
(a) Diagram of 2-day-old planula larva with cell classes coloured according to the atlas. (b) 2-D UMAP embedding of the cell atlas grouped by cell classes. (c) Heatmap of in situ gene candidates and diagnostic markers grouped by cell classes.

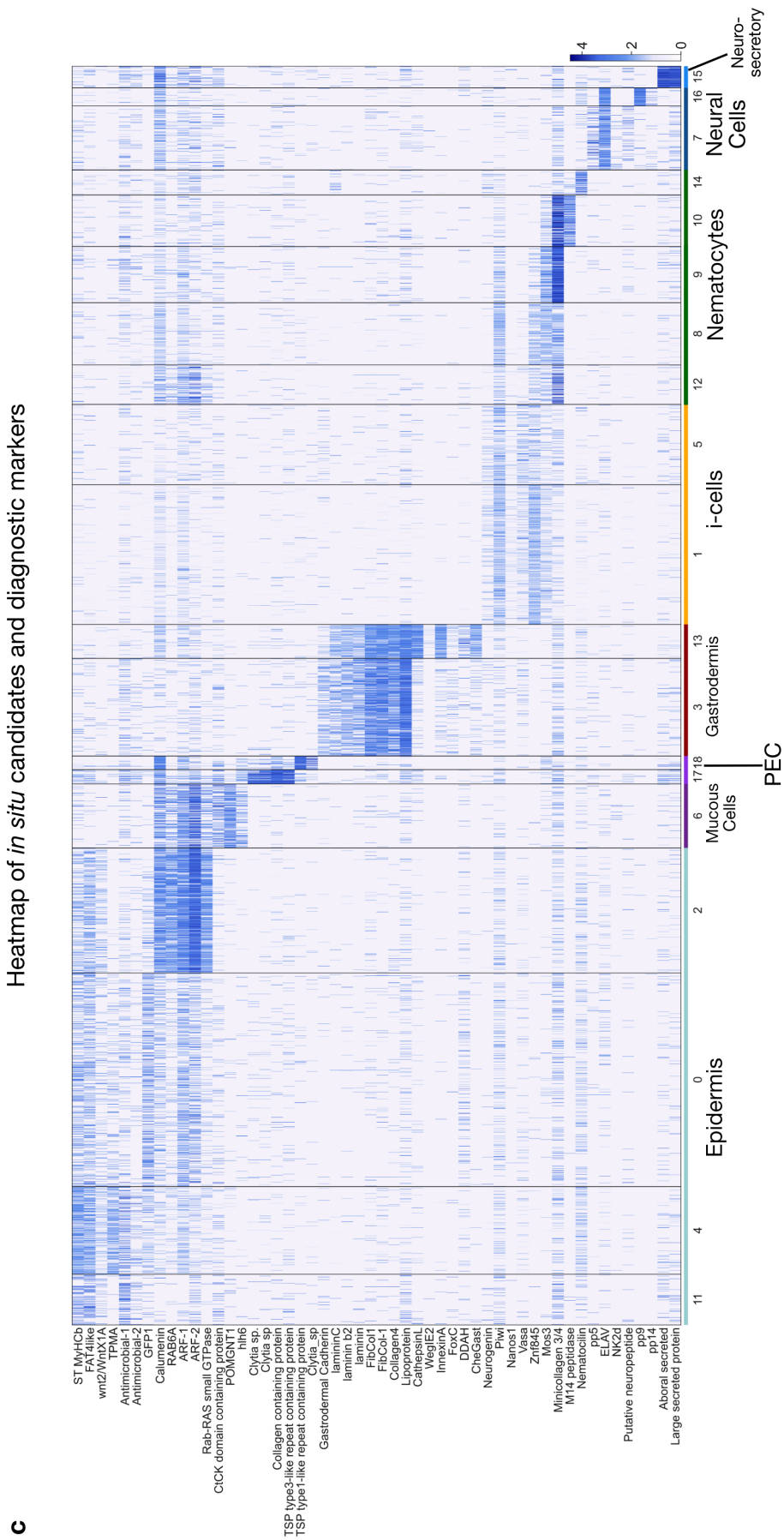


Figure 3.4c - Cell classes

3.2.1.1 - Ectodermal epithelial cells (Epidermis)

Ectodermal epithelial cells are the basic building blocks of the ectodermis and gastrodermis tissue layers in cnidarians (Leclère & Röttinger, 2017).
In the planula atlas I assigned four clusters of ectodermal epithelial cells (Fig. 3.5a). All the cells of the class are characterised by the expression of cadherin genes such as *Fat4like* (Magie & Martindale, 2008; Fig. 3.5b and c), a marker expressed throughout the life cycle (Fig. 3.8d). Within those, I could identify a sub-type of ectodermal cells enriched in muscle genes such as *Tropomyosin-A (TPMA)*, a marker expressed at medusa stage as well (Chari et al., 2021; Fig. 3.5e), which likely represents the epitheliomuscular cells of the planula (Leclère & Röttinger, 2017; Fig. 3.5b and c). The upregulated expression at 24hpf of general epitheliomuscular markers indicates the differentiation of the basic epitheliomuscular cells in early stages of planula development (Fig. 3.5c; panels “24 hpf”).

Among the four epidermis clusters I could identify two previously undescribed cell types. I assigned cluster 11 as an epidermal subtype expressing *Fat4like* and a planula specific putative antimicrobial protein (Fig. 3.5f). The cells are organised in a band in the aboral half of the larvae. The expression of this marker gene is only detectable at 48 and 72hpf (Fig. 3.5b and c). I assigned the second type as “basolateral epidermal cells”. These are characterised by a gene signature enriched in vesicle trafficking mediators including the Rab/RAS small GTPase marker. This population of cells is localised in the oral half of the larvae in the basal region of the ectoderm close in contact with the mesoglea (Fig. 3.5b and c). The lack of expression of the muscle markers TPMA suggests a non-muscular epidermal type. Cells of this cluster show a shared gene signature with the mucous cells suggesting a link between these cell types and/or shared secretory functions (see below). These last two markers (notably the putative antimicrobial marker and Rab/RAS) are specifically expressed at planula stage (Fig. 3.5f and g) suggesting that the epidermis subtype of the aboral ring and the basolateral epidermal cells are likely planula specific.

Marker genes expressed in putative planula specific epidermis subtypes are upregulated at later stages of larval development (Fig. 3.5b). We can speculate their involvement in specialised roles linked to behaviours that the planula exhibits towards the end of its lifetime. In the case of the unclassified epidermis subtype, one possibility is settlement, given the expression of a putative antimicrobial marker. Bacterial cues play a role in settlement and these cells might somehow contribute to this process. On the other hand, the basolateral epidermis subtype could be linked to signalling pathways which involve the mucous cells (see below), and potentially play a role prior to metamorphosis. At present, the function of those two subtypes remains to be determined. The marker genes I presented here can represent good candidates for future functional analyses.

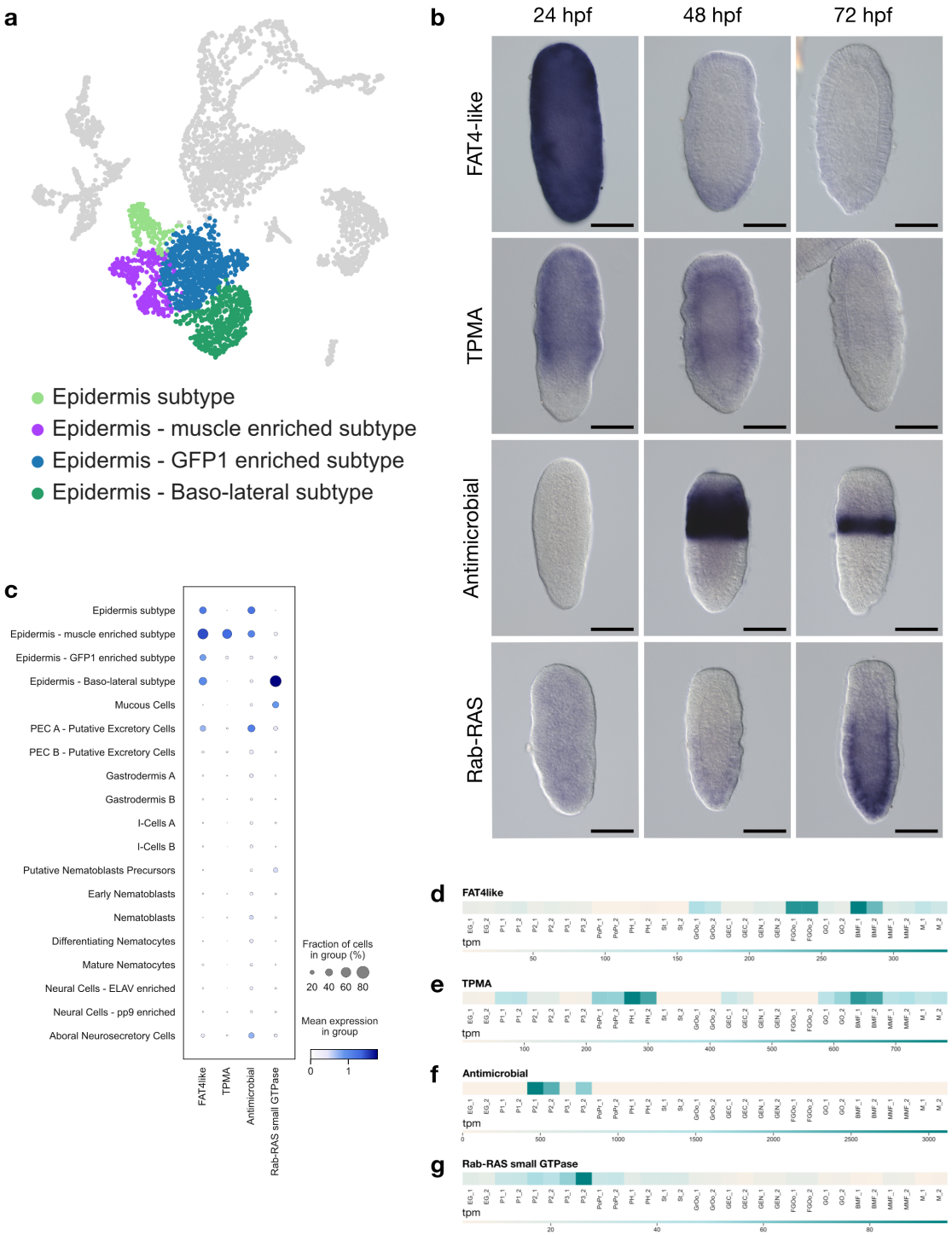


Figure 3.5 - Ectodermal epithelial cells

(a) 2D-UMAP plot coloured by epidermis cell types and legend with the cell type annotation. (b) *In situ* hybridization of epidermis marker genes at the three planula developmental stages. Scale Bars represent 100µm. (c) Dot Plot of expression of epidermis marker genes. (d, e, f, g) Expression of epidermis marker genes across developmental stages of *Clytia*. Transcripts per million (tpm) calculated from bulk RNA-seq data from Leclère et al., (2019). EG = Early gastrula, P1 = Planula 24 hpf = Planula 48 hpf, P3 = Planula 72 hpf, St = Stolon, PoPr = Primary polyp, PH = Polyp head, GO = Gonozooid, BMF = Baby medusa female, MMF = Mature medusa female, M = Mature medusa male

3.2.1.2 - Endodermal epithelial cells (Gastrodermis)

The inner epithelium of the planula or gastrodermis is the second classic cnidarian epithelium (Leclère & Röttinger, 2017). In the adult stages it has digestive function and usually gastrodermis cells surround digestive gland cells (see Chapter 2). In the gastrodermis of the planula there are numerous nuclei with irregular shapes which suggest that a process similar to phagocytosis is occurring (Fig. 1.10 - Chapter 1). Given that the planula does not feed, one possibility is that epithelial cells or yolk are digested by other cells to recycle cellular components used in other functions. How the cells of gastrodermis of the planula are involved in digestion or phagocytosis remains to be determined. One process that is active in the gastrodermis is the production of early stages of nematocytes. Additionally, i-cells also originate in the gastrodermis, potentially explaining why some of the planula i-cells share expressed genes with the gastrodermal epithelial cells.

In the planula atlas I could distinguish two clusters of gastrodermal epithelial cells (Fig. 3.6a). Analysis of the gene signatures revealed the expression of multiple proteins which likely represent components of the ECM/mesoglea such as laminins and cadherins (Magie & Martindale, 2008). We assigned the two cell types on the basis of previously described markers relating to intracellular digestion such as CathepsinL from Chari et al., (2021; Fig. 3.69b and c) as well as the taxon specific gene WegIE2 and transcription factor FoxC from Lapébie et al., (2014). The expression of CathepsinL and Innexin-A in gastrodermis-B compared to gastrodermis-A, suggests that type B might represent a more specialised subtype. However, the analysis of the *in situ* expression patterns did not highlight particularly striking subtypes. It is possible that the stage of development of the planula used for the generation of the data (notably 50-52 hpf) is not suitable to determine clear gastrodermis subtypes. The limited number of cells also likely hinders discrimination between types. The generation of additional data from the one used and from earlier stages of development would help to clarify the situation.

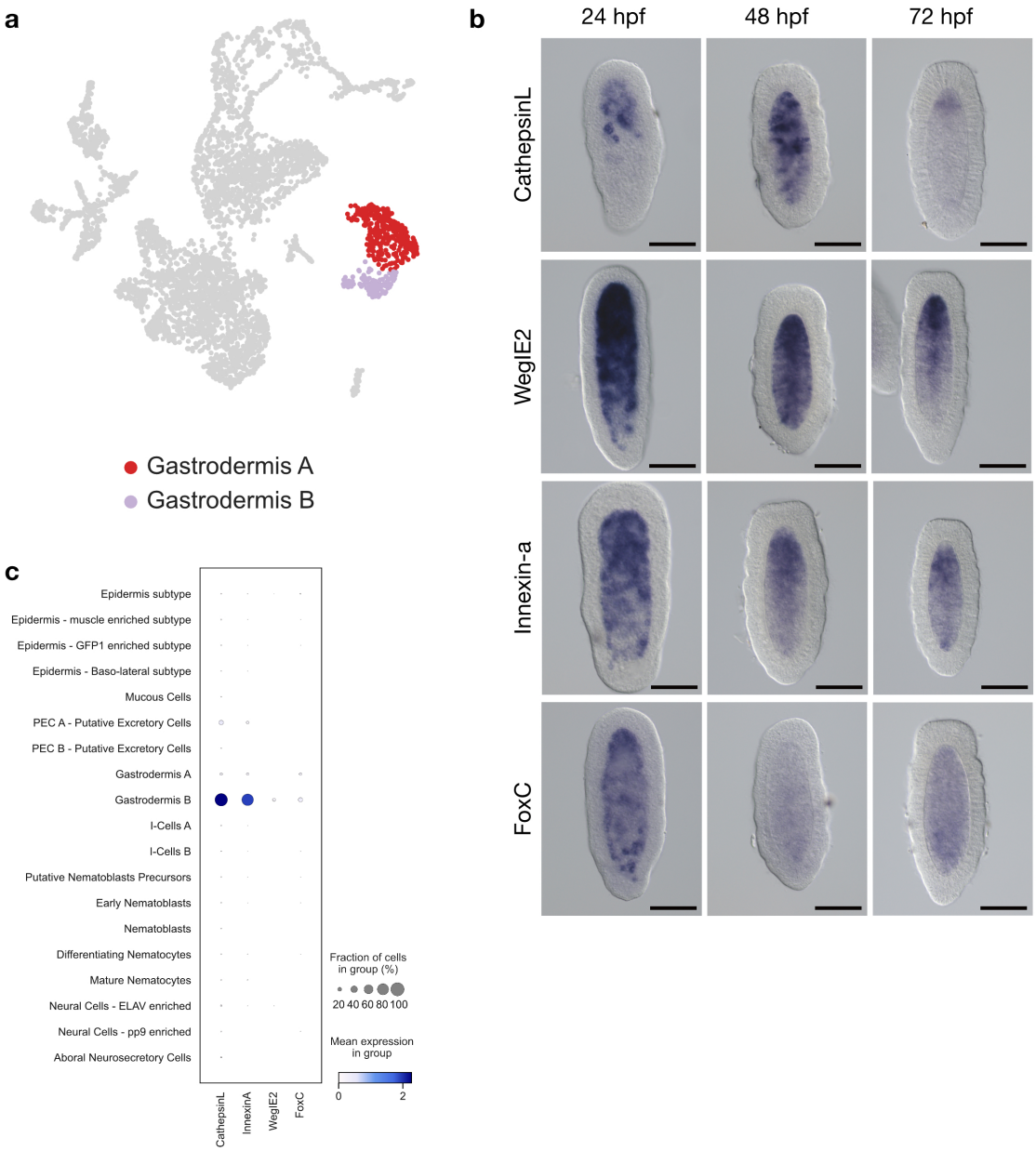


Figure 3.6 - Gastrodermis cells

(a) 2D-UMAP plot coloured by gastrodermis cell types and legend with the cell type annotation. (b) *In situ* hybridization of gastrodermis marker genes at the three planula developmental stages (b) Scale Bars represents 100µm. (c) Dot Plot of expression of gastrodermis marker genes.

3.2.1.3 - Hydrozoan Stem Cells (I-Cells)

The i-cells are hydrozoan stem cells that give rise to nematocytes, neural cells, gland cells and germ-line according to what has been described in *Hydra* (see Chapter 1; Bode, 1996; Watanabe et al., 2009).
In the planula cell atlas I assigned two clusters as i-cells (Fig.3.7a) on the basis of the expression of previously described i-cell marker genes such as *Piwi*, *Nanos1* and *Vasa* (Leclère et al., 2012; Fig. 3.7b and c). Analysis of the gene signatures of the clusters and expression of Piwi in i-cells and in putative nematoblasts precursors (Fig. 3.7c) suggests that a large population of i-cells is committed to producing nematoblasts at this stage of planula development. I attempted to assess whether these two clusters represented stages of i-cells differentiation but could not detect clear differences between the two. One hypothesis is that the two clusters might represent cell states instead of cell types of i-cell committed to become nematocytes and the subtle difference is not clearly recognisable in the data.

In the medusa, nematoblast as well as neural cells derive from a population of i-cell-like precursors. This is corroborated by pseudotime analysis of single cell transcriptomics data of isolated i-cell, neural cells and nematocytes (Chari et al., 2021). Indeed, in the medusa atlas presented in Chapter 2 we assigned two clusters of nematocyte and neural cell precursors.
I could not identify a population of i-cells sharing a gene signature with neural cells, /neural cell precursors in the planula atlas. Analysis of the marker genes did not reveal a striking connection between i-cells and neural cells as they did for nematocyte and i-cells (see below). One hypothesis is that neurogenesis in the planula occurs before the stage we sampled (50-52 hpf). In this case, the differentiation of an i-cell-like precursor would occur earlier, with neural cells completely differentiated at this stage hence not represented in our data. The low number of neural cells might also impact our ability to detect a potential precursor population. Finally, the lack of connection between neural cells and i-cells could reflect a parallel or alternative neurogenesis pathway operating at this stage, but not in the medusa. An alternative neurogenesis pathway that could be operating at planula stage is discussed in Chapter 4.

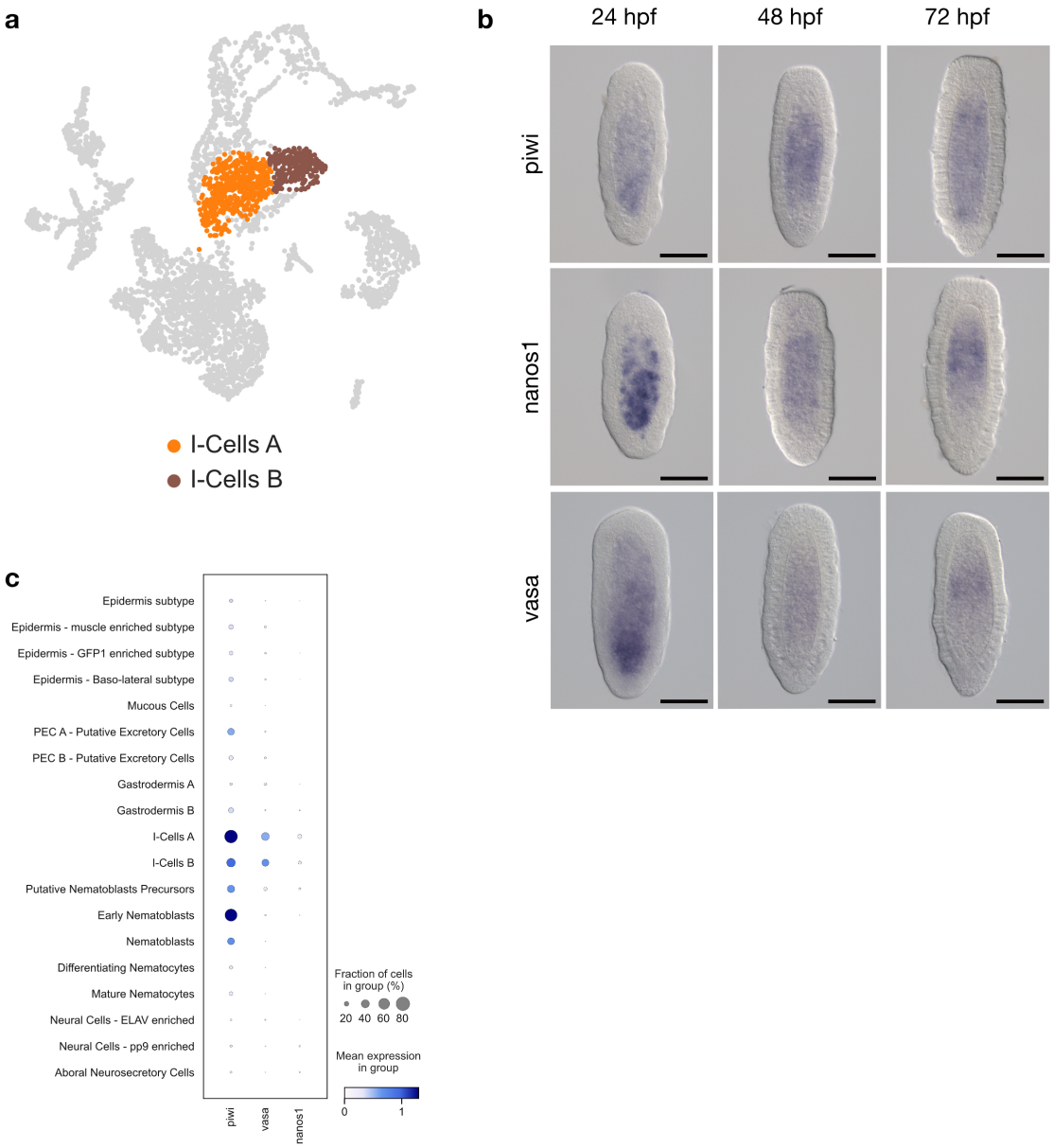


Fig.3.7 - I-Cells

(a) 2D-UMAP plot coloured by i-cell types and legend with the cell type annotation. (b) *In situ* hybridisation of i-cell marker genes at the three planula developmental stages (b) Scale Bars represents 100µm. (c) Dot Plot of expression of i-cell marker genes.

3.2.1.4 - Nematocytes

In the medusa atlas, the large number of cells allowed the characterisation of two phases of nematogenesis (Chari et al., 2021: see Chapter 2). We use the term nematoblast to characterize the cells in the initial phase of nematogenesis during which the stinging capsule (nematocyst) is formed (Chari et al., 2021; Condamine et al., 2019; Denker et al., 2008; see Chapter 2). The second phase of nematogenesis is characterised by migration of nematoblasts towards the distal portion of the medusa tentacle and the formation of the cnidocil (Tardent, 1995) or "nematocil". This is the mechanosensory structure that triggers the ejection of the nematocytes capsule upon stimulus to sting the prey (Chari et al., 2021; Tardent, 1995; see Chapter 2). The two phases of nematogenesis are characterised by distinct and consecutive transcriptional profiles and by spatial distribution of the cells in different developmental stages from proximal to distal parts of the tentacle bulb and then along the tentacle (Chari et al., 2021; see Chapter 2).

In the planula atlas we could identify five clusters corresponding to nematoblasts and nematocytes (Fig. 3.8a) which are grouped into the nematocyte class. Remarkably, the expression of marker genes in common with the medusa stage recapitulated the bi-phasic developmental trajectory (Fig. 3.8b and c). The dataset that I generated, using 2 day-old planula stage, allowed me to reconstruct the nematocyte developmental trajectory. Indeed, all the genes identified as marking different stages of nematogenesis could be mapped onto the atlas. As in medusa, expression of marker genes such as *Znf845*, *Mos3* and *Minicollagen 3/4* (*Mcol3/4)* characterise the nematoblasts phase (Fig. 3.8b). Interestingly, the expression of *Minicollagen 3/4* persists until later stages (Fig. 3.8b, 72 hpf panel) of planula development, when the planula is competent to metamorphose (Fig. 3.8b). This suggests that not all the planula nematoblasts are programmed to transition to mature nematocytes until later stages. Presumably the nematoblasts in the planula that do not differentiate into nematocytes represent a 'reservoir' for the polyp stage after metamorphosis. This observation is also consistent with previous observations of ongoing production of nematoblasts in later stages of planula development (Bodo and Bouillon, 1986).
The second phase of nematogenesis is characterised by the expression of M14-peptidase and Nematocilin, as in medusa stage (Chari et al., 2021). M14-peptidase expression marks the initial phase of differentiation into mature nematocytes in medusa (Chari et al., 2021). Surprisingly, this marker is expressed in fewer cells at early stages of planula development (24 hpf; Fig. 3.8b) and the signal is weaker than nematoblast genes. This is consistent with the idea that only few nematoblasts go through complete differentiation, with the transition to the second phase of nematogenesis occurring in early stages of planula development, and a large reservoir of nematoblasts persisting (Fig. 3.8b).

