# A cell-type atlas from a scyphozoan jellyfish *Aurelia coerulea* (formerly sp.1) provides insights into changes of cell-type diversity in the transition from polyps to medusae

Oliver Link[1], Stefan M. Jahnel[2], Kristin Janicek[1], Johanna Kraus[1], Juan Daniel Montenegro[1], Bob Zimmerman[1], Alison G. Cole[1]*, and Ulrich Technau[1,3]*

[1] Department of Neurosciences and Developmental Biology, Faculty of Life Sciences, University of Vienna, Djerassiplatz 1, 1030 Vienna.
[2] Institute of Molecular Biotechnology, Dr.-Bohr-Gasse 3, 1030 Vienna.
[3] Research platform Single Cell Regulation of Stem Cells, University of Vienna, Djerassiplatz 1, 1030 Vienna.

*corresponding authors: alison.cole@univie.ac.at and Ulrich.technau@univie.ac.at

**Keywords:**
Aurelia, Scyphozoa, strobilation, single cell RNAseq, developmental cell atlas, life cycle, jellyfish

## Abstract

We provide here a detailed single cell transcriptomic atlas covering the life cycle of the cosmopolitan scyphozoan *Aurelia coerulea*. We show that there is an increase in cell type diversity in the medusa stage, which is reflected by an increase in the number of unique transcripts expressed. We highlight parallels in both cell complement and specification pathways between *Aurelia* and the sea anemone *Nematostella*, two lineages separated by more than 500 MY. We find that several cell types of the neuroglandular lineages are specified by the same transcription factors that arose by gene duplications in the common ancestor of medusozoans and anthozoans. This confirms gene duplications as a source for the increase of cellular complexity during the evolution of cnidarian neuroglandular lineages dates to the common ancestor of all cnidarians. We validated spatial expression domains corresponding to molecular profiles of smooth and striated muscles and followed the development of the striated muscle in the medusa. Our data suggests that striated muscles in the subumbrella of the medusa derive directly from smooth muscles, and yet these two muscle cell phenotypes are characterized by distinct molecular profiles. Further, reconstruction of putative cell trajectories indicates multipotent cell states derive from tissues at the boundary between epithelial layers and highlight molecular characteristics not typical of other cnidarians.

## Teaser

Single cell transcriptomic atlas across the jellyfish life cycle reveals increased cellular complexity associated with the medusa form.

## Introduction

Cell-type diversification is a hallmark in the evolution of metazoans, and the principles that underlie these processes are still barely understood. Advances in single cell sequencing allow the cataloguing of genome-wide differential transcriptome usage across different cell types from entire organisms (reviewed in (1,2)). These collections offer an unprecedented quantity of data with which to compare gene usage across cell types in different phylogenetic lineages. Comparative methods for such data across phylogenetic distances remain a non-trivial task, but robust phylogenetic sampling will be required to make any clear evolutionary inferences from cross-organism comparisons. To understand the evolution of cell-type diversification that coincided with the emergence of complex body plans in the animal kingdom, members of basal branching animal phylum Cnidaria hold the key to reconstructing ancestral cell-type states present during the explosion of animal morphotypes that remain in modern day. Cnidarians have been shown to constitute the phylogenetic sister group to Bilateria (3–5), which comprise the vast majority of all animal phyla. In this regard understanding the phylum Cnidaria is key for the identification of key bilaterian traits. Moreover, the high plasticity and regeneration capacities of cnidarians have attracted scientists already 270 years ago and have had a deep impact on both developmental and cell-biology (6).