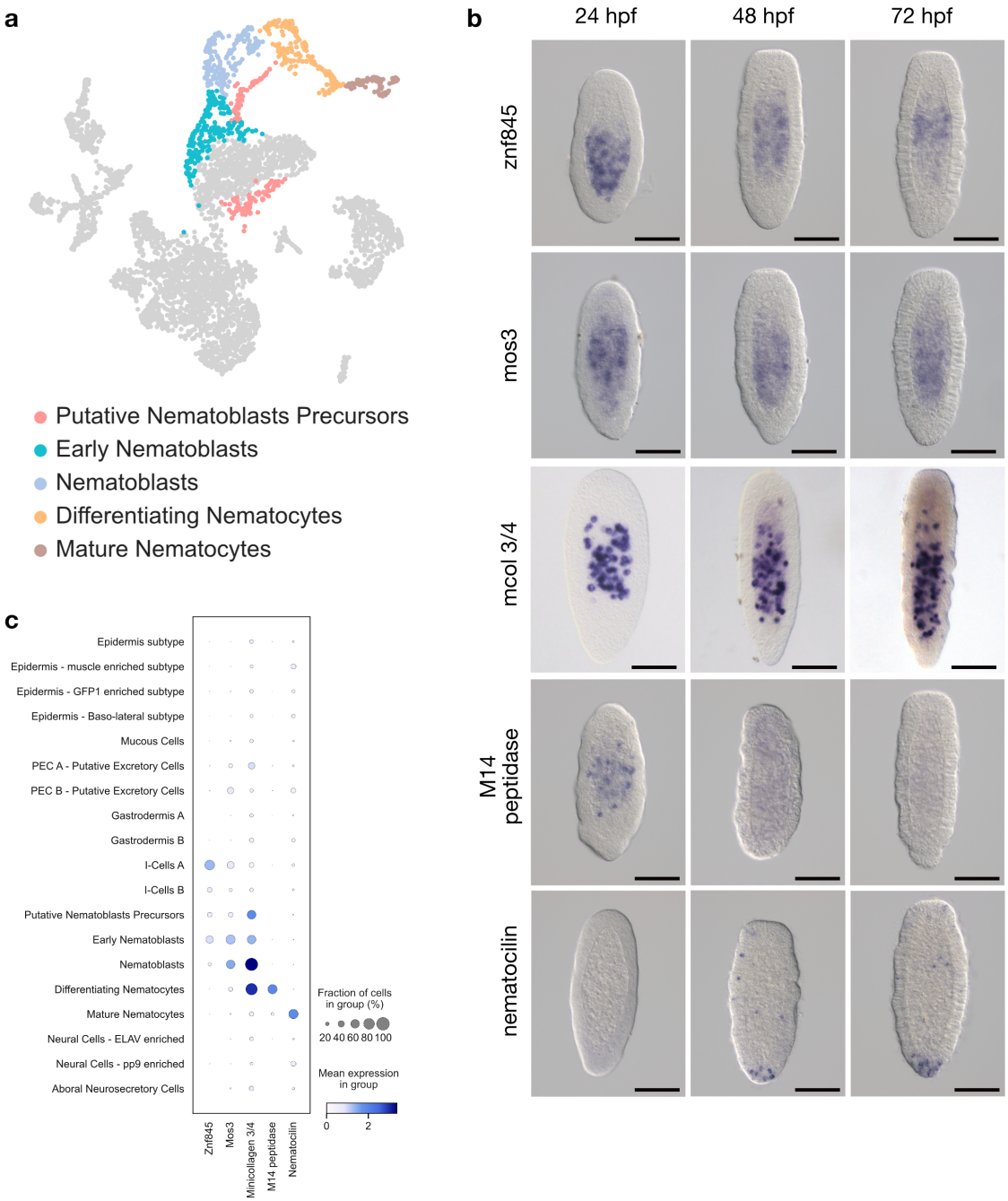


Figure 3.8 - Nematocytes

(a) 2D-UMAP plot coloured by nematocyte cell types and legend with the cell type annotation. (b) *In situ* hybridization of nematocyte marker genes at the three planula developmental stages (b) Scale Bars represents 100µm. *In situ* hybridization of Mcol 3/4 by Antonella Ruggiero. (c) Dot Plot of expression of i-cell marker genes. The images of Nematocilin at 48 and 72 hpf are generated by merging of several Z planes (c) Dot Plot of expression of nematocyte marker genes.

3.2.1.5 - Neural Cells and Aboral Neurosecretory Cells

In the planula cell type atlas I could identify two clusters of neural cells (Fig. 3.9a) expressing previously described markers such as ELAV (Nakanishi et al., 2012) and Pp9 (Chari et al., 2021; Fig. 3.9b, c, d, e). ELAV is expressed in a scattered population of neural cells (Fig. 3.9b). Analysis of the transcriptome signature of ELAV-enriched neural cells indicates that this cluster is a mix of putative neural subpopulation expressing previously described and predicted neuropeptide precursors such as Pp5, Pp11, Pp2 and others (Chari et al., 2021). Neural cells expressing *Pp9* are localised in the aboral half of the larva and the expression of the gene is higher at later stages of planula development indicating a differentiated neural cell population with potential specialised function.

A distinct population of neurosecretory cells is localised at the aboral pole (Fig. 3.9b). The transcriptome signature of these cells is enriched in genes encoding for secreted proteins expressed specifically at the planula stage (Fig. 3.9 e, f, g). The aboral neurosecretory cells show a particular “hourglass” shape with the elongated cellular body in contact with the mesoglea on the basal side and embedded in between epithelial cells. The nucleus is generally located in a basal position. The centre is shrunken and compressed between epithelial cells and it exhibits a slight enlargement at the apical side as shown in the electron micrograph in Figure 3.10. The presence of small vesicles throughout the cytoplasm, especially in the apical region, argues in favour of the classification as secretory cells (Fig. 3.10, arrows). Along with the presence of neural related components in the gene signature, the location of these cells at the aboral pole of the planula and the distinctive shape support the involvement in neural-like functions (Sinigaglia et al., 2015). Likely, the cells we designate as aboral neurosecretory cells correspond to those described by Bodo and Bouillon (1968) as ‘foamy’ glandular cells containing multiple irregular secretion vacuoles (Bodo & Bouillon, 1968; see Chapter 1).

I performed reclustering of neural cells and neurosecretory cells to search for potential subpopulations, as done for the medusa (Chari et al., 2021). For this I repeated preprocessing and clustering analysis (see Methods) on neural and neurosecretory cells only. This reclustering analysis included 360 cells which were computationally grouped into 10 clusters. (Fig. 3.11a). I extracted marker genes following the same approach used for the whole atlas to validate putative subpopulations (see Methods) and identified a few additional markers (Fig. 3.11b). However, the clusters obtained from this rather low number of cells probably do not correspond to discrete planula neural subpopulations. One indication of this was that the expression of some neuropeptide precursors (Pp5 and Pp9) that were previously observed to be specific for neural subpopulations in medusa (Chari et al., 2021), spanned more than one cluster (Fig. 3.11c). Furthermore, it was not always possible to select cluster specific genes (i.e. for clusters 7, 9 and 4, Fig 3.11c). This suggests that some clusters represent uninformative groupings of cells during reclustering analysis, very likely caused by the low number of cells. Nonetheless this reclustering atlas was useful for the identification of few new neural-specific genes and for the characterisation of some putative planula neural

subpopulation expressing Pp14, the transcription factor Nk2d and a putative neuropeptide (Fig. 3.11b).

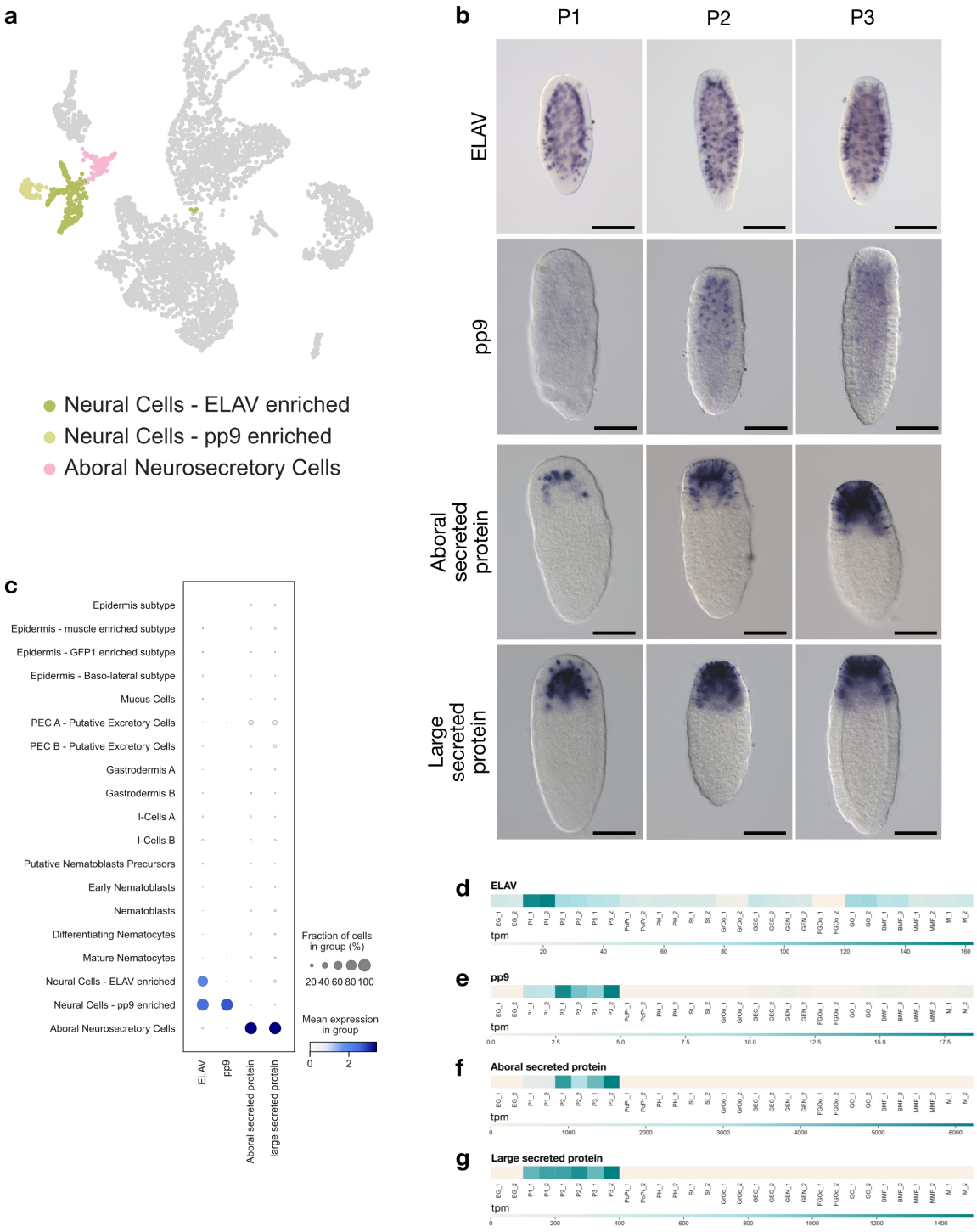


Figure 3.9 - Neural and Neurosecretory Cells

(a) 2D-UMAP plot coloured by neural and neurosecretory cell types and legend with the cell type annotation. (b) *In situ* hybridization of epidermis marker genes at the three planula developmental stages. Scale Bars represent 100µm. *In situ* hybridization of ELAV by Sandra Chevalier. (c) Dot Plot of expression of neural and neurosecretory marker genes. (d, e, f, g) Expression of neural and neurosecretory marker genes across developmental stages of *Clytia*. Transcripts per million (tpm) calculated from bulk RNA-seq data from Leclère et al., (2019). EG = Early gastrula, P1 = Planula 24

hpf = Planula 48 hpf, P3 = Planula 72 hpf, St = Stolon, PoPr = Primary polyp, PH = Polyp head, GO = Gonozooid, BMF = Baby medusa female, MMF = Mature medusa female, M = Mature medusa male

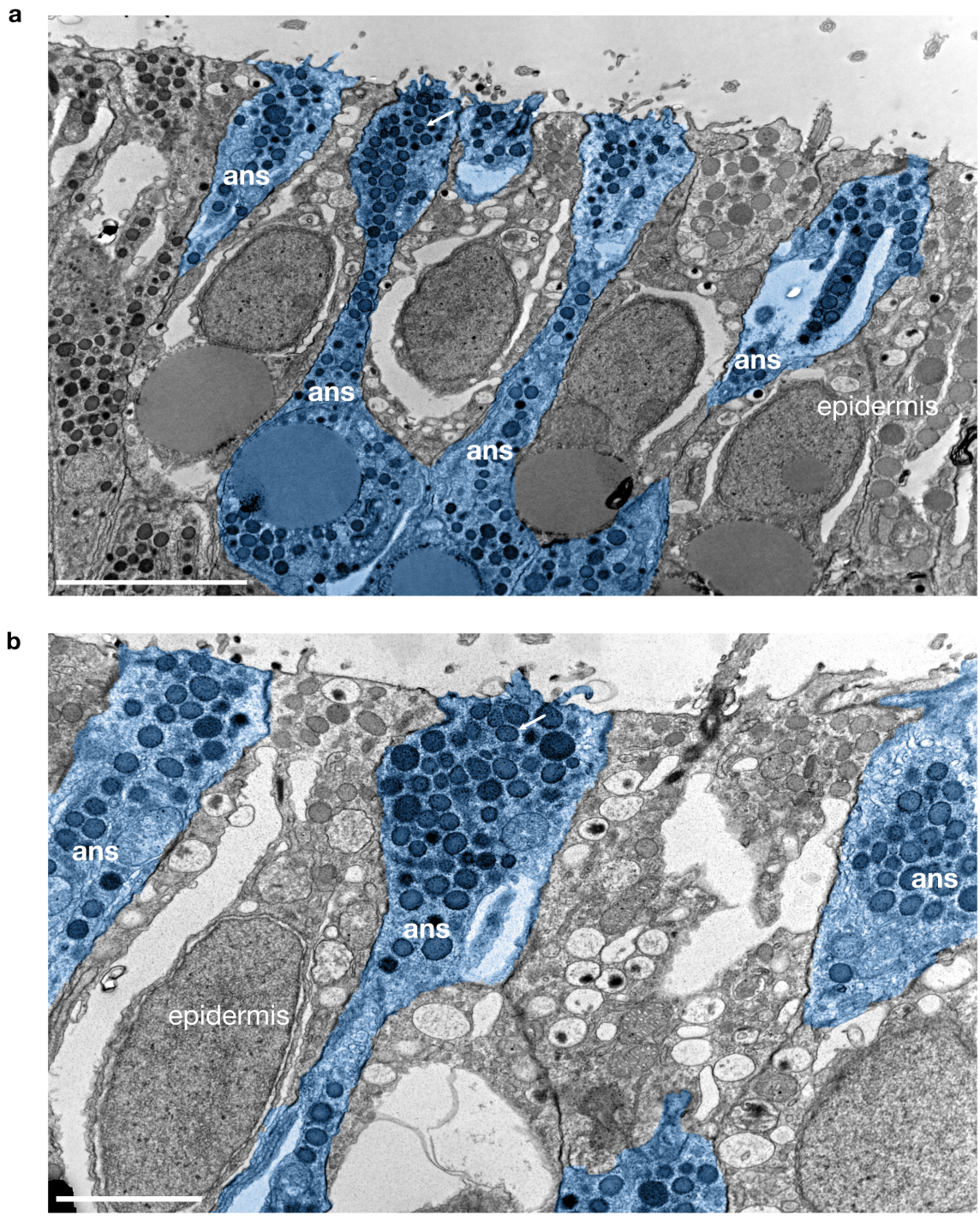


Figure 3.10 - Ultrastructure of aboral neurosecretory cells
Electron micrographs of sections through the aboral epidermis of a 3-day-old *Clytia* planula. (a) Organisation of aboral neurosecretory cells (ans, blue) in the aboral epidermis of the larva. Putative neurosecretory vesicles are concentrated at the apical side of the cells (arrow). (b) Higher magnification of another section of the aboral epidermis of the larva showing apical side of neurosecretory cells. Scale bar in (a) represents 5µm. Scale bar in (b) represents 2µm.

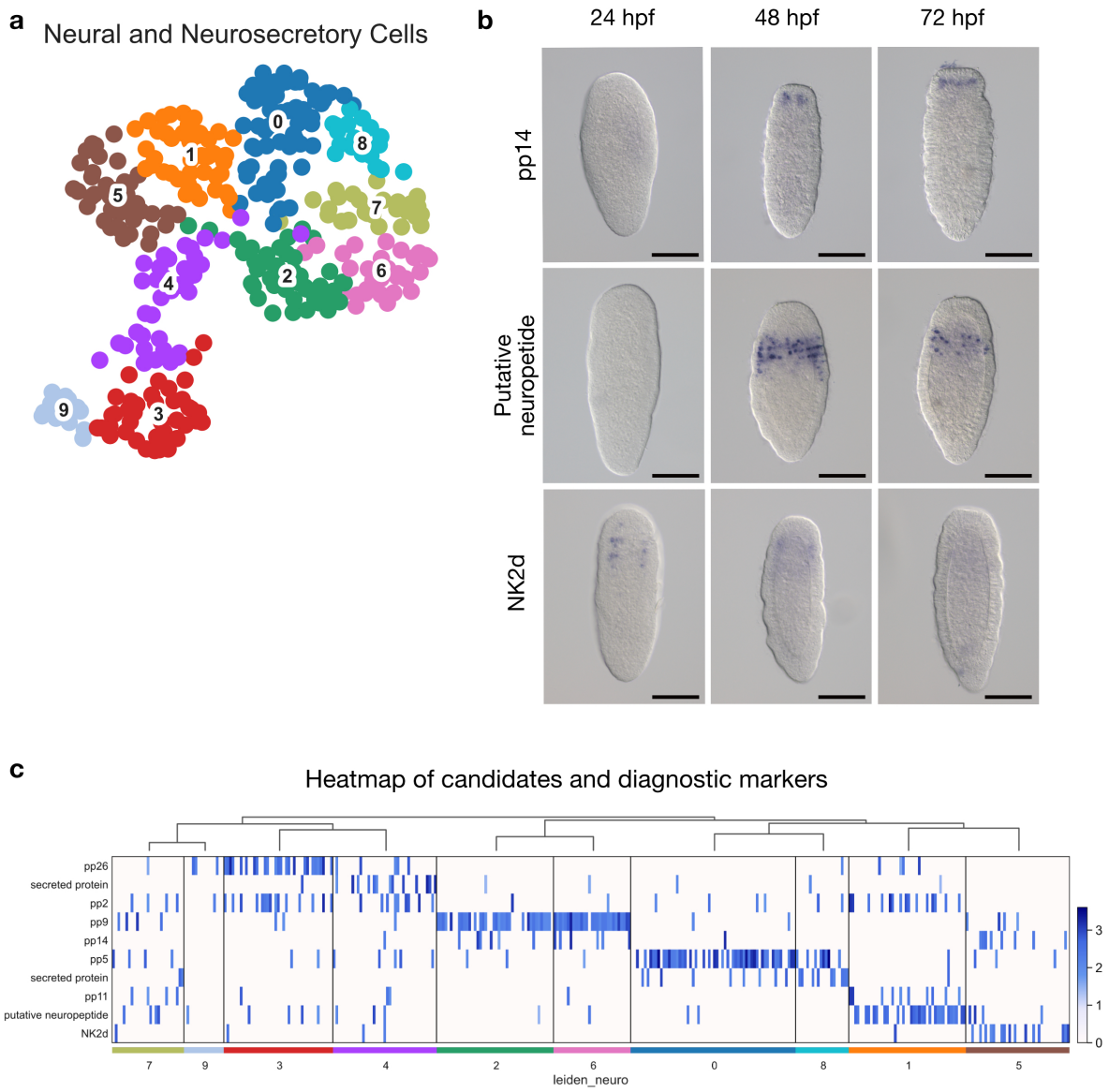


Figure 3.11 - Reclustering of neural and neurosecretory cells
(a) 2D-UMAP plot corresponding to Leiden clusters following reclustering analysis. (b) *In situ* hybridization of candidates at the three planula developmental stages. Scale bars represent 100μm. The image forPp14 and 'Putative neuropeptide' at 48 and 72 hpf are generated by merging of several Z planes. (c) Heatmap of expression of selected candidate genes across clusters..

3.2.1.6 - Mucous Cells

I assigned cluster 6 as mucous cells (Fig. 3.12a). Very likely this cell type corresponds to the mucous cells previously observed and described (Bodo and Bouillon, 1986; see Chapter 1) as glandular spheroulous cells, characterised by a muco-proteinic secretion. The transcriptional signature is characterised by highly expressed secreted proteins, including novel glycoproteins.

One of the candidate genes that we used for *in situ* experiments codes for is a C-terminal cysteine-knot-domain-containing protein (CtCK, Fig. 3.12b), a domain found in several mucins (Bythell & Wild, 2011). We confirmed the expression pattern with another marker, the O-linked-mannose beta-1,2-N-acetylglucosaminyltransferase-1 protein (POMGNT1), that shows expression highly specific to cluster 6, (Fig. 3.12b). Both markers showed that mucous cells are distributed through the planula epidermis, in lateral and oral regions, as well as in cells of the underlying basal epidermis. Several genes, including ones with predicted functions in vesicle trafficking are shared by this mucous cell cluster and cluster 2, the baso-lateral epidermis subtype. One hypothesis to explain these common features is that the mucous cells derive from a basal layer of ectoderm during development, along with some of the aboral neural and neural secretory cell types (See Chapter 4). The ultrastructure of mucous cells is clearly distinct from the epithelial cells (Fig. 3.13). The cytoplasm of the apical side faces the external space and is filled with numerous vesicles, similarly to descriptions of mucocytes in corals (Bakshani et al., 2018). It has been suggested that mucous in cnidarians, especially in corals, might be involved in various functions including feeding, protection against microbes and even larval settlement (Bakshani et al., 2018; Brown & Bythell, 2005).
Although we are not sure of the role of the mucous cells in the planula of *Clytia*, the high expression of many diagnostic genes specific to this stage (Fig. 3.12d and e) indicates that mucous cells are a specialised cell type of the planula, and might play a specific role at larval stage, for instance in interacting with biofilm substrates ahead of settlement.

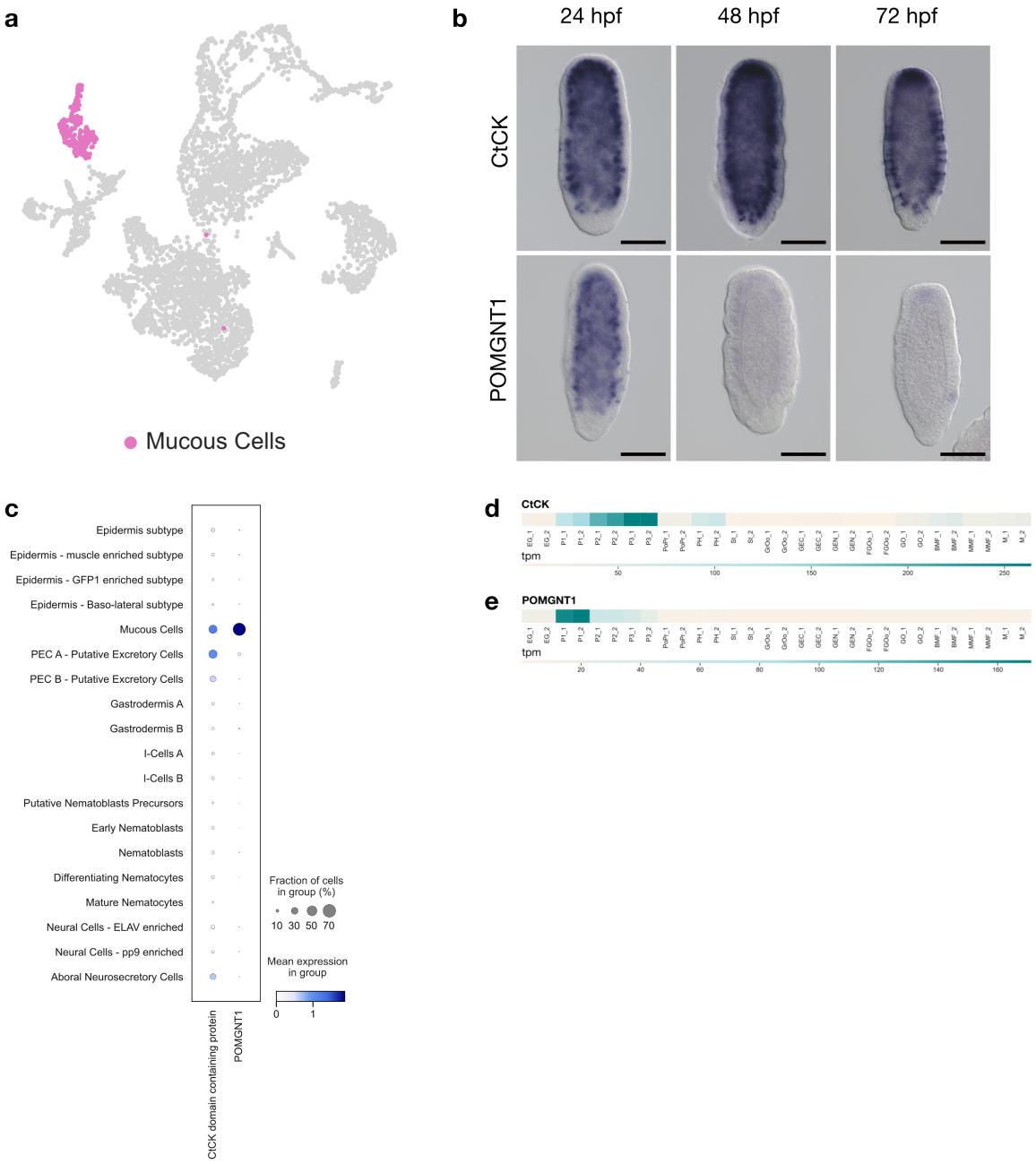


Figure 3.12 - Mucous cells

(a) 2D-UMAP plot coloured by mucous cell cluster and legend with the annotation. (b) *In situ* hybridisation of mucous cells marker genes at the three planula developmental stages. Scale Bars represent 100µm. The images for CtCK at 24, 48 and 72 hpf and og POMGNT1 at 24 hpf are merged from several z-planes. (c) Dot Plot of expression of mucous cells marker genes. (d, e, f, g) Expression of mucous cells marker genes across developmental stages of *Clytia*. Transcripts per million (tpm) calculated from bulk RNA-seq data from Leclère et al., (2019). EG = Early gastrula, P1 = Planula 24 hpf = Planula 48 hpf, P3 = Planula 72 hpf, St = Stolon, PoPr = Primary polyp, PH = Polyp head, GO = Gonozooid, BMF = Baby medusa female, MMF = Mature medusa female, M = Mature medusa male

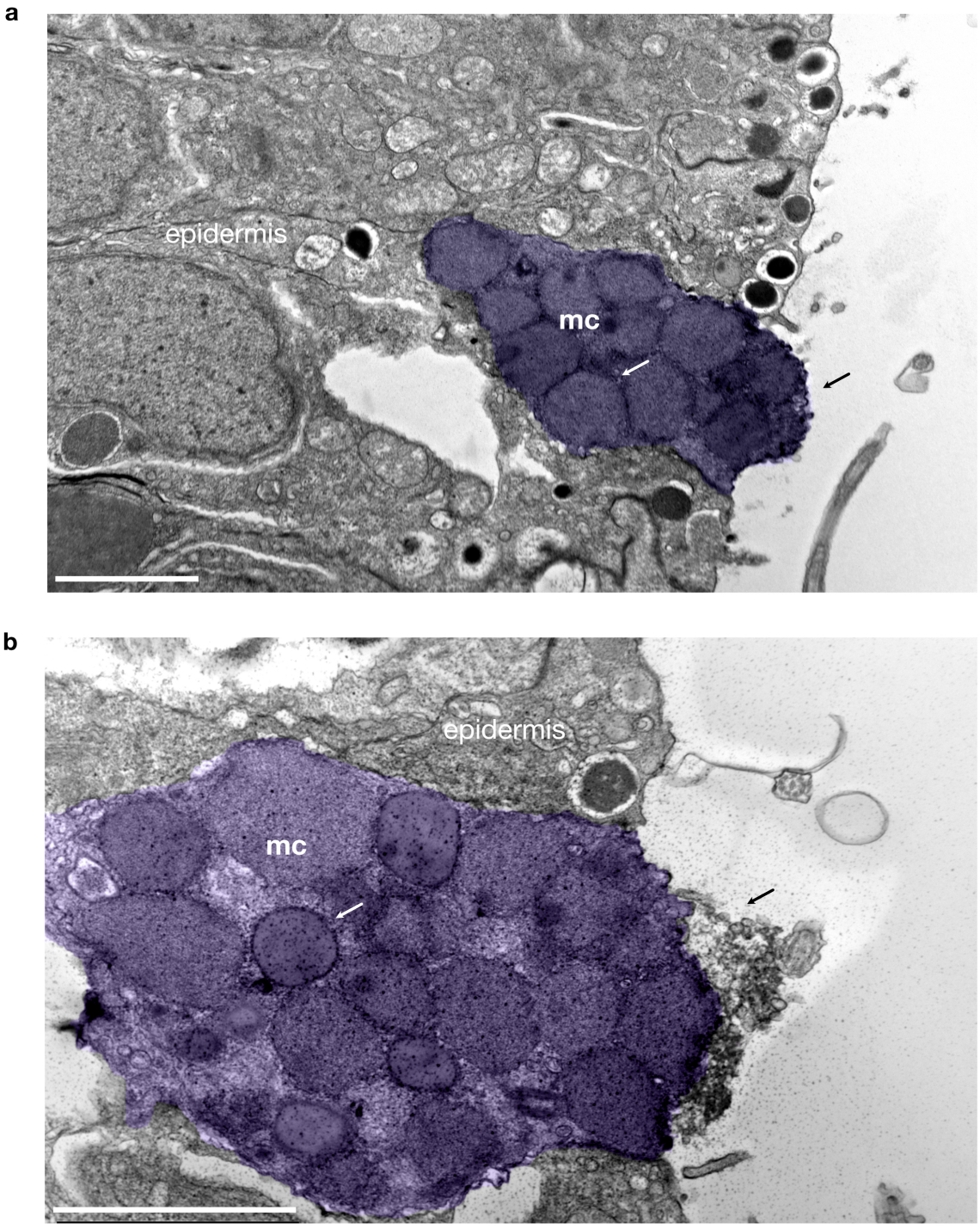


Figure 3.13 - Ultrastructure of mucous cells

Electron micrographs of sections through the epidermis of a 3-day-old *Clytia* planula. (a) isolated mucous cell (mc, purple) in the epidermis of the larva. White arrow indicates a secretory vesicle. Black arrow indicates the apical side of the cell which protrudes from the epidermis with intact apical membrane. (b) Higher magnification of another section of the epidermis of the larva showing isolated mucous cell (mc, purple). White arrow indicates a secretory vesicle. Black arrow indicates the apical side of the cell with irregular apical membrane which likely corresponds to an advanced stage of secretion compared to (a). Scale bars in (a) and (b) represent 2µm.

3.2.1.7 - Putative Excretory Cells (PEC)

Clustering analysis revealed two small clusters of cells that did not correspond to any of the medusa cell types. We labelled these as Putative Excretory Cells or PEC (Fig.3.14a) based on correspondence in position and morphology with previously described larval cell type (Bodo & Bouillon, 1968).

Analysis of the marker gene signature of these two clusters revealed several unclassified putative secreted proteins. Furthermore, the two cell types, designated PEC-A and PEC-B, show highly similar transcriptional profiles, coding for proteins containing Trypsin domains, Shk domains, and collagen domains as well as cadherins. Additionally, PECs express numerous proteins with spondin and thrombospondin domains, suggesting a possible association with cell adhesion and extracellular matrix.

Selected marker genes for *in situ* hybridisation included two *Clytia* specific and cell type specific genes for each cluster (chePEC-A and chePEC-B, Fig. 3.17b and c) strongly expressed in isolated gastrodermal cells in particular at 48 and 72 hpf. Similar expression patterns are obtained by *in situ* hybridisation of Collagen triple helix containing protein (CTHRC), a specific marker for PEC-A cells and thrombospondin type-1 repeat-containing protein (TSP type-1), a specific marker for PEC-B cells (Fig. 3.14c and d). All the selected candidates also show expression in the stolon of the polyp colony (Fig. 3.14e, f, g and h) indicating a possible role of these cells beyond planula stage.

A putative excretory cell type was described by Bodo and Bouillon (1986; see Chapter 1) as unique to later planula stages. These cells were mainly found in the ectoderm, and were characterised by the presence of putative excretion granules filled with possible products of yolk digestion (Bodo & Bouillon, 1968). The authors speculated that these cells participate in the digestion of yolk and other cellular components in the endoderm through endocytosis, and then migrate out through the ectoderm to excrete the contents.
To explore this hypothesis I carried out *in situ* hybridisation for CTHRC and TSP type-1 at 96 hpf (i.e. in 4 day-old planulae). The CTHCR positive cells appeared to relocate from the endoderm to the ectoderm at later stages (Fig. 3.17d). Interestingly, CTHCR expression persists at 96 hpf while TSP type-1 expression is almost completely undetectable (Fig. 3.17d). The persistence of PEC-A but not of PEC-B cells in ageing planulae suggests that they might have distinct functions.

I propose a putative excretory classification for the two clusters based on analysis of the expression patterns of candidate genes and similarities with previous descriptions (Bodo & Bouillon, 1968). However, additional single cell data and functional studies are needed in order to understand the nature of these cells.

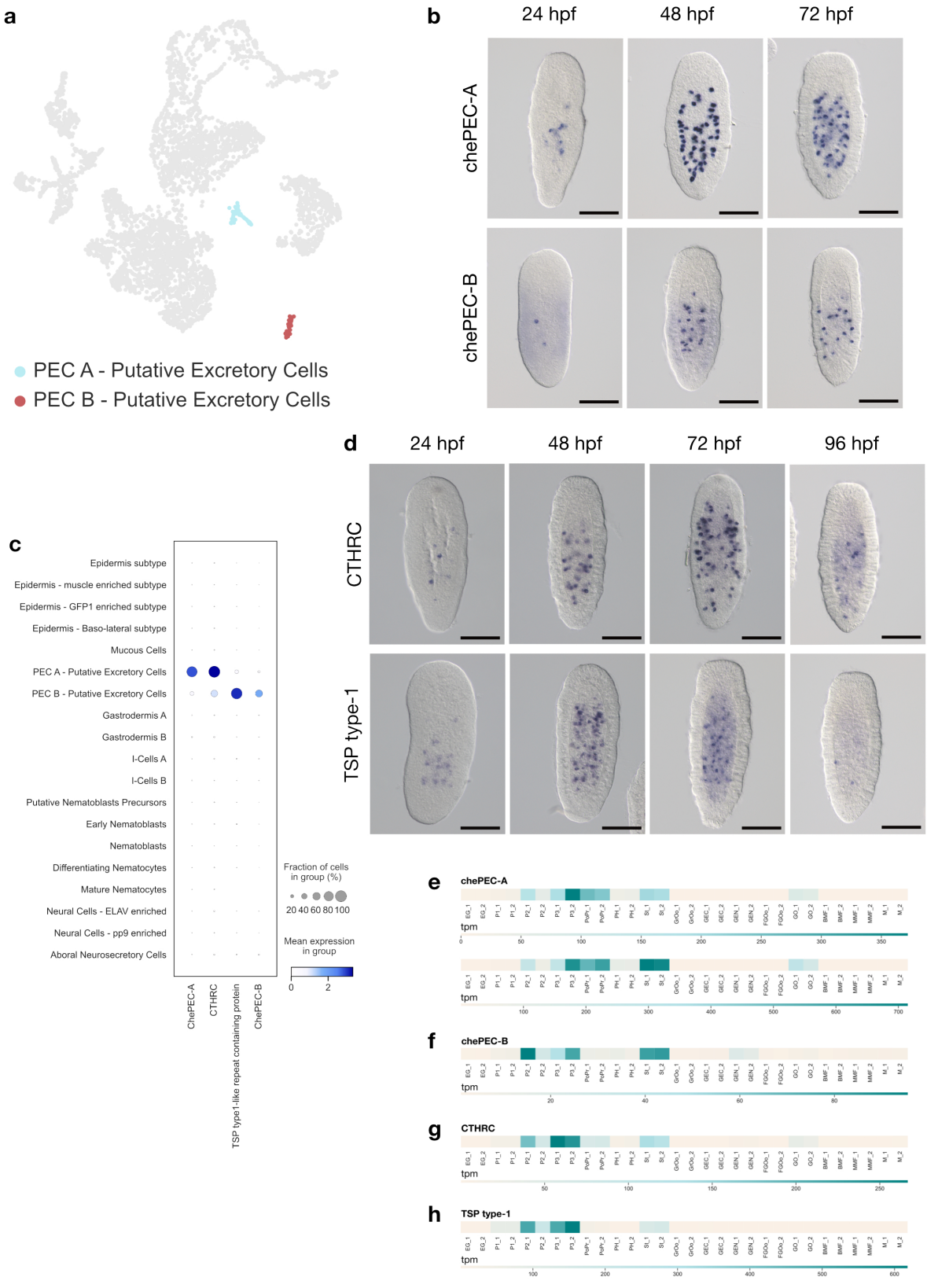


Figure 3.14 - Putative Excretory Cells

(a) 2D-UMAP plot coloured by putative excretory cells cluster and legend with the annotation. (b) *In situ* hybridization of putative excretory cells marker genes at the three planula developmental stages. Scale Bars represent 100μm. The images for chePEC-A and , chePEC-B at 24, 48 and 72 hpf as well as

CTHRC and TSP type-1 at 48, 72 and 96 hpf are merged from several z-planes. (c) Dot Plot of expression of PEC marker genes. (d) *In situ* hybridization of putative excretory cells marker genes at four planula developmental stages. Scale Bars represent 100µm. (e, f, g, h) Expression of PEC marker genes across developmental stages of *Clytia*. Transcripts per million (tpm) calculated from bulk RNA-seq data from Leclère et al., (2019). EG = Early gastrula, P1 = Planula 24 hpf = Planula 48 hpf, P3 = Planula 72 hpf, St = Stolon, PoPr = Primary polyp, PH = Polyp head, GO = Gonozooid, BMF = Baby medusa female, MMF = Mature medusa female, M = Mature medusa male

3.3 - METHODS

3.3.1 - Clytia planula culture

Fertilisation was carried out by mixing gametes from adult medusae (Z23 female strain and Z4C male strain). Planula larvae were cultured in filtered sea water with 1:25000 penicillin and streptomycin solution at 17°C. (see Chapter 1; Lechable et al., 2020))

3.3.2 - Clytia planula dissociation and cell fixation

Larvae were collected after 50 hpf and washed with Ca/Mg-free artificial seawater three times by serial transfer from sea water using medium sized petri dishes. 500 to 700 larvae were used for each experiment. Single cell dissociation was carried out incubating the larvae in Ca/Mg-free artificial sea water (Table 3.1) and 40U/ml SUPERase•In™ RNase Inhibitor (20 U/µL, Invitrogen) for ten minutes in a small sized plastic petri dish within a 40µm cell strainer and observed with a dissecting microscope (Fig. 3.15a).
After incubation, excess sea water was removed from the strainer and larvae were gently pressed against the mesh using the rubber plunger from a 1ml syringe (Fig. 3.15b). The cell strainer was washed several times with 100 µl of Low-Ca artificial sea water (Table 3.2; (Fig.3.15b) and 40U/ml SUPERase•In™ RNase Inhibitor (20 U/µL). Single cells were collected in a tube by pipetting the solution from the bottom of the strainer (Fig. 3.18c) until 1ml of cell suspension was collected.
To estimate cell concentration, 20 µl of cell suspension were stained with 1:1000 Hoechst 33258 (1 mg/ml stock; Sigma-Aldrich).
Cells were counted manually at A2 Axio Imager (Zeiss) binoculars with the use of a Neubauer improved counting chamber (Sigma-Aldrich BR717810-1EA) filling the two grids available with 10 µl of cell suspension each.
Cell mortality was estimated during previous separate test experiments diluting the cell suspension with 1:1 ErythrosinB 0.5mg/ml solution, which labels dead cells.
The mortality data were used to develop the final dissociation protocol with the aim of reducing cell mortality to 10-20% maximum.
Dissociated cells were immediately fixed with 80% ice cold methanol as in (Chari et al., 2021) or ice cold ACME solution (García-Castro et al., 2021; Table 3.3) and 40U/ml SUPERase•In™ RNase Inhibitor (20 U/µL), then incubated for 30 min at -20°C or on ice, depending on the fixation solution.
After 30 min, PBS-1%BSA and 40U/ml SUPERase•In™ RNase Inhibitor (20 U/µL) was added 1:1 to the cell suspension and mixed gently by pipetting in order to mobilise the cells. Cells were split in 3 or 4 aliquots depending on initial cell concentration and centrifuged for 10 min at 700 rcf in a swinging bucket centrifuge.
Supernatant (900µl) was discarded and the cell pellet was resuspended with 800µl of PBS-1%BSA and 100U/ml SUPERase•In™ RNase Inhibitor (20 U/µL).

To estimate cell concentration in each aliquot, 20 µl of cell suspension were stained with 1:1000 Hoechst 33258 (1 mg/ml stock; Sigma-Aldrich). Cells were counted manually using the Neubauer chamber as described above.
Finally, DMSO was added to a final concentration of 10% to preserve the cells during freezing. Repeated dissociation and cell counting experiments provided enough data to estimate that a 2-day-old larva consists of about 3000 to 4000 cells.

The solution was snap-frozen in liquid nitrogen and stored at -80°C until the day of the encapsulation.
Several aliquots containing at least 500 cells/µl fixed with ACME and 80% Methanol were collected over several days and finally shipped to the Sebé-Pedros lab at the CRG, Barcelona, where I went to perform encapsulation trials using different platforms (see below)

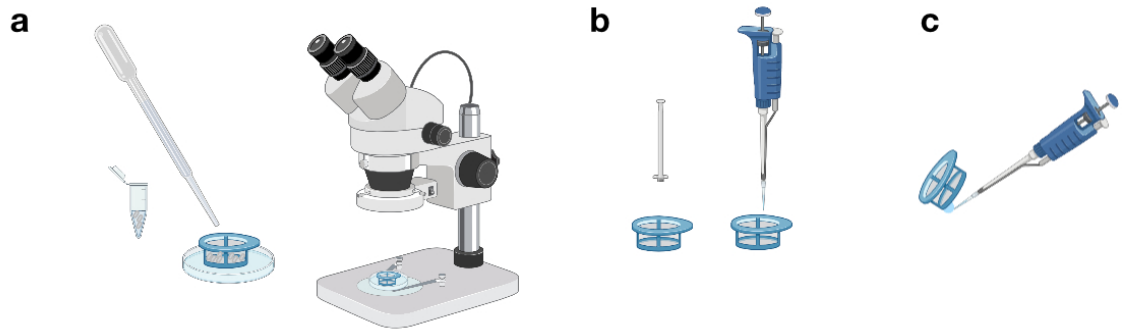


Figure 3.15 - Dissociation protocol
Diagram showing (a) larvae of *Clytia* collected in a tube and transferred to a plastic petri dish within a 40 µm cell strainer for observation with a binocular. (b) Dissociation with the plunger of a 1ml syringe and repeated washing. (c) Collection of single cell suspension from the bottom part of the cell strainer. Created with BioRender.com

Calcium/Magnesium - Free Artificial Sea Water	
NaCl	31 g
KCl	0.8 g
NaHCO3	0.29 g
Na2SO4	1.6 g
MilliQ H2O	Up to 1 litre

*final pH 8.0
Table 3.1 - Calcium/Magnesium - Free Artificial Sea Water

Low Calcium/Magnesium-Free Artificial Sea Water	
NaCl	26.88 g
KCl	0.74 g
CaCl2	0.16 g
HEPES	2.38 g
MilliQ H2O	Up to 1 litre

*final pH 7.6

Table 3.2 - Low-Calcium/Magnesium - Free Artificial Sea Water

ACME fixattive solution	
Low Calcium/Mg-Free ASW	2.25 ml
Glycerol	0.5 ml
Glacial acetic acid	0.5 ml
Ice cold 100% Methanol	0.75 ml

Table 3.3 - ACME fixative solution

3.3.3 - Cell Sorting, Encapsulation, Preparation of single cell libraries

Planula single cell aliquots were processed and sorted in collaboration with Marta Iglesias Garcia at CRG in Barcelona. Cells were encapsulated using the 10X Genomics Chromium system and cDNA libraries were prepared according to the Chromium Next GEM Single Cell 3′ library preparation protocol v3.1.

Cells were thawed on ice and tubes were centrifuged at 1000xG in a swinging bucket centrifuge. The supernatant was discarded and the cell pellet was washed with PBS-0.5%BSA and 40U/ml RNase inhibitor. Cells were washed two times to eliminate the DMSO. 3-4 aliquots were combined according to previous estimated cell concentration in each tube and the cells were finally transferred to Low Binding tubes in a total volume of 500ul PBS-0.5%BSA with 100U/ml RNase inhibitor.

To estimate cell concentration, 6 µl of cell suspension were stained with 1:1 PBS-1:100 DAPI and counted using a Countess machine (Invitrogen) .
The remaining cells were stained in the dark with nuclei and membrane staining (1:300 DRAQ5, Stock 5 mM, and 4 µl/mL Concanavalin-A, Stock 1 mg/ml) and finally FACSorted.
FACSorted target populations were collected into a round-bottom-96-well plate containing RT buffers prepared according to the Chromium Next GEM Single Cell 3′ library preparation protocol v3.1. Finally the mix was transferred to low binding tubes and reverse transcriptase enzyme was added. Cells encapsulation targeting 4000 to 6000 cells and cDNA library preparation according to the Chromium Next GEM Single Cell 3′ library preparation protocol v3.1 was performed by the CRG Genomics facility.
Sequencing of single cell libraries was carried out on NextSeq 150 Midoutput and NextSeq 75 High output kits at the CRG Genomics facility (see Troubleshooting).

3.3.4 - Mapping and Clustering analysis

Demultiplexing of sequencing output and quality check was carried out by the CRG Genomics facility.

Once I received the sequencing files (fastq files), I mapped the reads individually for each sample against the genome reference (Leclére et al., 2019) using STARSolo (Dobin et al., 2013). The gene reference including annotation of mitochondrial genes was made with STAR according to the STAR manual recommendation. Only cells with more than 100 UMIs were retained for the following steps. The resulting cell x gene count matrix was analysed with Scanpy 1.8.0 (Wolf et al., 2018) and Jupyter Notebook (https://jupyter.org/) .

Preprocessing and quality checking were performed according to the Scanpy tutorial (https://scanpy-tutorials.readthedocs.io/en/latest/pbmc3k.html) with the following modification: I concatenated the 6 libraries and filtered out 4760 low quality cells with less than 500 counts along with cells with more than 1% of mitochondrial gene counts, which are assumed to be poor quality cells (Lun et al., 2016). Additionally I filtered out 3619 genes that were detected in less than 2 cells. I decided on the final filtering thresholds after several tests.
Counts per cell were normalised, log-transformed and highly variable genes were computed retaining genes with a minimum mean value of 0.0125 and a maximum mean value of 3. Counts were scaled and Principal Components (PCs) were computed on highly variable genes. PCs were ranked by variance and the Neighbour graph was computed with 20 PCs and 20 neighbours. The Neighbour graph was embedded in two dimensions using UMAP (McInnes et al., 2018). Clustering of the Neighbour Graph was performed with the Leiden algorithm (Traag et al., 2019) and it resulted in 22 clusters. No batch effect was observed between the libraries at this stage (Fig. 3.1).
Candidate genes were selected from a list of 1000 genes for each cluster and ranked with Scanpy rank_genes_groups function using the Wilcoxon non parametric test (Mann & Whitney, 1947).

Clusters were annotated as cell types using previously described and novel genes (Fig. 3.2). Cluster 0 and Cluster 3 could not be assigned as cell types due to absence of specific marker genes, (see Troubleshooting section).
We excluded the corresponding barcodes and recomputed a final atlas using the same parameters described above, starting from the raw data object which contained all the genes not filtered out in the preprocessing.
We observed a rearrangement of the cells in 19 clusters and we re-annotated the cell types according to the expression of marker genes (Fig. 3.3). Cell types were manually grouped into eight Cell Classes based on gene expression pattern analysis (Fig.3.4)

3.3.5 - In situ Hybridisation

Larvae were fixed at 24, 48, 72 and 96 hours post fertilisation in a cold solution of 3.7% formaldehyde and 0.2% glutaraldehyde in PBS on ice for 40 min as described for medusa stage (Chari et al., 2021; see Chapter 2). Specimens were then washed with PBST (PBS + 0.1%Tween 20), dehydrated stepwise in methanol on ice, and stored in 100% methanol at –20°C.
In situ hybridisation was performed in a robot (Intavis AG, Bioanalytical Instruments) as previously described (Chari et al., 2021). Probes were generated by polymerase chain reaction (PCR) from cDNA clones of our expressed sequence tag (EST) collection (Chari et al., 2021; Chevalier et al., 2006) and from cDNA of planula. Oligonucleotide primers were designed with PrimerBlast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) for POMGNT1, Pp9, Pp14 nad Pp27 probes and successfully cloned in pGemT-easy vector according to the manufacturer recommendations. Sequences of the oligonucleotide primers and EST identification names are provided in Table C which is included in the annexed electronic documents for this manuscript.

3.3.6 - Transmission Electron Microscopy

Fixation of 48 and 72hpf larvae, embedding in Epoxy resin and sectioning for TEM were performed by Sophie Pagnotta at the Plateforme Commune de Microscopie Électronique, Université Côte d'Azur using the Osmium/glutaraldehyde fixation method described by Eisenman & Alfert, (1982). With Evelyn Houliston, I joined Sophie Pagnotta for the imaging sessions at the Plateforme Commune de Microscopie Électronique, Université Côte d'Azur to identify and document the cell types.

3.4 - TROUBLESHOOTING

3.4.1 - Cell dissociation troubleshooting

To dissociate planula cells I initially applied the same dissociation protocol I developed for the medusa stage. I performed a first encapsulation experiment by loading 10.000 freshly dissociated cells using the 10X Chromium System and prepared the libraries according to 10X Genomics technology recommendations (https://www.10xgenomics.com).
I performed downstream analysis of the sequenced reads using Cellranger (10x Genomics Cell Ranger 3.0.0) and the output revealed a recovery of around 70.000 barcodes (Fig. 3.16a) while I expected less than 10.000 cells. This indicates that almost all the droplets were saturated with transcripts. Clustering analyses revealed a non structured UMAP plot where almost no distinct clusters could be obtained. Analysis and visualisation of known marker genes indicated mixed profiles.