The phylum Cnidaria is composed of the subphylum Anthozoa, which mostly encompasses sea anemones and corals, and its sister group, the Medusozoa. Even though anthozoan sea anemones and corals evolved a large diversity of body shapes, their life cycle is biphasic: they are restricted to sessile polyp forms, with dispersion accomplished by broadcast or local spawning and the development of a short-lived planktonic larva, which eventually settles and metamorphoses into a polyp. Medusazoa differ from the anthozoans in that these lineages, with the exception of Staurozoa (7), possess a triphasic life cycle that includes a free-swimming larva, sessile polyp, and a free-swimming jellyfish stage in their life-histories(8,9). Sexual maturity occurs in the free-swimming medusa, although there are exceptions in which the medusa stage is lost, and sexual maturity occurs already in polyps(10). The most prominent example of this is the freshwater polyp *Hydra* (11). Notwithstanding that jellyfish are widely seen as plagues in marine ecosystems (12), their genomes turned out to be treasure vaults for geneticists and evolutionary biologists (11,13–15) and their life-histories are quite remarkable and flexible as some scyphozoan species completely lack a sessile stage in their life-cycle (16) or some were even noticed to reverse their ontogeny, which is an astonishing feature and very unique in the animal kingdom (17,18).

The drastic metagenesis from the polyp into free-swimming jellyfish in different medusozoan clades provides opportunities to take a deeper look into underlying evolutionary principles for

the development of new cell types which accompany the switch from a sedentary into a free-living lifestyle. One of these novel cellular traits are striated muscles that morphologically in cnidarians have only been identified in the medusozoan jellyfish (17). Additionally, a complex sensory organ, the rhopalia, and a specific set of motor neurons comprising the motor nerve net (MNN) develop during the polyp-medusa metamorphosis (17–19). Jellyfish also have gravity-sensing organs (statocysts), making them attractive also for studies in weightlessness, hence jellyfish were the first model used for studies in zero gravity on the SLS-1 mission (20). In scyphozoans, like *Aurelia* sp.1 metagenesis occurs via polydisc strobilation, an asexual horizontal fission process along the oral-aboral axis of the polyp which enables the organism to produce a large number of free-swimming individuals at once (21). Strobilation can be induced by cold temperature or by peptides activated by the retinoic acid pathway (22). However, we have very little information about the molecular and cellular changes that precede and accompany this drastic change of life form. Furthermore, since medusa formation occurs by lateral budding in hydrozoans, hence a very different process, it is still an open question, whether medusozoan jellyfish are homologous.

In this study, we provide a molecular foundation for investigating mechanisms of cell type diversification in Scyphozoa by generating a cell-type atlas of a cosmopolitan scyphozoan model *Aurelia coerulea* (formerly sp.1). We used single-cell RNA sequencing to produce whole animal, as well as tissue specific, single-cell RNAseq libraries that follow *Aurelia* life cycle stages with an emphasis on the metagenetic transition from polyp (also called scyphistoma) and medusa. We assess gene usage across the life cycle and characterise the composition of identified cell type families. We found remarkable parallels with anthozoan cellular complexity with regards to neural composition and cnidocyte specification pathways (23,24), and we identify cell types that define the jellyfish as free-swimming individuals, including various neuronal populations and striated muscles (17,25). These muscles enable jellyfish to propel through the water column and allow for a highly efficient swimming ability (26). We characterise the molecular signatures of the striated jellyfish muscles to identify genes putatively involved in the assembly and maintenance of this phenotype. Our data suggest a shared contractile unit and a possible transformation from smooth to striated fibre types. While the striated muscle appears to derive from pre-existing smooth muscle in the polyp, other cell types such as the stinging cells (cnidocytes), neurons, and various gland cells, derive from cells expressing a molecularly distinct transcriptomic profile, primarily associated with the inner cell layer of the polyp, and the manubrial arms of the jellyfish. We suggest these cells are analogous to the i-cells of hydrozoans and the putative stem cells in anthozoans.

# Results

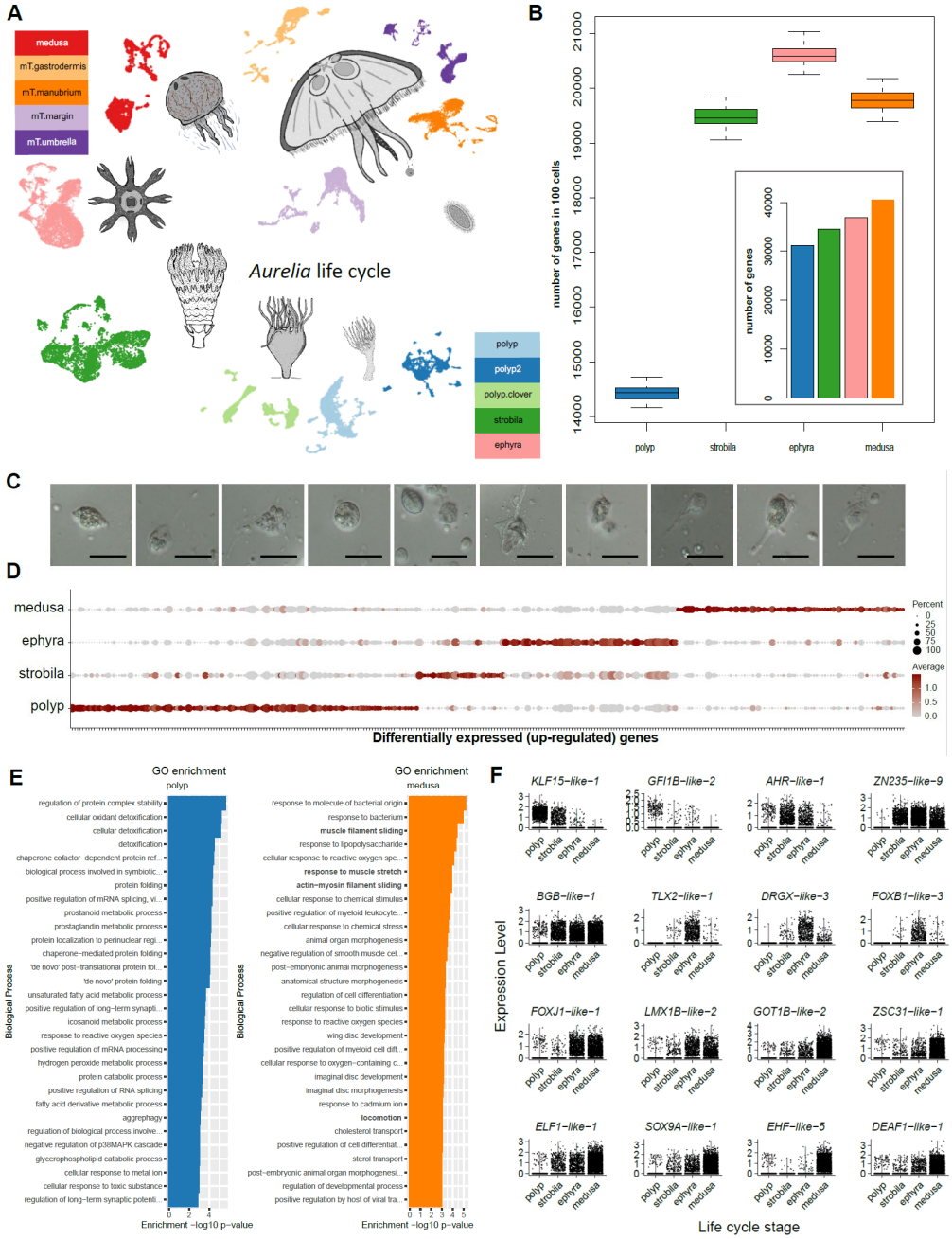



**Fig. 1 Single cell transcriptomic atlas covering the polyp-medusa life cycle transition demonstrates shifts in genome usage between key life history stages** (**A**) Ten single cell transcriptomic libraries were generated from cell dissociations of sequential life history stages (6) or tissue compartments from a sub-adult medusa (4). (**B**) Distribution of genome use across the life cycle. Boxplot representing 100 replicate sampling of 1000 cells over all expressed (min. 10 reads) genes from each stage. inset: total gene models recovered across sequencing libraries. (**C**) Random collection of diverse cell morphologies from whole animal tissue dissociations, representing partial diversity of cells captured from cell suspensions. Scale bars = 10µm. (**D**) Differentially expressed genes across the dataset. (**E**) Over-represented GO-terms represented in the up-regulated gene set highlights prostaglandin synthesis in polyps and over abundance of muscle in the medusa. (**F**) Stage-specific up-regulated transcription factors show differential activity across the life history stages.