I then developed a gentler dissociation protocol making sure to not excessively damage the cells and that no RNA degradation occurred during the process of dissociation. I repeated the encapsulation experiment as described above and obtained a similar output. I tried to salvage the data by extracting a matrix containing UMI counts per barcode to search for a threshold that would allow only 'true' cells to be selected. Analysis of the matrix indicated that the majority of the barcodes contained a large number of UMIs and it was not possible to discriminate between 'true' cells and background. I also carried out clustering analysis testing several filtering parameters with the purpose of retaining only cells with high UMI counts (Fig. 3.16b). However, previously identified cell-type-specific marker genes were expressed across the majority of the barcodes suggesting that the RNA captured in each droplet belonged to different cells (Fig. 3.16c).

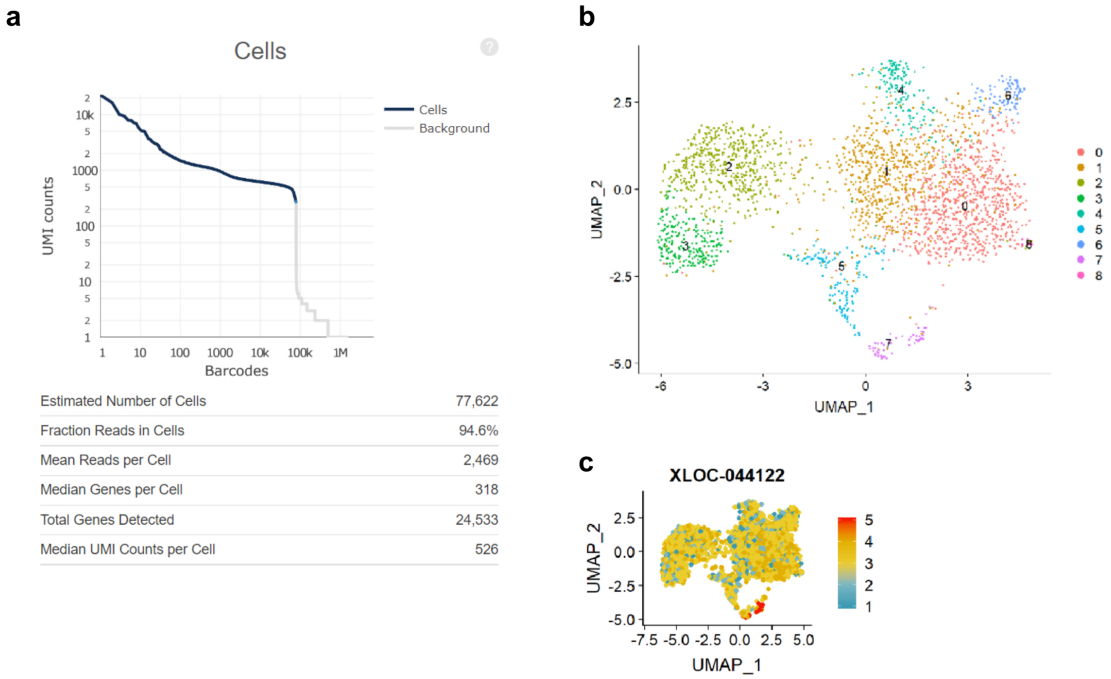


Figure 3.16 - Example of an unsuccessful planula scRNAseq experiment
(a) Detail of Cell Ranger 3.0 summary. (b) 2D-UMAP with stringent parameters consisting of 2672 cells. (c) Expression plot of the nematoblast-specific gene *Minicollagen 3/4*.

One hypothesis to explain the failure in these experiments was that planula cells are particularly sensitive to an osmotic type shock induced by mixing the cell suspension prepared in sea water with the reverse transcriptase (RT) reagents prepared in nuclease free water prior to the encapsulation. Accordingly, the lysis of the cells caused the release of RNA molecules in the solution resulting in the generation of droplets including RNA from various cells.

To overcome this issue I tested the InDrop (Indexed Droplets) system (Klein et al., 2015), a distinct droplet-based method, in collaboration with the Sebé-Pedros lab at the CRG in Barcelona (Fig. 3.17a). The InDrop microfluidic chip has four inlets, respectively for the hydrogel microspheres labelled with different barcodes, for the cells, for RT/lysis reagents and for oil, and an outlet for droplet collection (Klein et al., 2015, Fig. 3.17b). The three aqueous inlets are connected by a junction corresponding to the first contact point between hydrogel microspheres, RT/lysis reagents and cells.

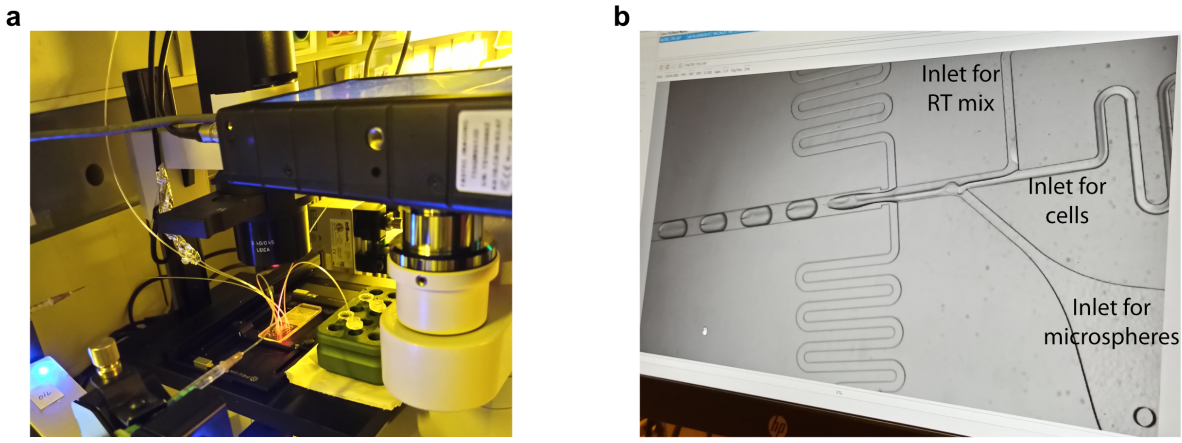


Figure 3.17 - InDrop system setup at CRG (Barcelona)
(a) Detail of the chip placement. (b) Snapshot of the encapsulation.

The advantage of this method is that the cells are kept in sea water until the encapsulation and the risk of osmotic shock followed by cell lysis should be reduced.
At the CRG we used InDrop to encapsulate freshly dissociated 2-day-old planula cells and prepared the gene expression library the next day with help from Sebé-Pedros lab members. Shortly afterwards we proceeded to sequencing. Unfortunately, the sequenced library showed an unsatisfactory complexity (i.e. few UMIs x Barcode).

Before visiting the Sebé-Pedros lab I optimised dissociation protocols involving a fixation step of planula single cells. I used two fixation methods, 80% methanol fixation, as in Chari et al. (2021) and ACME fixation, as in García-Castro et al. (2021).
I then prepared aliquoted single cell samples using the two fixation methods and I sent the samples to the Sebé-Pedros lab (see above).
I developed cell fixation protocols in case the experiment with freshly dissociated cells was unsuccessful and to explore how different methods would affect the encapsulation.
Also, I expected to accomplish a successful experiment with at least one of those methods.

With the help of Sebe-Pedros lab members, in particular Marta Iglesias García, we proceeded to process the fixed cells with both fixatives using 10X technology.
We included an additional sorting step (FACS) to eliminate any source of ambient RNA from the cell suspension and we encapsulated the sorted cells using the 10X Chromium Controller by aiming to encapsulate 4000 to 6000 cells for each experiment. The CRG genomics facility prepared the gene expression libraries according to the 10X Genomics recommendations and the libraries were sequenced on NextSeq 150 Midoutput and NextSeq 75 Highoutput.

After receiving the reads I mapped the data as described above and carried on with clustering analysis. Encapsulation experiments with FACsorted cells fixed with both fixatives were successful and we integrated all the produced libraries in our final atlas. Additionally, we were able to compare the effect of the two fixatives on planula cells after mapping (Table 3.4).

Sample	Fixative	Input read length	Uniquely mapped reads %	nUMIs	nCells	Mean UMI per Cell	Mean Gene per Cell
Planula 1	ACME	91	62.80%	6910892	2252	2192	621
Planula 2	ACME	91	60.37%	6703480	2149	2224	646
Planula 3	ACME	56	62.67%	4218581	686	2235	605
Planula 4	ACME	56	63.12%	5390430	1191	2211	588
Planula 5	ACME	56	60.04%	5293104	1349	2087	572
Planula 6	Methanol 80%	56	47.19%	9634406	3783	1388	627

Table 3.4 - Statistics of planula analysed libraries

We consider the quality of the ACME fixed cells better than methanol fixed cells, as confirmed during FACS experiments. The mapping rate against the genome was higher for ACME fixed cells ("Uniquely mapped reads", Table 3.4). The methanol-fixed cells yielded the highest number of total UMIs ("nUMIs", Table 3.4), althoug the values of UMIs per cell ("Mean UMI per Cell", Table 3.4) are higher for ACME-fixed samples, indicating that the integrity of the cells and the transcripts it's better preserved by ACME fixative.
Interestingly, we obtained higher recovery of cells from methanol samples. This is possibly because the cell membranes are less sticky when fixed in 80% methanol therefore less prone to form clumps and/or sticking to tubes or pipettes with respect to ACME fixed cells.

3.4.2 - Clustering analysis troubleshooting

Following clustering analysis of the integrated dataset including the 6 libraries we obtained an initial atlas consisting of 5347 cells and 22 clusters (Fig. 3.1). Clustering analysis, extraction of marker genes, selection of candidates and validation by in situ hybridization were performed as described above. We assigned 20 clusters as cell types.

Two of the clusters, clusters 0 and 3, were not assigned as cell types as they had a mixed transcriptome profile corresponding to a combination of the majority of genes specifically expressed in other cell types, especially in epidermis, neural cells and mucous cells (Fig. 3.2b). One hypothesis is that those clusters represent a cell state or alternatively a precursor of epidermal cells that eventually will differentiate in more specialised secretory cells (neural or mucous cells) in later planula developmental stages. On the other hand, I noticed the expression of other cell type specific markers, for instance CathepsinL, a marker for gastrodermis, suggesting that the mixed expression profile goes beyond epidermis and secretory cells.

I then computed co-expression of some cell type specific markers to address the co-expression of genes in the same cells. I used cell type specific markers with non-overlapping *in situ* expression patterns, also expressed in cluster 0 and 3 (Fig. 3.18c).
Strikingly, The UMAP plots of co-expressed genes (Fig. 3.18a and b) show that cell type specific marker genes are co-expressed in some barcodes belonging to cluster 0 and 3 indicating that those barcodes may very likely represent a technical artefact. We consider that they probably arise as an artefact of sample preparation, for instance due to ambient RNA associating with cell debris and subsequently being treated as cells.

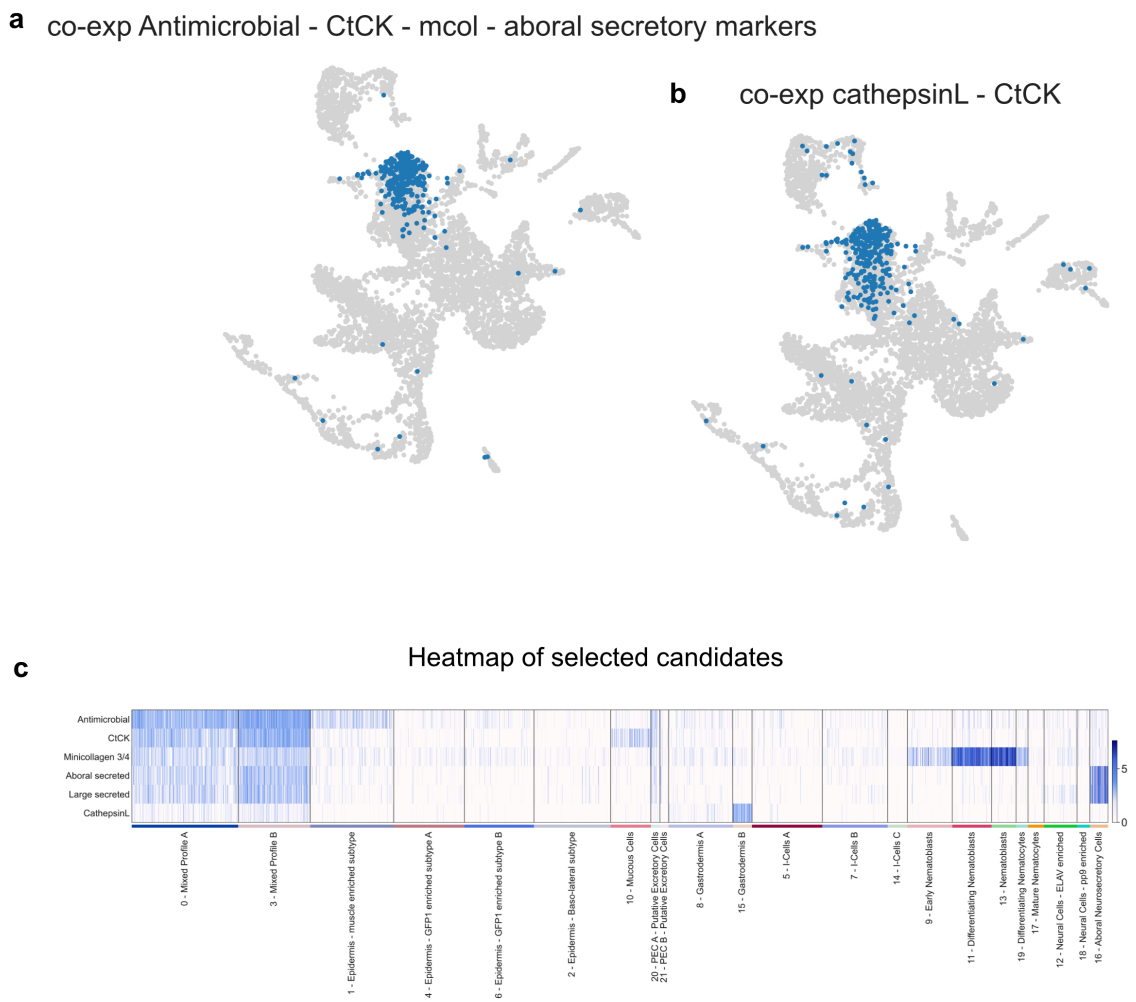


Figure 3.18 - Co-expression of marker genes
(a and b) 2D-UMAP plot of initial planula cell atlas. (a) The cells coloured in blue co-express the putative antimicrobial marker, CtCK, Minicollagen 3/4, and the Aboral Secreted marker. (b) The cells coloured in blue co-express CathepsinL and, CtCK markers. (c) Heatmap of candidate genes selected to address the co-expression .

The barcodes belonging to clusters 0 and 3 show a comparable number of counts to the other cells ranging from 500 counts to 3000 (Fig 3.19) so it is not possible or justified to filter them

out on the basis of the counts per cell. I manually removed barcodes belonging to clusters 0 and 3 and proceeded to recluster the remaining cells using the same parameters obtaining the atlas shown in Figure 3.3.

Figure 3.19 - Distribution of counts per cell across the clusters of the initial planula atlas
Distribution of counts in the initial planula atlas grouped by clusters/cell types. Numbers of the y axis correspond to clusters in Figure 3.2. Counts are shown on the x-axis and are log transformed.

The development of protocols to fix cells immediately after dissociation was crucial to obtain a satisfactory planula dataset. The unsuccessful encapsulation of freshly dissociated cells is very likely caused by augmented fragility of planula single cells compared to the medusa stage. The dataset obtained from the encapsulation of fixed planula cells still shows some features of previously unsuccessful experiments (notably cluster 0 and 3). However, the majority of the cells are of good quality and allowed the characterization of previously undescribed as well as providing further knowledge of known planula cell types.

3.5 - DISCUSSION

The *Clytia* planula cell atlas presented in this chapter represents the first cell type atlas of an hydrozoan larva. Four other cell atlas of cnidarian species have been published in the last few years, notably the hydrozoan polyp *Hydra* (Siebert et al., 2019), the anthozoan *Nematostella* (Sebé-Pedrós, Saudemont, et al., 2018), the soft coral *Xenia* (Hu et al., 2020) and the stony coral *Stylophora* (Levy et al., 2021) cell atlases. Only two of those include larva cell atlases: *Nematostella* and *Stylophora*.

Although the availability of cnidarian cell atlases has increased in the last few years, many questions concerning the evolution of cnidarian cell types remain open. Comparative analysis of hydrozoan and anthozoan cell atlases revealed a poor correlation between *Hydra* cell types and the others, concordant with the phylogenetic distance (Levy et al., 2021). Our planula cell atlas will allow comparisons and should help to gain insight on this issue. Furthermore, *Clytia* represents an advantage to address questions concerning the evolution of hydrozoan life cycle since the research on the model *Hydra* has focused on the adult polyp form.

I identified the hydrozoan planula cell types previously described in the literature, and gained additional understanding of their complexity along with uncharacterized types. I characterised transcriptional signatures for cells of the classical epitheliomuscular tissue layers, the epidermis and the gastrodermis, detecting additional epidermal subtypes potentially specific to the planula stage. Clustering analysis and validation via in situ hybridisation revealed the existence of cell types previously described on the basis of histology and morphology: the mucous cells, PECells and a population of aboral neurosecretory cells, originally termed glandular/granular cells (Bodo & Bouillon, 1968).

This *Clytia* planula cell atlas has proven useful for supporting existing knowledge of larval cell types as well as enhancing it with new information. One example concerns the nematocyte lineage. Strikingly, nematogenesis seems to employ similar sets of genes in both the medusa and the planula stages and is clearly a biphasic process. Furthermore, *in situ* hybridisation pattern analysis of nematogenic genes revealed that the first phase of nematogenesis (production of nematoblasts) is stalled for most cells at planula stage suggesting that these nematoblasts with intact nematocysts serve as a reserve to undergo final differentiation and deployment after metamorphosis at the polyp stage. The relatively small number of mature nematocytes in the planula may have the same fate, and/or have a defensive role already for the planula stage. The *Clytia* planula does not feed, therefore nematocytes are not needed to catch prey as they are in the adult stages. The presence of nematocyte capsules has been shown in the disk-like form into which the planula transforms a few hours after settlement (Krasovec et al., 2021), supporting this hypothesis.
Clustering analysis revealed two clusters of neural cells from the planula scRNAse data. Each of these clusters includes more than one neural subtypes, as shown by visualising the expression of several neuropeptide precursors previously described in medusa (Chari et al., 2021). We performed reclustering analysis of neural and neural-like cells, notably the aboral neurosecretory cells, to address the classification of planula neural subtypes. However, the

small number of cells used for reclustering yielded clusters of cells that could not be confidently assigned to neural subtypes. Analysis of the marker gene lists helped to identify few further neural cells-specific markers that very likely correspond to neural subpopulations.

The generation of the planula single cell dataset was more challenging compared to the medusa one. Despite several troubleshooting tests, I had to manually remove data associated with some UMI barcodes considered technical artefacts, which show similarities to the output obtained following unsuccessful experiments. Following discussion within the EvoCELL network, other PhD students encountered the same type of issue indicating that this type of output is common in such experiments starting with complex biological starting material such as whole animal larvae. For the final planula atlas, we excluded around 1000 cells reducing the number of cells to 4370. Given the relatively low number of cells, I consider that some additional cell diversity might be overlooked. Further single cell transcriptomics experiments are needed to increase the resolution of this atlas and possibly identify rare cell types. However, this planula cell atlas already makes a substantial contribution for expanding the knowledge on cnidarian cell types as well as representing a useful resource for the community. Furthermore, in the context of my PhD study, the planula atlas was crucial to allow the comparative analysis with the medusa atlas and to address the complexity of the medusa.

CHAPTER 4 -
Comparison of cell types across stages of *Clytia*

Overview of the chapter

Single cell transcriptomics allows the molecular characterisation of thousands of cells at the same time. When applied to whole organisms, this technique provides a large amount of information regarding gene regulation at the cell type level. In this chapter I present initial steps in a comparative study at the cell type level across life stages of the hydrozoan *Clytia hemisphaerica*, which represents the ultimate goal of my project. My main aim was to address the complexity of the medusa form by comparison with a simpler planula larva and detect differences or similarities at the cell type level within the hydrozoan life cycle.

Clytia displays a tri-phasic hydrozoan life cycle including a vegetatively propagating polyp colony, a free-swimming medusa form and a planula larva. Previous analysis of the genome and of bulk transcriptomes across the three life forms revealed distinct stage-specific gene expression programs indicating a potential diversification at the cell type level. The majority of stage specific transcription factors were detected within the medusa stage, suggesting a potential expansion of cell types in the adult of *Clytia*. For my comparison of cell types across life stages, I focused on the adult medusa and the planula larva of *Clytia* and generated single cell data. Besides being a resource for understanding the basis of the transcriptional regulation of the cell types, these data serve as a foundation to address similarities at the cell type level across the life cycle of *Clytia*.

Together with Richard Copley I generated a presence/absence matrix of expression of marker genes across the clusters using a method developed previously (Leclère et al., 2019). We then applied a phylogenetic statistical approach usually employed to predict gene “gain” and “loss”, in this case to generate a binary cell type tree which represents hierarchical relationships between cell clusters/types. We applied this workflow using presence/absence matrices computed with different combinations of the available atlases to test its robustness. We observed some degree of consistency between the results of different analyses at the level of cell classes, but relationships across cell types were not always coherent. Thus, I restricted my interpretations to relationships at a broad level, such as between cell classes that are consistent across the analyses.
These analyses allowed me to propose shared and unique cell classes across life stages. Cell classes potentially unique to stages, include the planula aboral neurosecretory cells and mucous cells, and the medusa digestive gland cells. The planula and medusa epidermis cell types do not show strong similarities in these analyses, likely reflecting the higher degree of specialisation of the medusa epidermal cell types. The analyses revealed consistent similarity of the gastrodermis class between planula and medusa, as well as of the neural cell class, with diversification into specialised subtypes detectable only in the medusa in both cases.
These analyses suggestd similarity at the level of cell types within two classes, one of these being the nematocytes. Cell types included in the nematocyte cell class, which represent successive stages of nematogenesis, show a high degree of similarity across life stages. Similarly, cell types belonging to the i-cell classes showed consistent relationships between stages across different analyses. However, the biological interpretation of the state of differentiation of planula i-cell types is not yet clear which makes the interpretation of the relationships difficult. Interestingly, some of the analyses suggested similarity between neural

cells and mature nematocytes, suggesting common elements in the transcriptional programs of these cell types as suggested in *Nematostella.*

As a parallel approach to compare cell types between medusa and planula I extracted a list of transcription factors for each cell type as a first step in identifying common gene regulatory programmes. As one concrete example I investigated which transcription factor families contributed to the similarity between neural cells and mature nematocytes allowing us to propose a tentative evolutionary scenario.

This comparative study is still in its early stage and it requires further work to extend the comparison to finer levels. Considering the phylogenetic position of *Clytia* as part of Cnidaria*,* this work will contribute to wider comparative studies concerning cell type evolution across metazoans.

4.1 - INTRODUCTION

Up until very recently, cell types have been classified on the basis of morphological observation and functional considerations. Single cell transcriptomics is a relatively recent innovation in rapid advancement. Currently it allows the characterisation of the transcriptome of thousands of cells at the same time representing a powerful tool to address the characterisation of the cell types at the molecular level (Hwang et al., 2018).

Differential gene expression profiles can be investigated to determine the physiological and functional features of given cell types while the regulation of those signatures is illustrated by a combination of transcription factors. Taken together, these co-expression modules constitute the blueprint of cell types (Arendt, 2008). Moving forward, cell type specific programs allow a cell type to evolve in a partial independent manner (Arendt et al., 2016). Thus, in an evolutionary framework a cell type can be considered as an evolutionary unit opening the way to comparative studies across animal phyla addressing the origin and evolution of cell types (Arendt, 2008; Arendt et al., 2016).

With these concepts in mind, I aimed to address similarities and differences among cell types in the context of a typical hydrozoan life cycle. Concretely, I exploited the cell atlases of the planula larva and the adult medusa of *Clytia* presented in the previous chapters to address the complexity of the medusa form against a simpler planula larva.

Unlike the well studied polyp-only species *Hydra*, *Clytia* displays a complete hydrozoan life cycle (Houliston et al., 2010). The three main life stages of *Clytia*, exhibit radically distinct body forms (see Chapter 1). Analysis of the genome and the bulk transcriptome across the three life stages revealed specific gene expression programs for each stage (Leclère et al., 2019). Accordingly, expression of a diversified set of transcription factors was detected across life stages (Leclère et al., 2019). The majority of stage-specific transcription factors were specific to the medusa stage, reflecting the complexity of the medusa with respect to the other life forms. Moreover, transcription factors expression patterns were often detected as scattered cells in specific regions of the medusa notably the manubrium, gonads, tentacle bulb and nerve rings suggesting a potential association with specific cell types (Leclère et al., 2019; see Chapter 1). The higher number of transcription factors expressed at the medusa stage can be explained by an evolutionary expansion of cell types to match the increased anatomic complexity. However, whether this implies a gain of certain cell types signatures which are completely absent at planula stage is an open question. A smaller number of transcription factors was detected as specific to the planula stage, suggesting independent evolution of some planula cell types.

The cell atlases introduced in the previous chapters set the stage for addressing these questions. In this chapter I present our comparative study at the cell type level across the planula larva and the adult medusa of *Clytia.* These analyses suggest the existence of unique cell type classes for both stages. On the other hand, several cell type classes displayed similarities at the level of transcriptional signature across life stages within each of these

common cell type classes, the medusa shows a richer diversity of cell types than the planula. Finally, these analyses support the existence of interesting similarities between neural cells and mature nematocytes which are shared across life stages.