The typical life cycle of medusozoan cnidarians involves the transition from a sessile polyp (scyphistoma) to the sexually active jellyfish (medusa). This not only is a dramatic reorganisation of the body plan, it may also be accompanied by the emergence of new cell types and potentially the disappearance of polyp-specific cell types. In order to unravel the cellular basis for this metagenesis, we generated ten single cell RNA seq libraries using the 10x Genomics platform: six from different stages of the life cycle of *Aurelia coerulea* including two replicates of the sessile polyp, a pre-strobilating polyp, strobilating polyp, newly liberated ephyra juvenile, and the newly metamorphosed medusa (1 month, ~1.5cm diameter); four tissue libraries from growing medusa (~8cm diameter) including the umbrella (excluding the margin), the margin (including tentacles), the manubrium (including the arms and mouth), and the gastrodermis (including future gonad) (**Fig. 1A**). Each library was sequenced to at least 70% saturation. We generated a custom mapping tool by combining an in-house transcriptome assembly with the genomic annotations from the available *Aurelia* genome showing the best alignment ((27): see methods). We detect 42,752 gene models, of which 32,200 are represented by 10 or more unique reads across the entire dataset.

## The transition from the sessile polyp to the free-swimming medusa is marked by an increase in expressed gene repertoire.

We first asked whether we could identify any gene sets specific for each life-cycle stage specific gene usage by comparing all samples against one another. To accomplish this, we sampled the same randomly selected number of cells from stage-specific samples (1000 each) from the set of genes with 10 or more reads across the entire dataset (100 replicates) and compared their proportional representation in the genome. Gene counts (at least 3 reads/ stage) were consistently lower in polyp samples with respect to the other life cycle stages, indicating that a portion of the coding genome is activated only in the medusa (**Fig. 1B**). We then queried the identity of the medusa-specific gene sets and found an enrichment in signalling and ion channels (**Fig. S1A, Data S1.1**). This suggests an increase in neuronal diversity within the medusa compared to the sessile polyp stage. We also examined differences in over-abundance of gene expression across the life cycle, and found that the polyp gene complement is enriched for prostaglandin synthesis (**Fig. 1C,D and Data S1.2**). Interestingly, it has been demonstrated that blocking prostaglandin signalling with indomethacin can induce strobilation in this species (28). In contrast, the medusa stage, with its large muscular swimming bell, is enriched in collagens and muscle-related biological processes (**Fig. 1C,D, Data S1.2**). We also note an upregulation of genes involved in mitosis and signalling pathway receptors within the strobila, consistent with the tissue reorganisation that is taking place at this stage (**Fig. 1C,D, S2, Data S1.2**). We assayed the representation of transcription factor families that are up regulated at each stage compared to the entire

dataset. We find several transcription factors that show stage-specific regulation (**Fig. 1E**). These genes are up-regulated in a life-cycle specific manner across all tissues (**Fig. S1B,C, Data S1.3**). Regulatory genes up-regulated in the strobila include those of the HGM (high mobility group) and bHLH (basic helix-loop-helix) protein families, while the homeodomain (HD) genes are then increased in the ephyra. These data suggest that cell type specification, which is at its maximum during the transition between the polyp and the medusa, is driven largely by bHLH, HMG, and Homeodomain proteins, like what has been demonstrated in bilaterians. To confirm gene identities, we constructed gene trees for these superfamilies from our transcriptome including the latest transcriptome for the anthozoan *Nematostella vectensis* (29) and bilaterian representatives (**Fig. S3**). We next sought to catalogue the cell diversity that is present at each point of the life cycle.

## Transition to free-swimming medusa is accompanied by increased cell-type complement.

All ten datasets were integrated using a reciprocal PCA anchor approach (30). This integration approach faithfully clusters cells of the same cell type as well as integrating the tissue layers across life cycle stages (**Fig. 2A,B, S4:S6**). Within the dataset, we identify 9 broadly defined cell populations, for which we assign identities by assessing up-regulated gene lists (**Fig. 2C, S7, Sup. Data S1.4**). Principal cell populations include cnidocytes (pink), two classes of neurons (orange, green), digestive (purple) and mucin (red) glands, striated muscle (dark blue), two large populations that represent the inner (yellow) and outer (light blue) cell layers, and an intermediate central population (grey) (**Fig. 2A**). This latter partition includes *foxA* positive tissue of the manubrium, *twist*-positive cells that line the boundary between the two layers (31), as well as contributions from all other libraries (**Fig. S5, 6B,C**). To obtain a full characterization of cell type dynamics in this system, we investigated the substructure of each population in terms of cell-type distribution, up-regulated genes and over-represented GO terms (**Fig. S8-S23, Data S1.5:20**). We then collated the set of cell-type specific transcription factors for all identified subtypes from each population (**Fig. 2D**). In total we find 12 cell states that are not represented in the polyp scRNA sequencing libraries (yellow highlight; **Fig. 2B)**.