4.2 - RESULTS

4.2.1 - Comparative analysis of cell types across life stages

To perform the comparative analyses we exploited the information included in our planula and medusa cell atlases. As described in the previous chapters the cells were computationally separated into 40 clusters for the medusa and 19 clusters for the planula, which we refer to as cell types. The cell types were defined by expression of a set of marker genes that show statistically significant variable expression between them (see Chapters 2 and 3). We used these marker genes as an appropriate basis to conduct comparative analysis across clusters in an unbiased way given that our atlases exhibit large differences in the number of cells and cell types. We exploited those sets of genes to generate a presence/absence binary matrix, combining information from both stages. We analysed the matrix using *biphy*, a method to infer phylogenies from binary character presence/absence data by modelling character gain and loss as a Poisson process (Pett et al., 2019; https://github.com/willpett/biphy; see below). We merged the results of 10 independent outputs and obtained a consensus tree depicting a hierarchy of shared marker gene content between clusters across planula and medusa stages. The resolution of the consensus tree can certainly be improved. However, based on well characterised cell types, some aspects make sense biologically and give us confidence in conclusions for other cell types. For instance the presence of strongly expressed markers such as *Minicollagen 3/4* likely contributes to the similarity detected between planula and medusa nematoblasts.

To test the robustness of our approach of deducing hierarchies of shared marker gene content, which has never been used before to perform comparative analysis of cell types, we applied the same workflow to several presence/absence binary matrices obtained from combinations of the available medusa and planula atlases. We generated presence/absence matrices using lists of genes extracted from the published medusa atlas (Chari et al., 2021) as well as my integrated medusa atlas, and combined these with lists of genes selected from our initial planula atlas as well as the recomputed one (which does not include cells belonging to the likely-artifactual clusters 0 and 3; see chapter 3). Finally we generated binary cell type trees out of each combination (Fig. 4.1; Annex 4a; Annex 4b).
We noticed that the topology of the trees varied depending on the combination of atlases used and was not consistent across analyses. What differed most was the arrangements at the level of the cell types. Therefore, our approach is not robust enough to assess clear relationships across cell types. Exceptions were the nematocyte and i-cell types of both planula and medusa which consistently showed similarity in the different analyses. Similarities across nematocyte and i-cell types are supported by other considerations, for instance, the genes used as markers to validate cell type identities are the same across stages. This suggests that our approach is promising, even though it certainly requires further work to understand the sources of variability. Accordingly, we observed a certain degree of consistency across the analyses at the level of cell classes. For instance neural cell types showed consistent similarity relationships independent of which atlas the data were drawn from, as did gastrodermis cell

types. Thus, for the moment I restrict my interpretations mainly to the level of the cell classes that showed consistencies across the analyses and that are supported by other considerations.

One example of a hierarchical cell type tree is shown in Figure 4.1. This particular tree is obtained by the analysis of the presence/absence matrix generated by using marker genes from the initial planula atlas (including cluster 0 and 3) and the medusa atlas published in Chari et al., (2021). The marker gene list of the published medusa atlas includes genes computed from a merged “Fed and Starved” and “Stimulation” datasets, which we used to select candidates for *in situ* hybridisation (Chari et al., 2021). So far, this is the only combination that generates a tree with a sensible biological interpretation. We do not yet understand the sources of variability in the trees we generated although one contributing factor is likely the batch effect. The published medusa atlas is not affected by batch effect, which may represent a source of alteration in the trees generated by using my integrated medusa atlas (see Chapter 2).
In Figure 4.1, each branch of the tree represents a cell type identified by the numbers assigned during clustering analysis. The ones coloured in blue belong to the medusa while the ones in red belong to the planula atlases. The dashed boxes highlight cell classes which show a consistent similarity across our analyses such as gastrodermis, neural cells, nematocytes and i-cells. I found that some cell classes consistently do not mix on the branches, such as aboral neurosecretory, mucous and putative excretory (PEC) cells of the planula as well as digestive gland cells of the medusa. We consider those as potentially unique to stages. However we can not rule out any other relationships between those and other classes, for instance whether they are more closely related to some classes than others, given their variable arrangement on the trees across the analyses.

In the following sections I discuss putative shared and unique cell classes across stages. I include some observations on putative regulatory programs that I could draw by analysing the list of transcription factors I extracted from a presence/absence matrix (Annex Table C; see methods). I only considered transcription factors that are uniquely detected in medusa or planula cell types or classes. I paid particular attention to transcription factors that are shared in nematocyte and neural cells and are potentially contributing to the similarity we observed in some but not all the analyses and the results of recent studies on *Nematostella.* I did not consider transcription factors that are shared across several cell classes and stages considering those as constitutively expressed and very likely not involved in regulatory programs that define cell types.

Figure 4.1 - Hierarchical cell type tree of medusa and planula cell types
Hierarchical representation of similarities between cell types and classes across the planula and the medusa of *Clytia.* Branches are labelled with the numbers assigned during clustering: Medusa cluster/cell types in blue; Planula clusters/cell types in red. Dashed boxes represent similarities that were consistently retrieved by analysis of different combinations of the available atlases (see text). Note that the 'i-cell and oocyte' box also includes progenitors of neural cells and nematoblasts.

4.2.2 - Cell types with unique transcriptional programs across stages

Analysis of the hierarchical cell type trees identified certain cell types with a low probability of sharing their expressed gene sets with any cell type from the other stage. I consider that these cell types are potentially unique to stages. Based on overall available evidence, two cell types can be considered planula-specific: the mucous cells and the aboral secretory cells, while the PEC cells are planula/stolon specific. For the medusa, the digestive gland cells, and also the bioluminescence cells both appear as specific cell types, while many sub-specialisations of the epidermis are also specific to this complex stage.

Concerning the planula, we found that the transcriptional signatures of the mucous cells and the aboral neurosecretory cells group as distinct from any cell type in the medusa in some of our analyses (annex 4b). In other analyses the mucous cells grouped with the planula ectoderm types and the aboral neurosecretory cells with the neural cells (Fig. 4.1; Annex 4a). The marker genes used to assign identities to both the mucous cells and the aboral neurosecretory cells are strongly and specifically expressed at planula stage (see Chapter 3), which argues in favour of the specificity of these two cell types to this stage. Very likely, these cell types are involved in specific planula functions such as settlement and metamorphosis. The mucous cells secrete a sticky glycoprotein rich substance which may help to associate the swimming planula with the substrate and the aboral secretory cells a glue-like substance that adheres the planula aboral pole tightly to it once settlement is triggered (Bodo & Bouillon, 1968).

Analysis of transcription factors in the presence-absence matrix revealed a planula-unique transcription factor *TBX8* (Lapébie et al., 2014; Leclère et al., 2019) which is expressed in both the mucous cells and aboral secretory cells. We can speculate that this shared expression reflects transcriptional regulation by TBX8 of a particular set of genes related to synthesis of planula-specific secreted glycoproteins. Alternatively it may relate to the common developmental origin of these cells, along with some neural cell types from the gastrula ectoderm (Ruggiero, Ferraioli, Chevalier et al. manuscript in preparation; see below). Functional studies would allow the role of TBX8 in the aboral neurosecretory and mucous cells in the planula of *Clytia* to be tested.

The putative excretory cells (PECs) of the planula showed a lack of similarity with any of the medusa cell types across the analyses (Fig. 4.1). Bulk transcriptome data reveal that marker genes for PECs are expressed strongly in planula and in some cases shared with the polyp stage, and more specifically in the stolon, but not in the medusa (see Chapter 3). Thus, it is plausible to consider PECs unique to planula stage compared with medusa cell types. Further studies are needed to determine the function of PECs in planula as well as in the polyp colony. They may be involved in excretion of digestion waste, as suggested by Bodo and Bouillon (1968). Among the set of transcription factors unique to planula stage we detected an uncharacterised Ets-domain containing transcription factor specifically expressed in PEC-A cells and in the subtype of the planula epidermis enriched in muscle genes, which could for instance reflect regulation of these cells a common signalling pathway (Röttinger et al., 2008). Interestingly, we detected the expression of a member of the *bHLH* family transcription factor

HLH6 in PECs along with mucous cells as well as in neural precursors and neural subtype expressing pp7 of the medusa (Annex Table C). *bHLH* transcription factors, including members of the Atonal superfamily to which Hlh6 belongs, such as its *Nematostella* orthologue Ath-like, are commonly associated with neurogenesis (Rentzsch et al., 2017). Analysis of the medusa cell type atlas shows that Hlh6 is expressed on the trajectory from i-cells to neural cells (Chari et al., 2021). Hypotheses concerning similarities and differences of neurogenesis in the planula and the medusa of *Clytia* are discussed in the following section.

Several medusa cell classes were consistently distinct from the planula classes in the various *biphy* analyses. A first example is represented by the medusa digestive gland cells (Fig. 4.1). These gland cells show transcriptional profiles dominated by strongly expressed digestive enzymes (see Chapter 2) and are likely to be involved in primary digestion processes. The planula does not feed, which fits with their absence at the larval stage. Interestingly in this particular analysis the group of digestive gland cells shows similarity with the neural cell group (including planula neurosecretory and neural cell types; Fig. 4.1). This may reflect the developmental relationship between gland cells and neural cells proposed from scRNAseq studies in *Hydra* and *Nematostella* (Siebert et al., 2019; Steger et al., 2022).
A second medusa cell type with a distinct transcriptional program from the cell types of the planula are the bioluminescent cells (Fig. 4.1). Bioluminescent cells of the medusa are characterised by the expression of GFP2 and Clytin2 (Fourrage et al., 2014; see Chapter 2) which are not expressed at planula stage. These bioluminescent cells are located along the tentacles and have been proposed to play a role in attracting prey. This medusa specific function supports their specificity to the medusa stage when compared with the planula.

Cell types belonging to the epidermis cell class for the planula and the medusa showed no consistent similarity relationships, i.e. a low probability of sharing transcriptional programs (Fig. 4.1). This could be explained by the high degree of specialisation of medusa epidermis in distinct subtypes such as smooth and striated muscles, exumbrella, tentacle bulb, gonad and manubrium epidermis which potentially are maintained by a distinct transcriptional programmes (see Chapter 2). The epidermis subtype enriched in muscle genes assigned to the planula atlas is presumably distinct in structure to any of the epitheliomuscular cell types of the medusa. The other two epidermis cell types of the planula are also apparently stage specific. The differences between these planula epidermis subtypes are not yet understood (see Chapter 3). Some or all of them contribute to different regions of the epidermis of the primary polyp upon metamorphosis, a specialised thin cell layer specialised to produce the extracellular theca of the colony.

4.2.3 - Cell types with shared transcriptional programs across stages

Analysis of the various hierarchical cell type trees consistently revealed close relationships for a number of cell classes across planula and medusa. This suggests with a high degree of probability that similar gene modules are shared between certain cell classes. We consider these as shared cell classes across life stages. For instance, we detected similarities between

groups of medusa and planula cell types within gastrodermis, I-cell , neural cell and nematocyte classes, as indicated by the dashed boxes in Figure 4.1. These relationships were consistent across the analyses independently of the combination of datasets used. In this section I discuss similarities concerning the gastrodermis and the i-cells classes. Similarities within neural cell and nematocyte classes are discussed in the following two sections.

Shared transcriptional signatures are detected among cell types of the gastrodermis class across the two stages as indicated in Figure 4.1. This is presumably supported by the expression of the enzyme Cathepsin-L, commonly expressed in gastrodermis cell types in medusa and in planula (see Chapter 2 and Chapter 3). A common transcriptional regulation is potentially driven by the transcription factor FoxC (Lapébie et al., 2014) which is detected in our presence/absence matrix as expressed in tentacle bulb gastrodermis and in the gastrodermis cell types of the planula (Annex Table C; see *in situ* pattern in Chapter 3). However, we observed a clear diversification in several gastrodermis subtypes in the medusa but not in the planula. Differences between the medusa gastrodermis subtypes likely relate to distribution of digestive functions between the manubrium, gonads and tentacle bulbs, as well as to their different contributions to other medusa-specific gastrodermis behaviours such as mesoglea deposition, autophagocytosis or starvation-induced cell mobilisation (Chari et al., 2021; Sinigaglia et al., 2020). The precise identity and function of each one remains to be established. Among the transcription factors unique to the medusa we detected the expression of a bZIP:popeye domain containing transcription factors specific to gastrodermis subtypes of the medusa (Annex Table C). Previous studies (Leclère et al., 2019) identified this protein as an ortholog of BVES, a transcription factor expressed in epithelial tissues including the intestine and involved in the regulation of programs associated with epithelial-to-mesenchymal-transition in mice (Choksi et al., 2018). Whether this transcription factor plays a role in the differentiation or function of particular gastrodermis cell subtypes in the medusa remains to be determined.

As expected, consistent transcriptome similarity across analyses was observed among planula and medusa i-cells (Fig 4.1; Annex Table C), which is supported by the expression of diagnostic markers such as *piwi, vasa* and *nanos1* (Leclère et al., 2012; see Chapter 2 and Chapter 3). Early neural and nematocyte precursors, which derive from i-cells and also express these genes, are included in the group (Fig. 4.1). Oocyte transcriptome signatures also showed strong similarity to those of the i-cells. The highest number of transcription factors unique to medusa in our matrix specific to a particular cell type are expressed in medium oocytes (Annex Table C). This can be presumably explained by the high content in oocytes of maternal transcripts with functions in early development (Annex Table C).

4.2.4 - Insight into developmental origin of neural cells

Precursors of medusa nematoblasts and neural cells along with precursors of planula nematoblasts showed strong evidence for shared transcriptional profiles, as well as some

degree of similarity with the i-cells across analyses (Fig. 4.1). This is not surprising considering that nematoblasts are clearly derived from the i-cell precursor population at both stages, as are neural precursors in the medusa (Chari et al., 2021). On the other hand, we have not clearly identified neural cell precursors or observed a clear relationship between i-cells and neural cells in our planula atlas (see Chapter 3). We propose that this reflects the presence of a parallel neurogenesis pathway in the planula, generating some neural cells as well as mucous and secretory cells from the gastrula ectoderm rather than from i-cells (Kraus et al., 2020; Thomas et al., 1987; Ruggiero, Ferraioli, Chevalier et al. in preparation). Further evidence for this hypothesis comes from analysis of the shared repertoire of transcription factors, which highlighted the expression of the SoxB family transcription factor *Sox10* in epidermis subtypes in planula and in neural cell precursors and subtypes of medusa. Ongoing experimental investigations within our laboratory involving lineage tracing and embryo bisection are confirming that some neural cell subtypes and mucous cells derive from the aboral ectoderm rather than from i-cells. The expression of *Sox10* in our planula dataset suggests that this process of neurogenesis from the ectoderm is still active at 48hpf of development.

At the medusa stage, some neural subtypes of the planula are clearly developing from i-cell like precursors. Among the set of transcription factors shared across stages we detected the expression of the bHLH transcription factor Neurogenin in the i-cells of the planula and in the neural cell precursors of the medusa. Neurogenin is a good candidate as a regulator ofthe i-cell-mediated neurogenesis pathway and our data indicate that this role is shared across life stages. Reclustering analysis of neural cells of the medusa revealed the expression of Neurogenin in a neural subpopulation classified as neural cell precursors (Chari et al., 2021). Generation of scRNA-seq data from earlier planula stages (i.e. 12hpf, 24hpf) would facilitate the identification of clear neural precursors and help in dissecting the developmental origin of the planula neural cells.

Concerning those neural subtypes that we consider differentiated neural cells, analysis of the cell-type trees revealed consistent similarities between the two life stages (Fig. 4.1). The shared transcriptional signatures are dominated by neuropeptide precursors which are commonly expressed in planula and medusa neural cells such as Pp5, Pp9 and Pp11 (Chari et al., 2021*;* see Chapter 2 and Chapter 3)*.* Furthemore, markers such as ELAV genes and Calmodulin*,* are expressed broadly in neural cell types in both planula and medusa (Chari et al., 2021). Similar to the case of the gastrodermis, in the medusa atlas we assigned several neural subtypes based on the expression of neuropeptide precursors which are not all clearly detected in planula; examples are Pp17, Pp27, Pp25. As explained in Chapter 3, the number of neural cells in the current planula dataset is limiting and likely causing a lack of resolution in the determination and comparison of explicit neural subtypes.

4.2.5 - Origin and development of nematocytes in *Clytia* and similarities with neural cell types across life stages

The nematocyte developmental trajectory exhibits a high degree of similarity across life stages. This is supported by the expression of a common set of diagnostic markers that recapitulate the two phases of nematogenesis in both life forms (see Chapter 2 and Chapter 3).

In medusa, nematoblasts originate from i-cell-like precursor pools located in the proximal region of the tentacle bulb epidermis (Denker et al., 2008). Consistently, we detected at this site the expression of the *Z*nf845 transcription factor, whose orthologues are commonly expressed in i-cells and early nematoblasts (Fig. 2.10c). The generation of the nematocyte capsule (nematocyst) occurs in slightly more distal regions of the tentacle bulb epidermis as indicated by the expression of specific markers such as *Minicollagen 3/4* (Fig. 2.10c). Once the nematocyst is complete, the cells move into the distal region of the tentacle bulb epidermis, towards the tentacle. At this stage this first transcriptional program is switched off and a process of nematocyte migration and maturation takes place. Maturing nematocytes expressing *M14-peptidase* migrate towards the end of the tentacle and the program including *SANS/USH-1G, Whirlin, Harmonin* is turned on. The migration may be driven by the active and continuous proliferation of the i-cells and early stage nematoblasts at the tentacle base (Denker et al., 2008). Finally mature nematocytes become oriented perpendicularly to the direction of migration and inserted into the tentacle epidermis. The expression of *Nematocilin* marks the production of the nematocil, the trigger for nematocyte discharge (see Chapter 2; Chari et al., 2021).

Nematogenesis in the planula occurs according to essentially the same programs (see Chapter 3). Indeed we can clearly match stages of differentiation across stages. Expression of the same sets of markers distinguish the two phases as corroborated by our hierarchical cell type tree analyses (Fig. 4.1) We found an unexpected developmental “stalling” of the nematogenesis programme in the planula. The first phase initiates in many cells as early as gastrulation, but only a small proportion of the Minicollagen3/4 expressing cells proceeds to the second phase. Thus, while differentiation of a small population of mature nematocytes is completed early during planula development, most cells remain at the ‘late nematoblast’ stage (see Chapter 3). In fact, early nematoblasts continue to proliferate until the end of the planula lifespan but the differentiation program is stalled before entry into the nematocyte maturation phase (see Chapter 3). Analysis of the transcription factors detected in the presence/absence matrix revealed the expression of several members of the Fox, bHLH, Tbox and DMRT families in nematoblasts and nematocytes uniquely at planula stage (Annex Table C). This suggests a particular regulation of the nematoblasts-nematocytes transition at the planula stage. Functional analyses will determine the role of those transcription factors in this regulation.

Studies in *Nematostella* have demonstrated the origin of nematocytes and neural cells from common neural-like-stem-cell precursors expressing the transcription factor SoxB2 (Richards & Rentzsch, 2014). Co-expression of Znf845 in these cells leads to the diversification of the

progenitor into nematocytes (Babonis et al., 2022). *Nematostella* Znf845 is also partially co-expressed with PaxA, which has been described as a regulator of the differentiation of nematocytes (Babonis & Martindale, 2017). However, recent studies indicate that Znf845 acts upstream PaxA, determining the switch of the progenitor cell towards the nematocyte lineage (Babonis et al., 2022). Consistent with these findings in *Nematostella*, our pseudotime analysis of i-cells and nematocytes lineages in *Clytia* have shown that *Znf845* is co-expressed in a population of progenitor stem cells which give rise to nematocytes (Chari et al., 2021). Analysis of the transcription factor list further revealed an ortholog of *PaxA* in both planula and medusa early and differentiating nematoblasts but not in I-cells (Annex Table C), supporting a conserved evolution of the Znf845/PaxA pathway in *Clytia*.

Interestingly, a particular combination of atlases used to generate the hierarchical cell-type trees, specifically the integrated medusa cell atlas and initial planula cell atlas (Annex 4a), suggested similarity between mature nematocytes and neural cells in both planula and medusa. Examination of the transcription factor list obtained from the presence absence matrix revealed that the expression of the homeobox *Pou4* (also known as Brn3) is specifically expressed in mature nematocytes and neural cells at both stages (Annex Table C). *Pou4* genes are required for the differentiation of several neuronal types, including sensory neurons in *C.elegans,* mice and *Drosophila*, (Leyva-Díaz et al., 2020; Serrano-Saiz et al., 2018). Functional studies have demonstrated that the downregulation of orthologs of *Pou4* results in the loss of the neurotransmitter identity in *C.elegans* and mice (Serrano-Saiz et al., 2018). This suggests a conserved regulatory function of this family in the differentiation and maintenance of the neural identity in Bilateria (Leyva-Díaz et al., 2020; Serrano-Saiz et al., 2018). Additionally, recent studies in *Nematostella* have demonstrated that *Pou4* is expressed in neural cells and it also regulates the generation of the nematocil and the maturation of mechanosensory “hair cells” (Ozment et al., 2021; Tournière et al., 2020). The detection of *Pou4* in neural subtypes and mature nematocytes across life stages of *Clytia*, contributes to the following hypothesis concerning nematocyte evolution. The role of *Pou4* genes in the determination of neural identity, in particular in distinct neural subtypes and sensory structures predates the Cnidaria-Bilateria divergence (Ozment et al., 2021; Tournière et al., 2020).

Taken together, these findings allow considerations of the origin and the evolution of the nematocytes. Possible evolutionary scenarios are discussed further in the following section. .

4.3 - METHODS

4.3.1 - Generation of the presence/absence matrix

Marker genes were computed as previously described for the medusa (Chari et al., 2021) and by using the Seurat function “FindAllMarkers” for the planula. The dataset used for the medusa is represented by a merged dataset including the “Fed and Starved” and the “Stimulation” data, which were integrated following a custom pipeline to eliminate batch effect (Chari et al., 2021;
https://github.com/pachterlab/CWGFLHGCCHAP_2021/tree/master/notebooks/L1Validation.
The dataset used for the planula is represented by the initial planula atlas including cluster 0 and 3 (Fig. 3.3 and Fig. 3.4).
Genes were computed with ranking tests (Wilcoxon non parametric test). Only genes associated with statistically significant expression value ($P < 0.05$ for the medusa and $P < 0.01$ for the planula genes), which we consider marker genes, were retained and used to generate a binary planula-medusa presence/absence matrix of unique genes expressed with custom Python scripts.

4.3.2 - Generation of the hierarchical cell-type tree

To infer cluster cell types based on shared marker gene expression, we applied a phylogenetic statistical approach to model gene gain and loss by performing Bayesian phylogenetic analysis, correcting for the inability to observe genes that were absent in both datasets (see manual) using the Biphy program (https://github.com/willpett/biphy; Pett et al., 2019). We performed a total of 10 iterations and merged the results in consensus trees using the Phylobayes package (Lartillot, 2020b, 2020a). The resulting consensus trees constitute what we consider a hierarchical cell-type tree, which allows us to visually inspect similarities and differences across clusters.
We performed these analyses on presence/absence matrices generated by combining lists of marker genes from different available planula and medusa cell atlases and measured consistency across the trees to test the robustness of the approach. The topologies of the generated trees were not consistent at the fine level of the cell type relationships. However we detected consistencies at the broader level of the cell classes. How the use of different combinations of the datasets affects the topology of the tree is still unclear.

4.3.3 - Generation of the transcription factor list

Transcription factors were isolated from the presence/absence matrix computed by using marker gene lists extracted by the recomputed planula atlas (Fig 3.5) and the integrated medusa atlas (Chapter 2). Genes were computed as described in the methods section of Chapter 2 and Chapter 3. A common filtering threshold was applied and only genes associated with a P value below 0.05 were retained for the following analysis (see above). A list of 550 transcription factors (Leclère et al., 2019) was used to extract positive matches from the presence/absence matrix with custom Python scripts. The final list consists of 239 transcription factors annotated by family (Leclère et al., 2019) and, where available, by name or clear ortholog names. Annotation by cell types and classes in which the genes were detected is also provided. Among those, 128 transcription factors were detected in cell types belonging to both stages, 85 were specific to cell types of the medusa and 28 were specific to cell types of the planula. The complete table is available as Annex Table C.

4.4 - DISCUSSION

The advent of single cell transcriptomics has opened the way to the characterisation of cell types in relatively complex groups of animals. In a much broader framework the information of molecular signatures of cell types represents an important tool to establish comparative studies across phyla and trace evolution (Arendt, 2008; Arendt et al., 2016).

We exploited our planula and medusa cell atlases to explore shared transcriptional signatures at cell type level across adult and larval stages of *Clytia.* The medusa is considered the most complex body form across Cnidaria (Leclère et al., 2019). Previous studies have identified a larger transcription factor repertoire associated with the medusa of *Clytia* with respect to other stages (Leclère et al., 2019). The molecular characterisation of the cell types presented in this study allows us to determine whether this complexity has a direct relationship with an increased diversity of cell types at this stage.

Our comparative study provides an initial overview of similarities and differences across life stages. Exploiting lists of marker genes we generated a presence/absence matrix which we analysed with a phylogenetic approach to explore similarities and differences across life stages (Fig. 4.1).

We detected the existence of putative stage specific cell types. The larva of *Clytia* has a lifespan of around 72 hours in laboratory conditions (Lechable et al., 2020; see Chapter 1). During this time, embryogenesis, cell differentiation and metamorphosis occur (Houliston et al., 2010; Kraus et al., 2020; see Chapter 1) and the larva likely undergoes major gene expression shifts. Our planula atlas corresponds to the 48 hpf stage allowing us to characterise the molecular signatures of the cell types present at this time of development. When compared to the medusa cell classes, the transcriptional signatures of the mucous and the aboral neurosecretory cell of the planula show low probabilities of being shared (Fig. 4.1). Among the transcription factors expressed uniquely at planula stage we detected *TBX8* transcription factor (Lapébie et al., 2014; Leclère et al., 2019) commonly expressed in these two cell types. TBX8 is thus potentially involved in the regulation of cell type-specific genes. Phylogenetic studies have demonstrated that the evolution of the T-box family transcription factors predates the origin of Metazoa. Although the DNA binding domain is highly conserved across Metazoa and non Metazoa, target genes differ, as demonstrated by sequence-based analysis of Brachyury. This suggests that an expansion of co-factors evolved with the outbreak of Metazoa (Sebé-Pedrós et al., 2013) and very likely corresponds with the emergence of novel cell types. In this scenario, T-box family transcription factors likely contribute to the differentiation of particular cell types, which would be represented by the aboral neurosecretory and mucous cells in the planula of *Clytia.* Furthermore, given the specificity to the stage of marker genes for these cell types (see Chapter 3), very likely they are both involved in functions specific to planula stage, such as settlement or metamorphosis.
Another cell class which shows low probability of similarity with any cell class from the adult medusa stage based on our marker gene comparisons is the putative excretory cells (PEC; Fig. 4.1). In this case the stage specificity likely extends to the polyp colony as well as the

planula stage, given the expression of some markers in the stolon bulk RNA seq (see Chapter 3). Concerning the medusa, a putative stage-specific transcriptome signature is detected when comparing digestive gland cells against planula cell classes, although there are hints of a relationship with the neural cell class (Fig. 4.1). The unique expression profile of the digestive gland cells is explained by the role of these cells in primary digestion, implying active feeding, which occurs in the medusa medusa but not in the planula. Feeding behaviour is observed at the polyp stage as well. However, given the current lack of a polyp (gastrozooid) cell atlas we can not rule out whether gland cells are present within the polyp stage.
A clear medusa-only cell class with many unique marker genes is the oocytes at different stages of oogenesis. These are clearly related to all the i-cell types from both stages and share a large set of genes with those (Fig. 4.1). We did not yet generate a cell type atlas from male medusae, but would expect spermatogenesis stages also to be medusa-specific. The complex behaviour exhibited by the medusa stage is partially due to unique-to-stage muscle fibres, in particular the fast contracting striated muscles, which allow the pulsating swimming motion (Leclère & Röttinger, 2017). In our analyses, we did not observe clear grouping epidermis cell classes across stages (Fig. 4.1). This was somewhat surprising but might partially be due to very specialised transcriptional signatures of some medusa cell types, notably striated muscles (see Chapter 2 and Fig. 4.1), but also by the sets of genes which constitute the uncharacterised planula-specific subtypes (see Chapter 3 and Fig 4.1).

Clear cell classe-level similarity between stages is detected for both the gastrodermis and the neural cells classes (Fig. 4.1). In each case, despite the broader level of similarity, the medusa displays an expanded set of cell subtypes when compared to the planula. Further investigations are needed to determine whether this expansion is dependent on the evolution of stage specific regulatory programs in the adult. A possible scenario is discussed in the General Discussion.

Analysis of the presence/absence matrix revealed that transcription factors associated with neurogenesis are shared across life stages (Leclère et al., 2019), although in some cases, the expression is not associated with the same cell types or classes. Members of the *SOX* and *bHLH* families, notably Sox10 and Neurogenin, are detected in I-cells and neural cell precursors in medusa as previously described (Chari et al., 2021; Leclère et al., 2019) supporting the hypothesis that neural cells originate from a stem cell like progenitor. *Neurogenin* expression in the i-cell of the planula is consistent to that observed at medusa stage. However, *Sox10* is detected in epidermis but not in i-cells in the larva, consistent with the likely parallel developmental origin of certain neural cells at this stage (see above; Annex Table C).

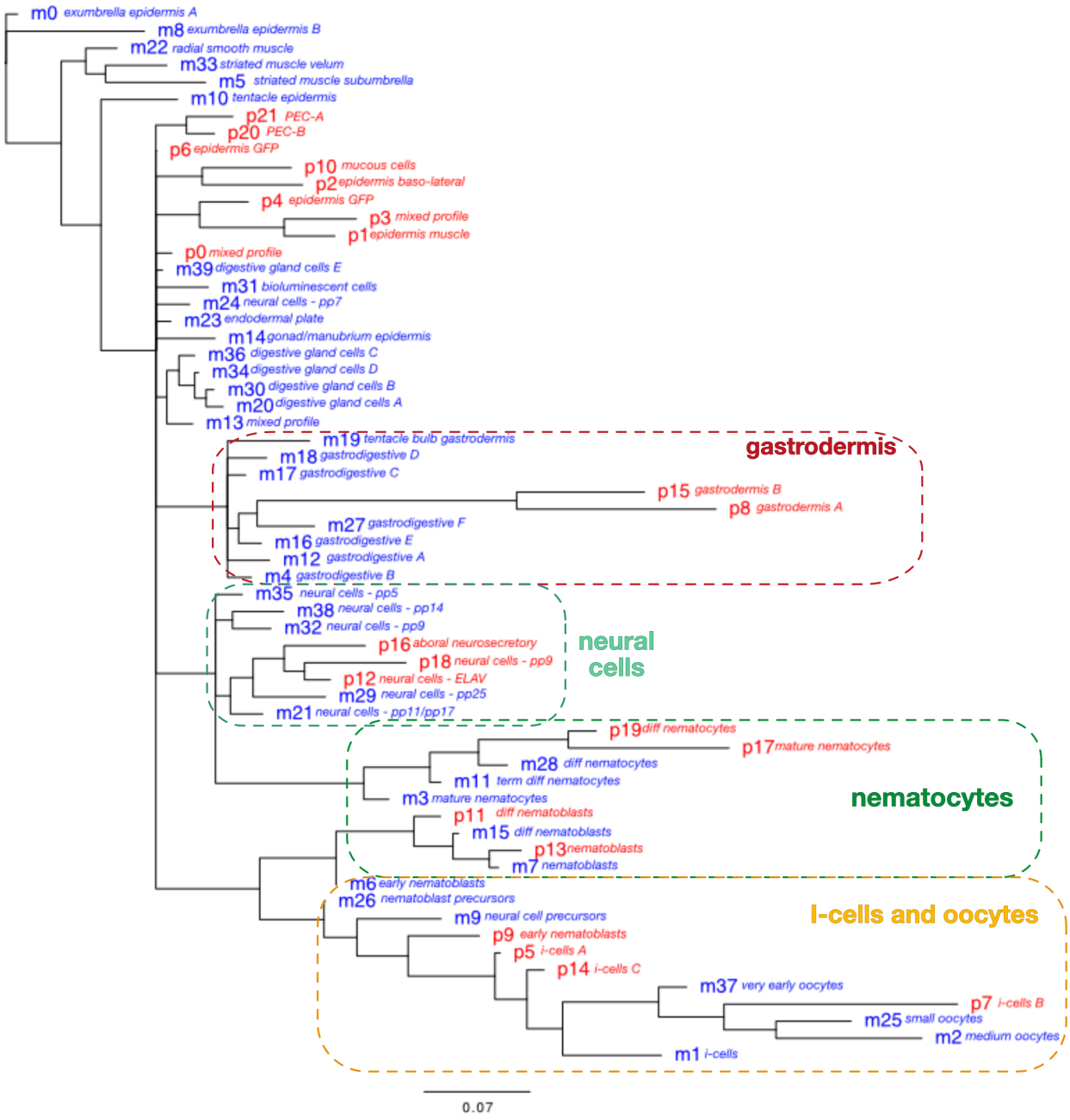
Finally, a remarkable similarity between transcriptional programs was detected within nematocyte developmental phases across the life cycle (Fig. 4.1). This represents the only example of finer level relationships that are consistent across our analysis. This evidence is corroborated by expression of candidate genes that recapitulate the two distinct phases of nematogenesis in both planula and medusa (see Chapter 2, Chapter 3 and Fig. 4.1). Nevertheless, expression of members of the *Fox, Tbox, bHLH* and *DMRT* families specific to planula nematoblasts suggest a difference in the regulation of the initial phase of nematogenesis in line with our experimental observations (Annex Table C and Chapter 3).

The transcriptional regulation of the origin and the differentiation of the nematocyte lineage in *Clytia* is similar to that reported for the anthozoan *Nematostella* (Babonis et al., 2022; Babonis & Martindale, 2017). Concerning the origin of this cnidarian specific cell type, previous studies reported the existence of a common neural-like-stem-cell ancestor, expressing the *Soxb2* transcription factor, which gave rise to neural cell and nematocytes in *Nematostella* (Richards & Rentzsch, 2014). Recent studies reported the characterisation of *Znf845* transcription factor in early stages of nematocyte development in *Nematostella* and *Clytia* (Babonis et al., 2022; Chari et al., 2021). A proposed evolutionary scenario suggests that the nematocyte lineage originated from the divergence of a sister cell type expressing a particular set of transcription factors, notably *Znf845* and *PaxA*, capable of modulating the expression of previous existing neural associated genes and recruiting novel ones (Babonis et al., 2022). The conservation of this pathway in *Clytia* argues in favour of an evolutionary scenario in which the nematocyte program evolved only once after the Cnidaria-Bilateria divergence.
Evidence of the expression of a conserved homeobox transcription factors across Cnidaria and Bilateria, notably *Pou4,* in neural cell and nematocytes in *Clytia* and *Nematostella,* supports the hypothesis of the existence of a progenitor cell type with sensory characteristics before the divergence (Ozment et al., 2021; Tournière et al., 2020; this thesis). Indeed, *Pou4* expression is detected in neural subpopulations of *ELAV*-expressing endodermal neural cells in *Nematostella* (Tournière et al., 2020). Additionally, *Nematostella Pou4* mutants fail to correctly assemble the nematocil and do not respond to tactile stimuli indicating a crucial role in the formation of sensory structures (Ozment et al., 2021; Tournière et al., 2020). Disruption of *NvPou4* resulted in polyp with a compromised ability to catch prey and feed caused by the lack of a complete assembly of fully functioning nematocytes (Tournière et al., 2020). Therefore young polyps were unable to reach maturity and it was not possible to observe a neural phenotype in adults (Tournière et al., 2020). On the other hand, sc-RNAseq data from adult *Nematostella* polyp reported the expression of *Pou4* in neural cells suggesting a possible involvement in the regulation of differentiated neural cells (Sebé-Pedrós, Saudemont, et al., 2018; Tournière et al., 2020). Taken together, these results suggest a conserved involvement of *Pou4* in the development of sensory structures and possibly neural identity across Cnidaria and Bilateria (Ozment et al., 2021). Consistently, Pou4 binding motifs are highly conserved across Cnidaria and Bilateria (Ozment et al., 2021). However, analysis of downstream targets indicates that *Pou4* regulates the expression of distinct sets of effector genes in hair cells and nematocytes in *Nematostella* (Ozment et al., 2021). Whether Pou4 regulated a distinct set of genes in the ancestral nematocyte or Pou4 targets diverged and became part of distinct transcriptional programs in different cell types remains to be determined (Ozment et al., 2021). During nematogenesis the first, distinct, transcriptional phase supporting nematocyst formation, would precede the common steps of neural differentiation regulated by Pou4.
This allows us to add to the already proposed scenario a possible preservation of a sensory program from the neural-like ancestor which is likely regulated by Pou4 in the terminal phases of nematogenesis in *Clytia*. This hypothesis would represent an explanation of the mechanosensory capabilities of the mature nematocytes.

To summarise, the scenario outlined above suggests that nematocytes evolved from a neural-like ancestor and specific transcription factors (such as Znf845) modulated the expression of existing genes in the sister cell type (Babonis et al., 2022). This allowed the recruitment of novel genes, for instance *Minicollagen* genes, into a first programme of nematogenesis, and

to the evolution of nematocytes as a unique cell type (Babonis et al., 2022). The detection of the conserved transcription factor Pou4 in neural cell subtypes and in mature nematocytes suggests that it may have regulated genes responsible for the sensory characteristics of the ancestor cell types, retained in nematocytes and sensory cells as “sister cell types'' (Arendt 2016). However, target genes of Pou4 in *Clytia* remain unknown, as are its interactions with other transcription factors. Analysis of Pou4 binding motifs and identification of downstream targets in *Clytia* will be useful to address this hypothesis and to shed light on the conservation of this program across Cnidaria.
Clytia is phylogenetically well placed to address the evolution of the nematocyte cell type in Hydrozoa as well as Medusozoa. In fact, this pathway appears to be shared between the planula larva and the more complex medusa. Conserved features of nematocyte origin and differentiation pathways across Anthozoa and Medusozoa strongly support a single origin of nematocytes after the divergence from Bilateria.

As already mentioned, further investigations are needed to support these observations concerning similarities and differences across life stages of *Clytia.* At the moment, our *biphy* analyses are not completely robust and the method certainly requires further testing and development. We can consistently detect relationships at the broad level of the cell classes but not at the finer one of the cell types. The one example of a finer level of comparison is within the nematocyte class, where we can consistently determine relationships across cell types in all our analyses. We anticipate that a detailed analysis of transcription factor and transcription factor-binding sites followed by the determination of downstream effector genes will enable a more solid comparative analysis by determining the “fingerprint” of the cell types (Arendt, 2008). At present, this comparative study provides further insights on the distribution of cell classes similarity relationships in *Clytia*. Once improved, our study can contribute to provide further knowledge concerning the evolution of the hydrozoan life cycle. Ultimately, our results can be exploited in the framework of a larger comparative analysis to address the evolution of cell types in the last cnidarian-bilaterian common ancestor.



ANNEX 4a - Hierarchical cell type tree of medusa and planula cell types Integrated medusa atlas versus initial planula atlas

Hierarchical representation of similarities between cell types and classes across the planula and the medusa of *Clytia.* Branches are labelled with the numbers assigned during clustering: Medusa cluster/cell types in blue; Planula clusters/cell types in red. Dashed boxes represent similarities.

ANNEX 4b - Hierarchical cell type tree of medusa and planula cell types Integrated medusa atlas versus recomputed planula atlas

Hierarchical representation of similarities between cell types and classes across the planula and the medusa of *Clytia.* Branches are labelled with the numbers assigned during clustering: Medusa cluster/cell types in blue; Planula clusters/cell types in red. Dashed boxes represent similarities.

GENERAL DISCUSSION

The studies I have presented in this thesis contribute to expanding the knowledge of the cell types in Cnidaria. In particular, the cell atlases I have generated of the adult medusa and the planula larva of *Clytia hemisphaerica* provide novel insights concerning the characterisation of the cell types of the medusozoan subgroup, Hydrozoa, to which *Clytia* belongs.

Clytia represents an emerging cnidarian model system which unlike the polyp-only hydrozoan model *Hydra*, exhibits a triphasic, or "complete", hydrozoan life cycle including a ciliated planula larva, sessile polyp forming colony and free-swimming medusa stages (Houliston et al., 2010). The medusa constitutes the fundamental innovation of the phylum as well as the most complex cnidarian body form. The medusa body is organised according to a tetra-radial symmetry and it exhibits distinct organs such as the gonads, the manubrium and the tentacle bulbs, which produce tentacles filled with stinging cells. Each of the organs displays an outer layer, or epidermis and an inner layer, or gastrodermis. Striated fast-contractile muscle fibres enable rapid contractions which characterise the pulsatile swimming movements. A well organised nervous system which innervates the subumbrella and is condensed into two nerve rings at the edge of the bell coordinates sophisticated medusa-specific behaviours, including swimming and feeding (Weissbourd et al., 2021). At the bell margin, nerve rings are connected with sensory structures or statocysts (Houliston et al., 2010).
The ease of culturing *Clytia* (Lechable et al., 2020) and the accessibility of all the life stages facilitated my work of establishing single cell transcriptomic approaches in order to address the complexity of the medusa against the much simpler planula stage at the cell type level. The planula exhibits a basic body organisation on an oral-aboral axis with the two tissue layers, the outer epidermis and the inner gastrodermis (Kraus et al., 2020). Prior to this study, our knowledge on the planula cell types relied on old morphologically-based descriptions (Bodo & Bouillon, 1968) and no molecular signatures were available.
Besides representing a foundation to perform comparative analysis at the cell type level across stages, the cell atlases of the planula and the medusa of *Clytia* provide new insights on previously uncharacterised cell types as well as supporting existing knowledge on those already described for both stages.

The availability of single cell transcriptomics and the progressive uncovering of detailed cell type molecular signatures sets the ground for a new perspective that envisages the exploitation of molecular information to study evolution at the cell type level (Arendt, 2008; Arendt et al., 2016). In detail, the association of cell type-specific transcriptional regulatory programs and effector genes constitutes a core regulatory complex, or CoRC. When specific CoRCs evolve in a partially independent manner this can be leveraged to determine the identity of a given cell type. As such, the cell type can be considered as an "evolutionary unit" and the molecular information can be exploited to "trace" its evolutionary history across species (Arendt et al., 2016).

My study is well placed in this context. *Clytia* is part of Cnidaria that diverged from Bilateria around 550 Mya. In a much broader comparative framework, the information concerning molecular signatures of cell types of *Clytia,* along with data from other studies in which the molecular characterisation of the cell types of other cnidarian species has been achieved, can inform on the evolution of Metazoa.

In the following sections I discuss the main findings of my study and how those contribute to further knowledge on the evolution of the hydrozoan life cycle and of the cell types in *Clytia*. I also discuss the challenges I encountered and the constraints of some experimental and computational approaches I used.

Briefly, my findings can be summarised as follows:
i) The generation of the first cell atlas of the medusa of *Clytia* in collaboration with the Pachter and Anderson groups at Caltech (Chari et al., 2021). Analysis of the atlas revealed a rich diversity of cell types within the single classes and allowed a detailed characterisation of developmental stages of nematogenesis and previously unreported neural cell types (Chari et al., 2021). This atlas the molecular characterisation of 36 cell types grouped in eight broad classes: the epitheliomuscular cells of the two body layers (epidermis and gastrodermis), the i-cells, neural cells nematocytes, digestive gland cells and bioluminescent cells. I integrated this atlas with data that I produced and finally generated an additional medusa atlas which represents a consolidation to the initial medusa cell atlas

ii) The generation of the cell atlas of the planula of *Clytia.* The development of the experimental dissociation protocol for the planula stage was a challenge and its refinement was crucial for the generation of the data. This atlas includes 19 cell types grouped into eight cell classes namely the epidermis, the gastrodermis, i-cells, neural cells, nematocytes as well as aboral neurosecretory cells, mucous cells and putative excretory cells (PEC), which were not previously characterised at the transcriptional level. Analysis of the atlas allowed us to identify cell types already reported (Bodo & Bouillon, 1968) and characterise them further. Interestingly, developmental stages of nematogenesis coincide with those in the adult. However analysis of their distributions during planula development suggests that some nematocyte stages of the planula might be regulated differently to those of the medusa.

iii) Analysis of the planula and medusa cell atlas data to explore shared transcriptional signatures at the cell type level. Previous studies had identified a larger set of transcription factors in the complex medusa stages of *Clytia* compared to the other life cycle stages (Leclère et al., 2019). This comparative analysis aimed to shed light on whether the increased molecular repertoire mirrors an expansion of the cell types in the adult. I have started to compare molecular signatures of the medusa cell types against the ones of the planula cell type and asked which of those were shared or unique.
This comparative study is still in its early stages, but it already has suggested some biological interpretations. It revealed some degree of similarity between certain planula and medusa cell classes, namely gastrodermis, i-cells, neural cells and nematocytes. On the other hand, other classes showed a lower likelihood of being shared between stages, such as the epidermis, mucous and putative excretory cells (PEC) of the planula, and the epidermis and gland digestive cells of the medusa. Thus this comparative approach provides an initial vision of the similarities across stages at the broader level of the cell classes. However, our methods require further assessment, and ideally more planula data to enable more robust comparisons at the cell type level. The abundance of nematocytes cells at various stages of differentiation in our data have enabled a finer level of comparison, revealing general equivalence of the trajectories. Indeed, stages of nematogenesis show a high degree of similarity of gene usage across the life cycle. As already mentioned, our data, supported by in situ hybridisation

profiles, suggest that the regulation of those sets of genes might differ between the two life stages, with nematogenesis in the planula being stalled for most cells at the transition from nematoblast stage to the nematocyte maturation phase. In parallel, I explored the expression of cell type/cell classes specific transcription factors. Taken together, the results of the initial exploration of regulatory programs along with the insights we could gather from the comparative analyses, allowed us to draw evolutionary scenarios on the origin and development of the cell types in *Clytia*.

Hydrozoan cell types: a viewpoint from Clytia

The medusa cell atlas I presented in Chapter 2 constitutes a tool to address the development and the evolution of the most complex form within Cnidaria. We uncovered a much higher degree of cell diversity compared with the hydrozoan model system *Hydra* and the anthozoan *Nematostella* (Chari et al., 2021; Sebé-Pedrós, Saudemont, et al., 2018; Siebert et al., 2019). Analysis of cluster-specific selected candidates and *in situ* expression pattern revealed that the epidermis cell class can be subdivided in seven cell types including basic structural epidermis types and unique smooth and striated muscle types. Previously underappreciated richness was detected within the gastrodermis that can be subdivided into eight cell classes. We characterised two of those as “structural” gastrodermis types that were enriched in extracellular matrix formation genes and are, therefore, likely involved in the production of the mesoglea. The remaining six cell types showed a compartmentalised distribution across the gastrodermis of the medusa and are very likely involved in different digestive functions the precise role of which remain to be determined (Chari et al., 2021). We identified the hydrozoan stem cell population, or i-cells, which are thoroughly described in *Hydra* and give rise to derivatives such as neural, gland, germ cells and nematocytes (Bode, 1996). We addressed i-cell-mediated neurogenesis and nematogenesis in *Clytia* by performing pseudotime analysis of isolated i-cells, neural cells and nematocytes. Our results indicate that in *Clytia,* as in *Hydra,* neurogenesis and nematogenesis initiate from i-cell-like precursors. We could not rule out that i-cell-like precursors are also responsible for generating digestive gland cells in *Clytia* (Chari et al., 2021). The lack of clear link between the i-cell and digestive gland cell clusters is possibly due to self-renewal mechanisms which contribute to the maintenance of digestive gland cells with a distinctive transcriptional profile (Siebert et al., 2019). We assigned six clusters as gland digestive cells which we characterised on the basis of strongly expressed digestive enzymes (Chari et al., 2021). The massive production of trypsin, ShKT, fibrinogen and chitinase domain containing proteins argues in favour of an involvement in primary digestion of these cells. The distinctions between each of these types is unknown. We assigned seven neural cell types based on the expression of ELAV (Nakanishi et al., 2012) and other neural expressed genes including several neuropeptide precursors. We predicted novel neuropeptide precursors by sequence-based analysis of neural markers (Chari et al., 2021) adding to the list of those already known from other studies (Quiroga Artigas et al., 2018). Among the neural cell types we identified one cluster as neural cell precursors and six clusters of differentiated neural cells producing neuropeptides. Reclustering analysis of isolated neural cells (Chari et al., 2021) and additional analysis of *in situ* expression patterns

of neuropeptide precursors revealed more putative neural subtypes present in the medusa. Whether those represent distinct subtypes with different functions or an artefact of the reclustering will require functional analysis of selected candidate genes to be resolved. Finally, as in Chari et al. (2021), I assigned seven cell types to the nematocyte class. The large number of cells recovered enabled the detailed characterisation of nematogenesis and nematocyte development. Indeed, we could reconstruct the nematocyte developmental trajectory and distinguish two distinct but continuous transcriptional programs. We characterised the initial program as the first phase of nematogenesis during which the nematoblasts differentiate from i-cells and produce the typical nematocyte organelle or stinging capsule. Following the production of the capsule, the nematocytes initiate a phase of differentiation and migration toward the tentacle which is characterised by an abrupt downregulation of previously expressed genes and by the expression of a set of genes which are enriched in the “stereovilli” of the vertebrate hair cells. This second phase terminates with the production of the nematocil, the trigger for nematocyte discharge.

The planula cell atlas presented in Chapter 3 constitutes the first cell type atlas of a hydrozoan larva. The presumed high fragility of the planula cells required several refinement steps of the dissociation protocol. Following optimisation of the experimental procedures I excluded two clusters of likely technical artefacts, decreasing the number of cells available for further analyses. The final atlas represents a useful resource for the characterisation of the planula cell types. It provides molecular signatures of as-yet uncharacterised subtypes of the epidermis and the gastrodermis. Analysis of *in situ* expression patterns of selected candidate genes suggests that the epidermis subtypes are potentially planula specific. Their involvement in specific functions remains to be determined. I assigned two clusters as neural cells on the basis of the expression of ELAV proteins (Nakanishi et al., 2012). Analysis of the molecular signatures indicate that each of those cell types express multiple neuropeptide precursors. I performed reclustering analysis of isolated neural cells to explore them further, however, the low number of cells used in those analyses yielded sub-clusters which could not be confidently assigned to neural subtypes. *In situ* hybridisation profiles of neuropeptide precursor genes suggested further diversity in neural subpopulations, which needs to be confirmed. Interestingly, the nematocyte markers used to determine the developmental stages of nematocyte differentiation recapitulated the two distinct phases of nematogenesis in planula. Indeed, our comparative analyses show an almost complete match of nematocyte developmental stages across the life cycle. Surprisingly, analysis of the *in situ* expression patterns indicated that the differentiation occurs in early stages of planula development. Indeed, this analysis indicates that the initial program is stalled for most of the cells and that nematoblasts continue to accumulate until later planula stages. Furthermore the expression of *Nematocilin* is detected in only a few cells, which are concentrated at the oral pole of the planula at 72 hpf. Mature nematocytes are present in the disk-like arrangement around the aboral pole of the planula after settlement (Krasovec et al., 2021) even though the non-feeding planula has no use for nematocytes to catch prey, in contrast to the polyp and the medusa. This suggests that the two nematocyte programs are active before metamorphosis and that nematoblasts with a fully formed capsule as well as completely mature nematocytes serve as a reservoir for the future primary polyp. Analysis of transcription factors revealed the presence of Fox, bHLH, Tbox and DMRT families in nematoblasts and nematocytes uniquely at planula

stage. This suggests that a difference in regulation might be responsible for the delayed programs in the planula with respect to the medusa.

Origin and evolution of the cnidarian stinging cells

The detailed characterisation of the molecular signatures of nematocytes developmental stages allowed a comparison with the nematogenesis pathways described in *Nematostella*. In *Nematostella*, nematocytes as well as neurons differentiate from a common neural-like-stem-cell precursor which is characterised by the expression the *Sox* family transcription factor Soxb2 (Richards & Rentzsch, 2014). An additional transcription factor, *Znf845*, is co-expressed in those cells, likely downstream of Soxb2, and it is responsible for the triggering of the diversification into nematocytes by inducing the nematocyte phenotype and potentially repressing the neural phenotype (Babonis et al., 2022). Znf845 is also partially co-expressed with PaxA which is described as a regulator of the differentiation of nematocytes (Babonis & Martindale, 2017) and a recent study demonstrated that Znf845 acts upstream of PaxA (Babonis et al., 2022). Thus, Znf845 is essential to nematogenesis. These findings suggested an evolutionary scenario in which the nematocyte lineage would originate with the emergence of sister cell types expressing a set of regulators, such as Znf845 and PaxA. These genes would then modulate the expression of neural-like programs previously existing within the ancestor and recruit novel genes which are responsible for the generation of the nematocyte capsule (Babonis et al., 2022). Consistent with these findings, our analysis of the single cell data of and of *in situ* expression patterns in *Clytia* demonstrated that Znf845 is expressed in a population of i-cell progenitors (Chari et al., 2021) andis likely playing a similar role to that described in *Nematostella.* Furthermore, analysis of the presence/absence matrix revealed the expression of an ortholog of PaxA in early and differentiated nematoblasts but not in i-cells for both planula and medusa. This supports a conserved evolution of the Znf845/PaxA pathway in *Clytia* and a common origin of the nematocyte lineage across Cnidaria.

A particular combination of markers extracted from our medusa and planula atlas showed similarity between mature nematocytes and neural cells in our hierarchical cell-type tree. Thus, we returned to our presence/absence matrix to investigate whether this similarity was driven by particular transcription factors. Interestingly, we detected the expression of the homeobox transcription factor Pou4 (also known as Brn3) in mature nematocytes and neural cells across stages. *Pou4* genes are responsible for conserved functions in Bilateria, such as the differentiation and the maintenance of neural identity which has been demonstrated in mice, *Drosophila* and *C.elegans* (Leyva-Díaz et al., 2020; Serrano-Saiz et al., 2018). Recent studies have demonstrated that Pou4 genes are involved in the regulation of the nematocyte cilium and in the maturation of mechanosensory “hair cells" in *Nematostella* (Ozment et al., 2021; Tournière et al., 2020). Furthermore, single cell transcriptomic data of the *Nematostella* adult polyp have reported the expression of Pou4 in neural cells suggesting a possible role in the regulation of their differentiation (Sebé-Pedrós et al., 2018; Tournière et al., 2020). Taken together, these results indicate that *Pou4* plays a role in the development of sensory structures and possibly neural identity across Cnidaria as it does in Bilateria (Ozment et al., 2021). The

detection of Pou4 in neural cells and mature nematocyte in *Clytia* supports the already proposed scenario which suggests that the evolution of the role of *Pou4* in the determination of neural identity and sensory structures predated the Cnidaria-Bilateria divergence (Ozment et al., 2021; Tournière et al., 2020). Indeed, Pou4 binding motifs are highly conserved across Cnidaria and Bilateria although downstream effector genes differ depending on the cell type (Ozment et al., 2021). What is still unresolved is whether a Pou4 transcription factor regulated a distinct set of genes in the ancestral nematocytes or if targets of *Pou4* diverged and evolved independently in distinct cell types (Ozment et al., 2021). Furthermore, Pou4 is defined as a "terminal selector", such as transcription factors considered responsible for the regulation of cell type identities acting through repression of alternative types (Hobert, 2008). One possibility is that Pou4 regulated "sensory genes" in the ancestor neural-like-stem cells of nematocytes and that those characteristics have somehow been retained in nematocytes and neural sensory cells as "sister cell type" where Pou4 modulates the expression of different effector genes (Leyva-Díaz et al., 2020). However at present, target genes as well as interaction with other transcription factors of Pou4 have not been identified in *Clytia.* The characterisation of these interactions will enable this hypothesis to be tested and expand the knowledge on the conservation of this program in Cnidaria.

Complexity of the medusa form compared to the simpler planula

We identified a large diversity at the cell type level in the medusa. However, a certain degree of diversity can be also observed at planula stage considering that certain planula cell classes do not show similarity with any of the medusa ones. For instance the epidermis and the sets of secretory cell classes of the two stages, which include mucous and putative excretory cells (PEC) of the planula and gland digestive cells of the medusa, showed a consistent dissimilarity in our analyses. These cell classes are likely associated with behaviours that are specific to the stages. For instance, the gland digestive cells of the medusa are likely involved in primary digestion of prey, indicating their involvement in feeding behaviours which are a feature of the medusa stage. On the other hand, mucous cells are probably involved in the secretion of molecules that contribute to securing the planula to the ground during settlement, which is a behaviour specific to the planula. Similarly, PEC might be involved in unknown stage-specific processes, likely common to the polyp but not to the medusa stage. Likewise, the epidermis cell types of the medusa, notably muscle fibres, are involved in specialised behaviour characteristic of the medusa stage only, such as the sophisticated swimming motion which is not present at the planula stage. This evidence enables us to speculate on the existence of putative ancestral epidermis and secretory cell types which would have diversified into sister cell types that have further differentiated in the two life stages in line with their behavioural complexity. At present, our comparative study has not resolved the evolutionary origin of cell classes unique-to-stage. The refinement of our comparative analyses is essential to enable the comparison at the finer level of the cell types and resolve the relationships within cell classes. This will also enable the relationships across classes to be investigated, for instance putative relationships across secretory cells and epidermis or I-cells, which would allow us to gain insights on their evolutionary origin.

Is the larger set of transcription factors observed at medusa stage directly associated with an expansion of cell types?

Overall my analyses indicate that an overall richer diversity of cell types is associated with the medusa stage with respect to the planula. These do not correspond to completely novel cell types but rather to specialised subtypes. Indeed the specialised subtypes observed at medusa stage belong to cell classes that consistently show similarity in our analyses, notably the gastrodermis, the i-cells, the neural cells and the nematocyte classes. Accordingly, we can speculate that the larger set of transcription factors at medusa stage would be involved in the emergence of diversified regulatory pathways following the evolution of putative ancestors of those cell classes. In more detail, the newly evolved regulatory pathways along with the recruitment of effector genes would have led to the emergence of highly specialised subtypes within those classes at medusa stage. Open questions are, for instance, which transcription factors control the emergence of the diversified subtypes at medusa stage? Are those involved in the regulation of subtypes at planula stage? Is the diversity of subtypes associated with different regulatory programs such as distinct sets of transcription factors and effector genes responsible for the emergence of novel subtypes? Or, In contrast, do the same transcription factors recruit and regulate distinct sets of effector genes in the different subtypes across life stages? The determination of cell type specific transcription factors, the analysis of binding domains and the identification of the effector genes are crucial steps that must be pursued to resolve those questions.

Evolution of the hydrozoan life cycle: perspectives from neural cell evolution

Speculations on the evolution of the hydrozoan life cycle can be drawn on the basis of the similarity within the neural cell class that we consistently observed across life stages in our comparative analyses. However some evidence suggests that from a developmental point of view neural cells might derive from different precursors at the two stages. Firstly, our comparative analyses show some degree of similarity between neural precursors of the medusa and i-cell classes as expected. This is supported by the expression of neurogenic transcription factors belonging to bHLH and Sox families such as *bHLH6* and *Sox10* transcription factors on the pseudotime trajectory from i-cells to neural cells (Chari et al., 2021). Thus, neural cells originate from an i-cell-like precursor at medusa stage. By contrast ,at the planula stage expression of *Sox10* and *bHLH6* are not detected in the i-cells but rather in epidermis, suggesting an epidermal neurogenesis. This is supported by the detection of *Sox10* expression in particular cells at the base of the ectodermal layer at gastrula stage (Kraus et al., 2020). The hypothesis of the epidermal origin of neural cells at planula stage is the subject of an ongoing investigation in the lab. Results are confirming the conclusion of

Thomas et al. (1987) that certain neural subtypes, along with mucous cells, derive from aboral and lateral epidermis. In parallel, some neural subtypes likely originate from i-cell-like precursors as in medusa, as supported by the expression of the bHLH family transcription factor Neurogenin in i-cells at the planula stage (confirmed by *in situ* hybridisation: S. Chevalier, unpublished). Neurogenin is likely involved in the regulation of the i-cell-mediated neurogenic pathway in the medusa as indicated by its detection in neural cell precursors at medusa stage (Chari et al., 2021). As reported in Chapter 4, our data indicate that this pathway is shared across life stages. Thus, neural cells at medusa stage probably originate exclusively (or at least dominantly) from i-cell-like precursors, while at planula stage they derive both from i-cells and by ectodermal neurogenesis.

In most animals neurogenesis starts from epithelial cells (reviewed in Hartenstein & Stollewerk, 2015). In *Nematostella* neural cells originate from epithelial cells of the ectoderm as well as endoderm (Nakanishi et al., 2012). Additionally, analysis of single cell transcriptomics data of *Nematostella* adult and larval neurons revealed that broad neural signatures are shared across *Nematostella* and *C.elegans* supporting the existence of a neural-like cell type in the common Cnidaria/Bilateria ancestor. Although signatures of subtypes do not show similarities suggesting that the assembly of specific molecular components which determine and maintain distinct neural subtypes evolved independently in each lineage (Sebé-Pedrós, Saudemont, et al., 2018). The ectodermal neurogenesis at planula stage of *Clytia* indicates that the larva partially retained a conserved neurogenic program that predated the Cnidaria-Bilateria divergence. The lack of evidence of this neurogenic program at medusa stage, which has been replaced by an i-cell-mediated neurogenesis, corroborates that the i-cells represent a derived trait evolved at the onset of Hydrozoa. Furthermore, considering the partial retention of an ancient neurogenic program, we can speculate that a planula-like body form was present before the emergence of the Hydrozoa and probably of all cnidarian (Nielsen, 1998; Salvini-Plawen, 1978).

Regardless, improving the sampling of cnidarian species at the cell type level is essential to resolve this hypothesis.

FUTURE PERSPECTIVES and CONCLUSIONS

In the general discussion I outlined that uncovering and comparing the molecular signatures of the cell types in *Clytia* allowed inferences on the complexity of the medusa stage and the evolution of the cnidarian life cycle to be drawn. Furthermore, this study provides novel insights on the evolution of individual cell classes for instance the nematocytes and neural cells. However, my study is limited to the characterisation of the cell types in *Clytia* at the molecular and, to some extent, morphological levels, therefore questions concerning specific roles and involvement in certain functions of the cell types remain to be resolved.
The experimental protocols presented in this work will serve as a basis to perform single cell transcriptomics in other cnidarian species. This would increase the availability of cnidarian cell atlases and contribute to expanding the knowledge on cnidarian cell types. In the context of *Clytia,* single cell transcriptomics for other stages is now possible. One possibility would be to explore cell types at earlier stages of planula development which would shed light on the dual evolutionary origin of the neural cells at planula stage. A second possibility would be to examine cell type molecular signatures at polyp stage, which would result in the characterisation of the cell types of the whole life cycle and in the resolution of unique and shared cell types across stages. Furthermore, the atlases presented in this study constitute a tool for the *Clytia* and the hydrozoan community. They provide detailed molecular information and several candidates that can be exploited for future functional analysis. Additionally, the initial characterisation of the ultrastructure of the cells I achieved during this work, added a further morphological level of characterisation of the cell types of *Clytia*. In the future, this can be expanded to the characterisation of the 3D structure of the cells which would allow better understanding of their specialised compositions as well as reconstruction of cell-to-cell interactions within tissues.

Some aspects concerning the computational analysis are not completely resolved. For instance, it is essential to refine the batch effect correction analysis to enable easy reclustering of specific cells and further investigation of subtypes. Additionally, increasing the robustness of our comparative analysis would establish a novel approach for the comparison of the cell types and the determination of similarities at the finer level of the cell type. This would lead to gain further insights into the evolution of the hydrozoan life cycle.
Finally, this study can be integrated in a much broader comparison with similar data of species belonging to other Cnidaria and Bilateria taxa. Ultimately, it will enable the hypothesis of the reconstruction of the evolutionary history of the cell types to be tested and potentially resolved.

BIBLIOGRAPHY

Appeltans, W., Ahyong, S. T., Anderson, G., Angel, M. V., Artois, T., Bailly, N., Bamber, R., Barber, A., Bartsch, I., Berta, A., Błażewicz-Paszkowycz, M., Bock, P., Boxshall, G., Boyko, C. B., Brandão, S. N., Bray, R. A., Bruce, N. L., Cairns, S. D., Chan, T.-Y., … Costello, M. J. (2012). The Magnitude of Global Marine Species Diversity. *Current Biology*, *22*(23), 2189–2202. https://doi.org/10.1016/j.cub.2012.09.036

Arendt, D. (2003). Evolution of eyes and photoreceptor cell types. *The International Journal of Developmental Biology*, *47*(7–8), 563–571.

Arendt, D. (2005). Genes and homology in nervous system evolution: Comparing gene functions, expression patterns, and cell type molecular fingerprints. *Theory in Biosciences*, *124*(2), 185–197. https://doi.org/10.1016/j.thbio.2005.08.002

Arendt, D. (2008). The evolution of cell types in animals: Emerging principles from molecular studies. *Nature Reviews Genetics*, *9*(11), 868–882. https://doi.org/10.1038/nrg2416

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

Babonis, L. S., Enjolras, C., Ryan, J. F., & Martindale, M. Q. (2022). A novel regulatory gene promotes novel cell fate by suppressing ancestral fate in the sea anemone *Nematostella vectensis*. *Proceedings of the National Academy of Sciences*, *119*(19), e2113701119. https://doi.org/10.1073/pnas.2113701119

Babonis, L. S., & Martindale, M. Q. (2017). PaxA, but not PaxC, is required for cnidocyte development in the sea anemone Nematostella vectensis. *EvoDevo*, *8*(1), 14. https://doi.org/10.1186/s13227-017-0077-7

Bakshani, C. R., Morales-Garcia, A. L., Althaus, M., Wilcox, M. D., Pearson, J. P., Bythell, J. C., & Burgess, J. G. (2018). Evolutionary conservation of the antimicrobial function of mucus: A first defence against infection. *Npj Biofilms and Microbiomes*, *4*(1). https://doi.org/10.1038/s41522-018-0057-2

Balasubramanian, P. G., Beckmann, A., Warnken, U., Schnölzer, M., Schüler, A., Bornberg-Bauer, E., Holstein, T. W., & Özbek, S. (2012). Proteome of Hydra nematocyst. *Journal of Biological Chemistry*, *287*(13), 9672–9681. https://doi.org/10.1074/jbc.M111.328203

Ball, E. E., Hayward, D. C., Reece-Hoyes, J. S., Hislop, N. R., Samuel, G., Saint, R., Harrison, P. L., & Miller, D. J. (2002). Coral development: From classical embryology

to molecular control. *The International Journal of Developmental Biology*, *46*(4), 671–678.

Bode. (1996). *The interstitial cell lineage of hydra: A stem cell system that arose early in evolution*. 10.

Bode, H. R. (2003). Head regeneration inHydra. *Developmental Dynamics*, *226*(2), 225–236. https://doi.org/10.1002/dvdy.10225

Bode, H. R., Berking, S., David, C. N., Gierer, A., Schaller, H., & Trenkner, E. (1973). Quantitative analysis of cell types during growth and morphogenesis in Hydra. *Wilhelm Roux Archiv F*[illegible]*r Entwicklungsmechanik Der Organismen*, *171*(4), 269–285. https://doi.org/10.1007/BF00577725

Bode, H. R., Heimfeld, S., Chow, M. A., & Huang, L. W. (1987). Gland cells arise by differentiation from interstitial cells in Hydra attenuata. *Developmental Biology*, *122*(2), 577–585. https://doi.org/10.1016/0012-1606(87)90321-6

Bodo, F., & Bouillon, J. (1968). *ÉTUDE HISTOLOGIQUE DU DÉVELOPPEMENT EMBRYONNAIRE*.

Boero, F., Bouillon, J., & Piraino, S. (2005). The role of Cnidaria in evolution and ecology. *Italian Journal of Zoology*, *72*(1), 65–71. https://doi.org/10.1080/11250000509356654

Bosch, T. C. G., & David, C. N. (1987). Stem cells of Hydra magnipapillata can differentiate into somatic cells and germ line cells. *Developmental Biology*, *121*(1), 182–191. https://doi.org/10.1016/0012-1606(87)90151-5

Bosch, T. C. G., Klimovich, A., Domazet-Lošo, T., Gründer, S., Holstein, T. W., Jékely, G., Miller, D. J., Murillo-Rincon, A. P., Rentzsch, F., Richards, G. S., Schröder, K., Technau, U., & Yuste, R. (2017). Back to the Basics: Cnidarians Start to Fire. *Trends in Neurosciences*, *40*(2), 92–105. https://doi.org/10.1016/j.tins.2016.11.005

Bouillon, J. (1966). *LES CELLULES GLANDULAIRES DES HYDROIDES ET DES HYDROMÉDUSES. LEUR STRUCTURE ET LA NATURE DE LEURS SÉCRÉTIONS. p a r*.

Bouillon, J., & Boero, F. (2000). *THE HYDROZOA: A NEW CLASSIFICATION IN THE LIGTH OF OLD KNOWLEDGE. 24*, 44.

Bridge, D., Cunningham, C. W., DeSalle, R., & Buss, L. W. (1995). Class-level relationships in the phylum Cnidaria: Molecular and morphological evidence. *Molecular Biology and Evolution*, *12*(4), 679–689. https://doi.org/10.1093/oxfordjournals.molbev.a040246

Brown, B., & Bythell, J. (2005). Perspectives on mucus secretion in reef corals. *Marine Ecology Progress Series*, *296*, 291–309. https://doi.org/10.3354/meps296291

Browne, E. N. (1909). The production of new hydranths in Hydra by the insertion of small grafts. *Journal of Experimental Zoology*, *7*(1), 1–23. https://doi.org/10.1002/jez.1400070102

Brusca, R. C., & Brusca, G. J. (1996). *Invertebrates, First edition*.

Bythell, J. C., & Wild, C. (2011). Biology and ecology of coral mucus release. *Journal of Experimental Marine Biology and Ecology*, *408*(1–2), 88–93. https://doi.org/10.1016/j.jembe.2011.07.028

Campbell, R. D. (1967). Tissue dynamics of steady state growth in Hydra littoralis. *Developmental Biology*, *15*(5), 487–502. https://doi.org/10.1016/0012-1606(67)90039-5

Campbell, R. D., & Bode, H. R. (1983). Terminology for Morphology and Cell Types. In H. M. Lenhoff (Ed.), *Hydra: Research Methods* (pp. 5–14). Springer US. https://doi.org/10.1007/978-1-4757-0596-6_2

Chapman, J. A., Kirkness, E. F., Simakov, O., Hampson, S. E., Mitros, T., Weinmaier, T., Rattei, T., Balasubramanian, P. G., Borman, J., Busam, D., Disbennett, K., Pfannkoch, C., Sumin, N., Sutton, G. G., Viswanathan, L. D., Walenz, B., Goodstein, D. M., Hellsten, U., Kawashima, T., … Steele, R. E. (2010). The dynamic genome of Hydra. *Nature*, *464*(7288), 592–596. https://doi.org/10.1038/nature08830

Chari, T., Weissbourd, B., Gehring, J., Ferraioli, A., Leclère, L., Herl, M., Gao, F., Chevalier, S., Copley, R. R., Houliston, E., Anderson, D. J., & Pachter, L. (2021). Whole-animal multiplexed single-cell RNA-seq reveals transcriptional shifts across Clytia medusa cell types. In *Sci. Adv* (Vol. 7, p. 1683). https://www.science.org

Charles Darwin. (1859). *On the origin of Species*. http://darwin-online.org.uk/

Chevalier, S., Martin, A., Leclère, L., Amiel, A., & Houliston, E. (2006). Polarised expression of FoxB and FoxQ2 genes during development of the hydrozoan Clytia hemisphaerica. *Development Genes and Evolution*, *216*(11), 709–720. https://doi.org/10.1007/s00427-006-0103-6

Chiori, R., Jager, M., Denker, E., Wincker, P., Da Silva, C., Le Guyader, H., Manuel, M., & Quéinnec, E. (2009). Are hox genes ancestrally involved in axial patterning? Evidence from the hydrozoan Clytia hemisphaerica (cnidaria). *PLoS ONE*, *4*(1). https://doi.org/10.1371/journal.pone.0004231

Choksi, Y. A., Reddy, V. K., Singh, K., Barrett, C. W., Short, S. P., Parang, B., Keating, C. E., Thompson, J. J., Verriere, T. G., Brown, R. E., Piazuelo, M. B., Bader, D. M., Washington, M. K., Mittal, M. K., Brand, T., Gobert, A. P., Coburn, L. A., Wilson, K. T., & Williams, C. S. (2018). BVES is required for maintenance of colonic epithelial integrity in experimental colitis by modifying intestinal permeability. *Mucosal Immunology*, *11*(5), 1363–1374. https://doi.org/10.1038/s41385-018-0043-2

Collins, A. G. (2002). *Phylogeny of Medusozoa and the evolution of cnidarian life cycles* (J. EVOL. BIOL).

Collins, A. G. (2009). *Recent Insights into Cnidarian Phylogeny*. *38*, 11.

Condamine, T., Jager, M., Leclère, L., Blugeon, C., Lemoine, S., Copley, R. R., & Manuel, M. (2019). Molecular characterisation of a cellular conveyor belt in Clytia medusae. *Developmental Biology*. https://doi.org/10.1016/j.ydbio.2019.09.001

David, C. N., & Challoner, D. (1974). Distribution of Interstitial Cells and Differentiating Nematocytes in Nests in *Hydra attenuata*. *American Zoologist*, *14*(2), 537–542. https://doi.org/10.1093/icb/14.2.537

David, C. N., & Murphy, S. (1977). Characterization of interstitial stem cells in hydra by cloning. *Developmental Biology*, *58*(2), 372–383. https://doi.org/10.1016/0012-1606(77)90098-7

Davis, L. E., Burnett, A. L., Haynes, J. F., Osborne, D. G., & Spear, M. L. (1968). Histological and ultrastructural study of the muscular and nervous systems inHydra. II. Nervous system. *Journal of Experimental Zoology*, *167*(3), 295–331. https://doi.org/10.1002/jez.1401670305

Deng, Q., Ramsköld, D., Reinius, B., & Sandberg, R. (2014). Single-Cell RNA-Seq Reveals Dynamic, Random Monoallelic Gene Expression in Mammalian Cells. *Science*, *343*(6167), 193–196. https://doi.org/10.1126/science.1245316

Denker, E., Manuel, M., Leclère, L., Le Guyader, H., & Rabet, N. (2008). Ordered progression of nematogenesis from stem cells through differentiation stages in the tentacle bulb of Clytia hemisphaerica (Hydrozoa, Cnidaria). *Developmental Biology*, *315*(1), 99–113. https://doi.org/10.1016/j.ydbio.2007.12.023

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Dunn, C. W., Giribet, G., Edgecombe, G. D., & Hejnol, A. (2014). Animal Phylogeny and Its Evolutionary Implications. *Annual Review of Ecology, Evolution, and Systematics*, *45*(1), 371–395. https://doi.org/10.1146/annurev-ecolsys-120213-091627

Eisenman, E. A., & Alfert, M. (1982). A new fixation procedure for preserving the ultrastructure of marine invertebrate tissues. *Journal of Microscopy*, *125*(1), 117–120. https://doi.org/10.1111/j.1365-2818.1982.tb00327.x

Fourrage, C., Swann, K., Garcia, J. R. G., Campbell, A. K., & Houliston, E. (2014). An endogenous green fluorescent protein-photoprotein pair in clytia hemisphaerica eggs shows co-targeting to mitochondria and efficient bioluminescence energy transfer. *Open Biology*, *4*(APRIL). https://doi.org/10.1098/rsob.130206

Frank, U., Nicotra, M. L., & Schnitzler, C. E. (2020). The colonial cnidarian Hydractinia. *EvoDevo*, *11*(1), 7. https://doi.org/10.1186/s13227-020-00151-0

Freeman, G. (2005). The effect of larval age on developmental changes in the polyp prepattern of a hydrozoan planula. *Zoology*, *108*(1), 55–73. https://doi.org/10.1016/j.zool.2004.11.002

Freeman, G., & Ridgway, E. B. (1988). The role of cAMP in oocyte maturation and the role of the germinal vesicle contents in mediating maturation and subsequent developmental events in hydrozoans. *Roux's Archives of Developmental Biology*, *197*(4), 197–211. https://doi.org/10.1007/BF02439427

García-Castro, H., Kenny, N. J., Iglesias, M., Álvarez-Campos, P., Mason, V., Elek, A., Schönauer, A., Sleight, V. A., Neiro, J., Aboobaker, A., Permanyer, J., Irimia, M., Sebé-Pedrós, A., & Solana, J. (2021). ACME dissociation: A versatile cell fixation-dissociation method for single-cell transcriptomics. *Genome Biology*, *22*(1), 89. https://doi.org/10.1186/s13059-021-02302-5

Gillespie, P. G., & Müller, U. (2009). Mechanotransduction by Hair Cells: Models, Molecules, and Mechanisms. *Cell*, *139*(1), 33–44. https://doi.org/10.1016/j.cell.2009.09.010

Gold, D. A., Katsuki, T., Li, Y., Yan, X., Regulski, M., Ibberson, D., Holstein, T., Steele, R. E., Jacobs, D. K., & Greenspan, R. J. (2019). The genome of the jellyfish Aurelia and the evolution of animal complexity. *Nature Ecology and Evolution*, *3*(1), 96–104. https://doi.org/10.1038/s41559-018-0719-8

Gurska, D., & Garm, A. (2014). Cell Proliferation in Cubozoan Jellyfish Tripedalia cystophora and Alatina moseri. *PLoS ONE*, *9*(7), e102628. https://doi.org/10.1371/journal.pone.0102628

Hand, C., & Uhlinger, K. R. (1992). The Culture, Sexual and Asexual Reproduction, and Growth of the Sea Anemone *Nematostella vectensis*. *The Biological Bulletin*, *182*(2), 169–176. https://doi.org/10.2307/1542110

Hartenstein, V., & Stollewerk, A. (2015). The evolution of early neurogenesis. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2015.02.004

Hie, B., Bryson, B., & Berger, B. (2019). Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. *Nature Biotechnology*, *37*(6), 685–691. https://doi.org/10.1038/s41587-019-0113-3

Hobert, O. (2008). Regulatory logic of neuronal diversity: Terminal selector genes and selector motifs. *Proceedings of the National Academy of Sciences*, *105*(51), 20067–20071. https://doi.org/10.1073/pnas.0806070105

Hobmayer, B., Jenewein, M., Eder, D., Eder, M. K., Glasauer, S., Gufler, S., Hartl, M., & Salvenmoser, A. W. (2012). Stemness in Hydra—A current perspective. *International*

Journal of Developmental Biology, *56*(6–8), 509–517. https://doi.org/10.1387/ijdb.113426bh

Houliston, E., Leclère, L., Munro, C., Copley, R. R., & Momose, T. (2022). Past, present and future of Clytia hemisphaerica as a laboratory jellyfish. In *Current Topics in Developmental Biology* (Vol. 147, pp. 121–151). Academic Press Inc. https://doi.org/10.1016/bs.ctdb.2021.12.014

Houliston, E., Momose, T., & Manuel, M. (2010). Clytia hemisphaerica: A jellyfish cousin joins the laboratory. *Trends in Genetics*, *26*(4), 159–167. https://doi.org/10.1016/j.tig.2010.01.008

Hu, M., Zheng, X., Fan, C.-M., & Zheng, Y. (2020). Lineage dynamics of the endosymbiotic cell type in the soft coral Xenia. *Nature*, *582*(7813), 534–538. https://doi.org/10.1038/s41586-020-2385-7

Hwang, B., Lee, J. H., & Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformatics pipelines. *Experimental and Molecular Medicine*, *50*(8). https://doi.org/10.1038/s12276-018-0071-8

Ikegami, S., Honji, N., & Yoshida, M. (1978). Light-controlled production of spawning-inducing substance in jellyfish ovary. *Nature*, *272*(5654), 611–612. https://doi.org/10.1038/272611a0

Jager, M., Quéinnec, E., Le Guyader, H., & Manuel, M. (2011). Multiple Sox genes are expressed in stem cells or in differentiating neuro-sensory cells in the hydrozoan Clytia hemisphaerica. *EvoDevo*, *2*(1). https://doi.org/10.1186/2041-9139-2-12

Jaitin, D. A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., & Amit, I. (2014). Massively Parallel Single-Cell RNA-Seq for Marker-Free Decomposition of Tissues into Cell Types. *Science*, *343*(6172), 776–779. https://doi.org/10.1126/science.1247651

Kayal, E., Bentlage, B., Sabrina Pankey, M., Ohdera, A. H., Medina, M., Plachetzki, D. C., Collins, A. G., & Ryan, J. F. (2018). Phylogenomics provides a robust topology of the major cnidarian lineages and insights on the origins of key organismal traits. *BMC Evolutionary Biology*, *18*(1). https://doi.org/10.1186/s12862-018-1142-0

Khalturin, K., Shinzato, C., Khalturina, M., Hamada, M., Fujie, M., Koyanagi, R., Kanda, M., Goto, H., Anton-Erxleben, F., Toyokawa, M., Toshino, S., & Satoh, N. (2019). Medusozoan genomes inform the evolution of the jellyfish body plan. *Nature Ecology and Evolution*, *3*(5), 811–822. https://doi.org/10.1038/s41559-019-0853-y

Kim, H. M., Weber, J. A., Lee, N., Park, S. G., Cho, Y. S., Bhak, Y., Lee, N., Jeon, Y., Jeon, S., Luria, V., Karger, A., Kirschner, M. W., Jo, Y. J., Woo, S., Shin, K., Chung, O., Ryu, J. C., Yim, H. S., Lee, J. H., … Yum, S. (2019). The genome of the giant

Nomura's jellyfish sheds light on the early evolution of active predation. *BMC Biology*, *17*(1). https://doi.org/10.1186/s12915-019-0643-7

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

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P. ru, & Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature Methods*, *16*(12), 1289–1296. https://doi.org/10.1038/s41592-019-0619-0

Krasovec, G., Pottin, K., Rosello, M., Quéinnec, É., & Chambon, J. P. (2021). Apoptosis and cell proliferation during metamorphosis of the planula larva of Clytia hemisphaerica (Hydrozoa, Cnidaria). *Developmental Dynamics*, *250*(12), 1739–1758. https://doi.org/10.1002/dvdy.376

Kraus, J. E., Fredman, D., Wang, W., Khalturin, K., & Technau, U. (2015). Adoption of conserved developmental genes in development and origin of the medusa body plan. *EvoDevo*, *6*(1). https://doi.org/10.1186/s13227-015-0017-3

Kraus, Y., Chevalier, S., & Houliston, E. (2020). Cell shape changes during larval body plan development in Clytia hemisphaerica. *Developmental Biology*, *468*(1–2), 59–79. https://doi.org/10.1016/j.ydbio.2020.09.013

Lapébie, P., Ruggiero, A., Barreau, C., Chevalier, S., Chang, P., Dru, P., Houliston, E., & Momose, T. (2014). Differential Responses to Wnt and PCP Disruption Predict Expression and Developmental Function of Conserved and Novel Genes in a Cnidarian. *PLoS Genetics*, *10*(9). https://doi.org/10.1371/journal.pgen.1004590

Lartillot, N. (2020a). *PhyloBayes: Bayesian Phylogenetics Using Site-heterogeneous Models*. 17.

Lartillot, N. (2020b). *The Bayesian Approach to Molecular Phylogeny*. 18.

Layden, M. J., Boekhout, M., & Martindale, M. Q. (2012). *Nematostella vectensis achaete-scute* homolog *NvashA* regulates embryonic ectodermal neurogenesis and represents an ancient component of the metazoan neural specification pathway. *Development*, *139*(5), 1013–1022. https://doi.org/10.1242/dev.073221

Layden, M. J., Johnston, H., Amiel, A. R., Havrilak, J., Steinworth, B., Chock, T., Röttinger, E., & Martindale, M. Q. (2016). MAPK signaling is necessary for neurogenesis in Nematostella vectensis. *BMC Biology*, *14*(1). https://doi.org/10.1186/s12915-016-0282-1

Lechable, M., Jan, A., Duchene, A., Uveira, J., Weissbourd, B., Gissat, L., Collet, S., Gilletta, L., Chevalier, S., Leclère, L., Peron, S., Barreau, C., Lasbleiz, R., Houliston, E., &

Momose, T. (2020). An improved whole life cycle culture protocol for the hydrozoan genetic model Clytia hemisphaerica. *Biology Open*, *9*(11). https://doi.org/10.1242/bio.051268

Leclère, L., Horin, C., Chevalier, S., Lapébie, P., Dru, P., Peron, S., Jager, M., Condamine, T., Pottin, K., Romano, S., Steger, J., Sinigaglia, C., Barreau, C., Quiroga Artigas, G., Ruggiero, A., Fourrage, C., Kraus, J. E. M., Poulain, J., Aury, J. M., … Copley, R. R. (2019). The genome of the jellyfish Clytia hemisphaerica and the evolution of the cnidarian life-cycle. *Nature Ecology and Evolution*, *3*(5), 801–810. https://doi.org/10.1038/s41559-019-0833-2

Leclère, L., Jager, M., Barreau, C., Chang, P., Le Guyader, H., Manuel, M., & Houliston, E. (2012). Maternally localized germ plasm mRNAs and germ cell/stem cell formation in the cnidarian Clytia. *Developmental Biology*. https://doi.org/10.1016/j.ydbio.2012.01.018

Leclère, L., & Röttinger, E. (2017). Diversity of cnidarian muscles: Function, anatomy, development and regeneration. *Frontiers in Cell and Developmental Biology*, *4*(JAN). https://doi.org/10.3389/fcell.2016.00157

Levy, S., Elek, A., Grau-Bové, X., Menéndez-Bravo, S., Iglesias, M., Tanay, A., Mass, T., & Sebé-Pedrós, A. (2021). A stony coral cell atlas illuminates the molecular and cellular basis of coral symbiosis, calcification, and immunity. *Cell*, *184*(11), 2973–2987. https://doi.org/10.1016/j.cell.2021.04.005

Leyva-Díaz, E., Masoudi, N., Serrano-Saiz, E., Glenwinkel, L., & Hobert, O. (2020). Brn3/POU-IV-type POU homeobox genes—Paradigmatic regulators of neuronal identity across phylogeny. *WIREs Developmental Biology*, *9*(4). https://doi.org/10.1002/wdev.374

Lun, A. T. L., McCarthy, D. J., & Marioni, J. C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data [version 1; referees: 5 approved with reservations]. *F1000Research*, *5*. https://doi.org/10.12688/F1000RESEARCH.9501.1

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A., & McCarroll, S. A. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, *161*(5), 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002

Magie, C. R., & Martindale, M. Q. (2008). *Cell-Cell Adhesion in the Cnidaria: Insights Into the Evolution of Tissue Morphogenesis*.

Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, *18*(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Marioni, J. C., & Arendt, D. (2017). How Single-Cell Genomics Is Changing Evolutionary and Developmental Biology. *Annual Review of Cell and Developmental Biology*. https://doi.org/10.1146/annurev-cellbio-100616

Marlow, H. Q., Srivastava, M., Matus, D. Q., Rokhsar, D., & Martindale, M. Q. (2009). Anatomy and development of the nervous system of *Nematostella vectensis* , an anthozoan cnidarian. *Developmental Neurobiology*, *69*(4), 235–254. https://doi.org/10.1002/dneu.20698

Martin, V. J., Littlefield, C. L., Archer, W. E., & Bode, H. R. (1997). Embryogenesis in Hydra. *The Biological Bulletin*, *192*(3), 345–363. https://doi.org/10.2307/1542745

Martindale, M. Q. (2005). The evolution of metazoan axial properties. *Nature Reviews Genetics*, *6*(12), 917–927. https://doi.org/10.1038/nrg1725

Mcconnell, C. H. (1932). Memoirs: The Development of the Ectodermal Nerve Net in the Buds of Hydra. *Journal of Cell Science*, *s2-75*(299), 495–509. https://doi.org/10.1242/jcs.s2-75.299.495

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

Metchnikoff, E. (1886). *EMBRYOLOGISCHE STUDIEN AN MEDUSEN. EIN BEITRAG ZUR GENEALOGIE DER PRIMITIV-ORGANE.*

Miller, D. J., & Ball, E. E. (2008). Cryptic complexity captured: The Nematostella genome reveals its secrets. *Trends in Genetics*, *24*(1), 1–4. https://doi.org/10.1016/j.tig.2007.10.002

Miranda, L. S., Collins, A. G., & Marques, A. C. (2010). Molecules Clarify a Cnidarian Life Cycle – The “Hydrozoan” Microhydrula limopsicola Is an Early Life Stage of the Staurozoan Haliclystus antarcticus. *PLoS ONE*, *5*(4), e10182. https://doi.org/10.1371/journal.pone.0010182

Momose, T., Derelle, R., & Houliston, E. (2008). A maternally localised Wnt ligand required for axial patterning in the cnidarian Clytia hemisphaerica. *Development*, *135*(12), 2105–2113. https://doi.org/10.1242/dev.021543

Momose, T., & Houliston, E. (2007). Two Oppositely Localised Frizzled RNAs as Axis Determinants in a Cnidarian Embryo. *PLoS Biology*, *5*(4), e70. https://doi.org/10.1371/journal.pbio.0050070

Munro, C., Cadis, H., Pagnotta, S., Houliston, E., & Huynh, J.-R. (2022). *Conserved meiotic mechanisms in the cnidarian* Clytia hemisphaerica *revealed by Spo11 knockout* [Preprint]. Developmental Biology. https://doi.org/10.1101/2022.01.05.475076

Nakanishi, N., Renfer, E., Technau, U., & Rentzsch, F. (2012). Nervous systems of the sea anemone *Nematostella vectensis* are generated by ectoderm and endoderm and

shaped by distinct mechanisms. *Development*, *139*(2), 347–357. https://doi.org/10.1242/dev.071902

Nielsen, C. (1998). Origin and evolution of animal life cycles. *Biological Reviews of the Cambridge Philosophical Society*, *73*(2), 125–155. https://doi.org/10.1017/S0006323197005136

Nilsson, D.-E., Gislén, L., Coates, M. M., Skogh, C., & Garm, A. (2005). Advanced optics in a jellyfish eye. *Nature*, *435*(7039), 201–205. https://doi.org/10.1038/nature03484

Nishimiya-Fujisawa, C., & Kobayashi, S. (2012). Germline stem cells and sex determination in Hydra. *The International Journal of Developmental Biology*, *56*(6-7–8), 499–508. https://doi.org/10.1387/ijdb.123509cf

Ohdera, A., Ames, C. L., Dikow, R. B., Kayal, E., Chiodin, M., Busby, B., La, S., Pirro, S., Collins, A. G., Medina, M., & Ryan, J. F. (2019). Box, stalked, and upside-down? Draft genomes from diverse jellyfish (Cnidaria, Acraspeda) lineages: Alatina alata (Cubozoa), Calvadosia cruxmelitensis (Staurozoa), and Cassiopea xamachana (Scyphozoa). *GigaScience*, *8*(7), giz069. https://doi.org/10.1093/gigascience/giz069

Ormestad, M., Martindale, M. Q., & Röttinger, E. (2011). A comparative gene expression database for invertebrates. *EvoDevo*, *2*(1), 17. https://doi.org/10.1186/2041-9139-2-17

Ozment, E., Tamvacakis, A. N., Zhou, J., Rosiles-Loeza, P. Y., Escobar-Hernandez, E. E., Fernandez-Valverde, S. L., & Nakanishi, N. (2021). Cnidarian hair cell development illuminates an ancient role for the class IV POU transcription factor in defining mechanoreceptor identity. *ELife*, *10*, e74336. https://doi.org/10.7554/eLife.74336

Pett, W., Adamski, M., Adamska, M., Francis, W. R., Eitel, M., Pisani, D., & Wörheide, G. (2019). The Role of Homology and Orthology in the Phylogenomic Analysis of Metazoan Gene Content. *Molecular Biology and Evolution*, *36*(4), 643–649. https://doi.org/10.1093/molbev/msz013

Polański, K., Young, M. D., Miao, Z., Meyer, K. B., Teichmann, S. A., & Park, J.-E. (2019). BBKNN: Fast batch alignment of single cell transcriptomes. *Bioinformatics*, btz625. https://doi.org/10.1093/bioinformatics/btz625

Putnam, N. H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., Kapitonov, V. V., Jurka, J., Genikhovich, G., Grigoriev, I. V., Lucas, S. M., Steele, R. E., Finnerty, J. R., Technau, U., Martindale, M. Q., & Rokhsar, D. S. (2007). Sea Anemone Genome Reveals Ancestral Eumetazoan Gene Repertoire and Genomic. *New Series*, *317*(5834), 86–94. https://doi.org/10.1126/science.ll43254

Quiroga Artigas, G., Lapébie, P., Leclère, L., Takeda, N., Deguchi, R., Jekely, G., Momose, T., & Houliston, E. (2018). *A gonad-expressed opsin mediates light-induced spawning in the jellyfish Clytia*. https://doi.org/10.7554/eLife.29555.001

Rentzsch, F., Layden, M., & Manuel, M. (2017). The cellular and molecular basis of cnidarian neurogenesis. *Wiley Interdisciplinary Reviews: Developmental Biology*, *6*(1). https://doi.org/10.1002/wdev.257

Richards, G. S., & Rentzsch, F. (2014). Transgenic analysis of a SoxB gene reveals neural progenitor cells in the cnidarian Nematostella vectensis. *Development (Cambridge)*, *141*(24), 4681–4689. https://doi.org/10.1242/dev.112029

Richards, G. S., & Rentzsch, F. (2015). Regulation of Nematostella neural progenitors by SoxB, Notch and bHLH genes. *Development (Cambridge)*, *142*(19), 3332–3342. https://doi.org/10.1242/dev.123745

Röttinger, E., Dahlin, P., & Martindale, M. Q. (2012). A Framework for the Establishment of a Cnidarian Gene Regulatory Network for “Endomesoderm” Specification: The Inputs of ß-Catenin/TCF Signaling. *PLoS Genetics*, *8*(12), e1003164. https://doi.org/10.1371/journal.pgen.1003164

Röttinger, E., Saudemont, A., Duboc, V., Besnardeau, L., McClay, D., & Lepage, T. (2008). FGF signals guide migration of mesenchymal cells, control skeletal morphogenesis and regulate gastrulation during sea urchin development. *Development*, *135*(4), 785–785. https://doi.org/10.1242/dev.020016

Salvini-Plawen, L. V. (1978). On the origin and evolution of the lower Metazoa. *Journal of Zoological Systematics and Evolutionary Research*, *16*(1), 40–87. https://doi.org/10.1111/j.1439-0469.1978.tb00919.x

Schmid, B., Schmid, V., & Tardent, P. (1974). The umbrellar growth process in the leptomedusaePhialidium hemisphaericum (Syn.Campanularia johnstoni). *Experientia*, *30*(12), 1399–1400. https://doi.org/10.1007/BF01919657

Sebé-Pedrós, A., Ariza-Cosano, A., Weirauch, M. T., Leininger, S., Yang, A., Torruella, G., Adamski, M., Adamska, M., Hughes, T. R., Gómez-Skarmeta, J. L., & Ruiz-Trillo, I. (2013). Early evolution of the T-box transcription factor family. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(40), 16050–16055. https://doi.org/10.1073/pnas.1309748110

Sebé-Pedrós, A., Chomsky, E., Pang, K., Lara-Astiaso, D., Gaiti, F., Mukamel, Z., Amit, I., Hejnol, A., Degnan, B. M., & Tanay, A. (2018). Early metazoan cell type diversity and the evolution of multicellular gene regulation. *Nature Ecology and Evolution*, *2*(7), 1176–1188. https://doi.org/10.1038/s41559-018-0575-6

Sebé-Pedrós, A., Saudemont, B., Chomsky, E., Plessier, F., Mailhé, M. P., Renno, J., Loe-Mie, Y., Lifshitz, A., Mukamel, Z., Schmutz, S., Novault, S., Steinmetz, P. R. H.,

Spitz, F., Tanay, A., & Marlow, H. (2018). Cnidarian Cell Type Diversity and Regulation Revealed by Whole-Organism Single-Cell RNA-Seq. *Cell*, *173*(6), 1520-1534.e20. https://doi.org/10.1016/j.cell.2018.05.019

Serrano-Saiz, E., Leyva-Díaz, E., De La Cruz, E., & Hobert, O. (2018). BRN3-type POU Homeobox Genes Maintain the Identity of Mature Postmitotic Neurons in Nematodes and Mice. *Current Biology*, *28*(17), 2813-2823.e2. https://doi.org/10.1016/j.cub.2018.06.045

Shapiro, E., Biezuner, T., & Linnarsson, S. (2013). Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nature Reviews Genetics*, *14*(9), 618–630. https://doi.org/10.1038/nrg3542

Sheng, M., & Greenberg, M. E. (1990). The regulation and function of c-fos and other immediate early genes in the nervous system. *Neuron*, *4*(4), 477–485. https://doi.org/10.1016/0896-6273(90)90106-P

Shook, D., & Keller, R. (2003). Mechanisms, mechanics and function of epithelial–mesenchymal transitions in early development. *Mechanisms of Development*, *120*(11), 1351–1383. https://doi.org/10.1016/j.mod.2003.06.005

Siebert, S., Farrell, J. A., Cazet, J. F., Abeykoon, Y., Primack, A. S., Schnitzler, C. E., & Juliano, C. E. (2019). Stem cell differentiation trajectories in Hydra resolved at single-cell resolution. *Science*, *365*(6451). https://doi.org/10.1126/science.aav9314

Singla, C. L. (1975). Statocysts of Hydromedusae. In *Cell Tiss. Res* (Vol. 158, pp. 391–407). Springer-Verlag.

Sinigaglia, C., Busengdal, H., Lerner, A., Oliveri, P., & Rentzsch, F. (2015). Molecular characterization of the apical organ of the anthozoan Nematostella vectensis. *Developmental Biology*, *398*(1), 120–133. https://doi.org/10.1016/j.ydbio.2014.11.019

Sinigaglia, C., Peron, S., Eichelbrenner, J., Chevalier, S., Steger, J., Barreau, C., Houliston, E., & Leclère, L. (2020). Pattern regulation in a regenerating jellyfish. *ELife*, *9*, e54868. https://doi.org/10.7554/eLife.54868

Steger, J., Cole, A. G., Denner, A., Lebedeva, T., Genikhovich, G., Ries, A., Reischl, R., Taudes, E., Lassnig, M., & Technau, U. (2022). *Single cell transcriptomics identifies conserved regulators of neurosecretory lineages* [Preprint]. Developmental Biology. https://doi.org/10.1101/2022.05.11.491463

Steinmetz, P. R. H. (2019). A non-bilaterian perspective on the development and evolution of animal digestive systems. *Cell and Tissue Research*, *377*(3), 321–339. https://doi.org/10.1007/s00441-019-03075-x

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P., & Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell*, *177*(7), 1888-1902.e21. https://doi.org/10.1016/j.cell.2019.05.031

Suga, H., Tschopp, P., Graziussi, D. F., Stierwald, M., Schmid, V., & Gehring, W. J. (2010). Flexibly deployed *Pax* genes in eye development at the early evolution of animals demonstrated by studies on a hydrozoan jellyfish. *Proceedings of the National Academy of Sciences*, *107*(32), 14263–14268. https://doi.org/10.1073/pnas.1008389107

Takeda, N., Kon, Y., Artigas, G. Q., Lapébie, P., Barreau, C., Koizumi, O., Kishimoto, T., Tachibana, K., Houliston, E., & Deguchi, R. (2018). Identification of jellyfish neuropeptides that act directly as oocyte maturation inducing hormones. *Development*, dev.156786. https://doi.org/10.1242/dev.156786

Tardent, P. (1995). The cnidarian cnidocyte, a hightech cellular weaponry. *BioEssays*, *17*(4), 351–362. https://doi.org/10.1002/bies.950170411

Technau, U., & Steele, R. E. (2012). Evolutionary crossroads in developmental biology: Cnidaria. *Development*, *139*(23), 4491–4491. https://doi.org/10.1242/dev.090472

Thomas, M. B., Freeman, G., & Martin, V. J. (1987). The embryonic origin of neurosensory cells and the role of nerve cells in metamorphosis in Phialidium gregarium (Cnidaria, Hydrozoa). *International Journal of Invertebrate Reproduction and Development*.

Tournière, O., Dolan, D., Richards, G. S., Sunagar, K., Columbus-Shenkar, Y. Y., Moran, Y., & Rentzsch, F. (2020). NvPOU4/Brain3 Functions as a Terminal Selector Gene in the Nervous System of the Cnidarian Nematostella vectensis. *Cell Reports*, *30*(13), 4473-4489.e5. https://doi.org/10.1016/j.celrep.2020.03.031

Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, *9*(1), 5233. https://doi.org/10.1038/s41598-019-41695-z

Trembley, A., Bayer, F. M., Lyonet, P., Pronk, C., & Schley, J. van der. (1744). *Mémoires pour servir à l'histoire d'un genre de polypes d'eau douce, à bras en forme de cornes /*. Chez Jean & Herman Verbeek,. https://doi.org/10.5962/bhl.title.64073

Tucker, R. P. (2010). Expression of Usherin in the Expression of Usherin in the Anthozoan Nematostella vectensis. In *Source: Biological Bulletin* (Vol. 218, Issue 2, pp. 105–112).

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

Weissbourd, B., Momose, T., Nair, A., Kennedy, A., Hunt, B., & Anderson, D. J. (2021). A genetically tractable jellyfish model for systems and evolutionary neuroscience. *Cell*, *184*(24), 5854-5868.e20. https://doi.org/10.1016/j.cell.2021.10.021

Wikramanayake, A. H., Hong, M., Lee, P. N., Pang, K., Byrum, C. A., Bince, J. M., Xu, R., & Martindale, M. Q. (2003). *An ancient role for nuclear b-catenin in the evolution of axial polarity and germ layer segregation*. *426*, 5.

Wolf, F. A., Angerer, P., & Theis, F. J. (2018). SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biology*, *19*(1), 15. https://doi.org/10.1186/s13059-017-1382-0

Wolf, F. A., Hamey, F. K., Plass, M., Solana, J., Dahlin, J. S., Göttgens, B., Rajewsky, N., Simon, L., & Theis, F. J. (2019). PAGA: Graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biology*, *20*(1), 59. https://doi.org/10.1186/s13059-019-1663-x

Zapata, F., Goetz, F. E., Smith, S. A., Howison, M., Siebert, S., Church, S. H., Sanders, S. M., Ames, C. L., McFadden, C. S., France, S. C., Daly, M., Collins, A. G., Haddock, S. H. D., Dunn, C. W., & Cartwright, P. (2015). Phylogenomic analyses support traditional relationships within Cnidaria. *PLoS ONE*, *10*(10). https://doi.org/10.1371/journal.pone.0139068