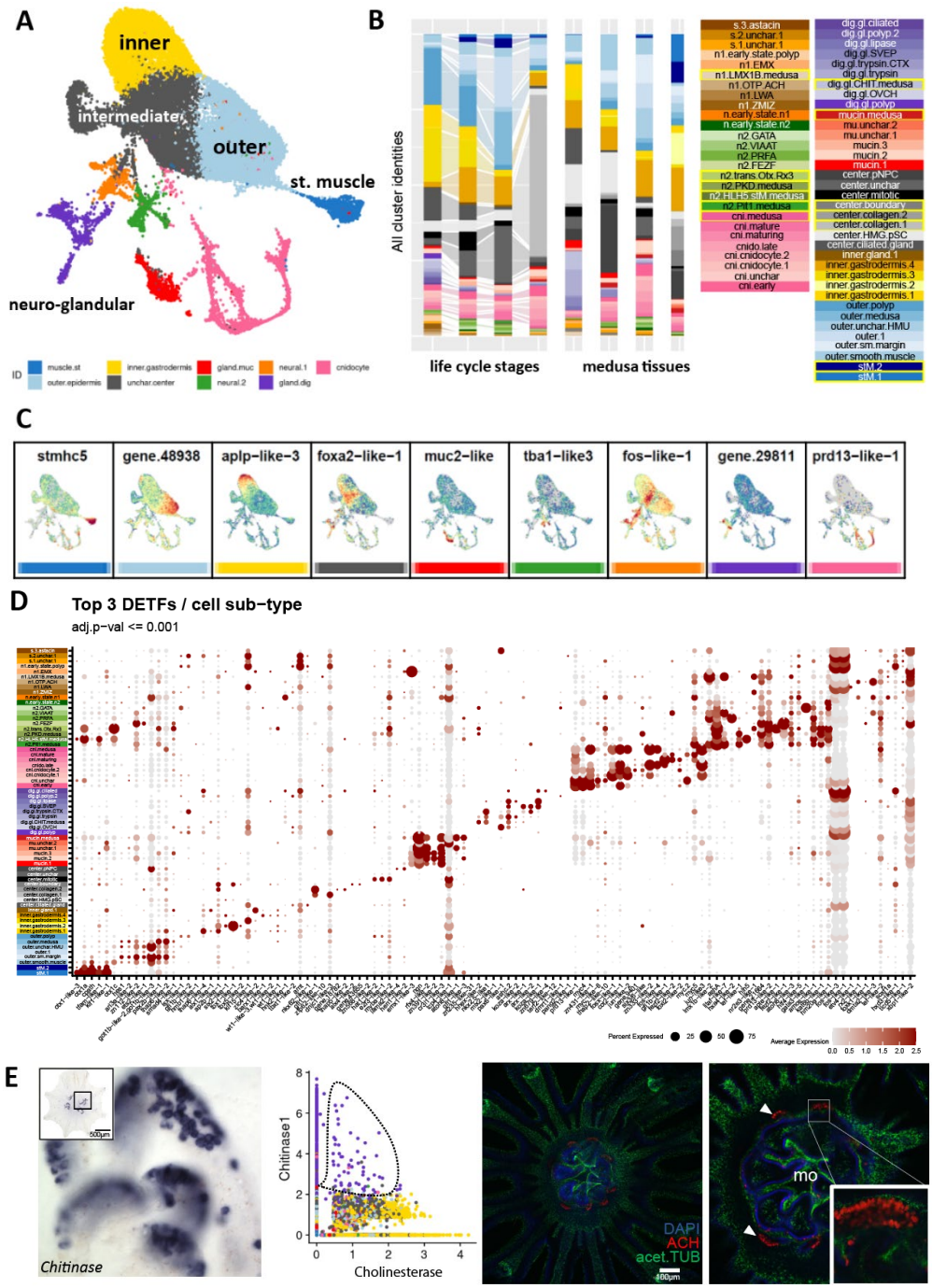



**Fig. 2 Integrated cell atlas permits identification of cell states across the life cycle.** (**A**) Distribution of cell sub-types across all life cycle stages as a fraction of all cells captured (upper) and plotted as a 2d UMAP with cells coloured according to their population identity (lower). (**B**) Relative numbers of differentiated neuro-glandular cell types across each life history stage and medusa tissue. Medusa-specific cell states are highlighted in yellow. (**C**) Expression of population-specific genes indicative of each major partition. (**D**) Differentially expressed transcription factors across all cell states in the atlas. (**E**) A subset of *chitinase*-positive cells co-express Cholinesterase. *left: Chitinase* positive cells are found within the gastric filaments of the medusa stage. *center:* Cells within the digestive gland cluster (purple) co-express both genes. *right:* Acetylcholine positive immunostaining (red) colocalizes within the gastric filaments surrounding the mouth (mo) of a developing ephyra.

Overall, the relative proportion of cells identified as the inner and outer epithelia remains constant across the life cycle. Similar proportions of mitotically active cells are present across all stages and tissue locations, with a slight increase during strobilation, and are generally associated with the outer epithelia and the intermediate cluster (**Fig2B:** black). The polyp stage is characterised by having a larger set of digestive gland cell types (**Fig2A,B** shades of purple), which become localised to the gastric filaments (**Fig. 2E**), manubrial arms, or mantle margin in the medusa (**Fig. S14**). There are no digestive gland cells captured from the umbrella sample, and the relative numbers of digestive gland cells captured during the non-feeding transition stages is reduced. Digestive gland cell types are united by expression of two *achaete-scuteC* (*ashC)* orthologs (**Fig. 2D:** purples) but have differing protease profiles (**Fig. S15C**). Cell types that are expected to have high turnover rates such as the stinging cells, or cnidocytes, show a clear set of specific transcription factor expression (**Fig. 2D**; pinks). All states of the cnidocytes are present across the life cycle, with the exception of a small medusa-specific population that is restricted to the margin and subumbrella (**Fig. 3A:** cni.medusa). Similarly, the mucin-producing population also contains two medusa-specific subtypes. All mucin subtypes (**Fig2:** reds) express a *neurogenin* (*ngn)*-ortholog (**Fig.2D**; reds), which is characteristic for neuronal cell populations in other systems ((23,32–34)). One of the medusa mucin-producing sub-types (mu.medusa) shows over-abundant expression of genes involved in the Notch signalling pathway, and is enriched for biological processes related to hair cells (**Fig. S17; Data S1**). These features suggest that this cell type may be neuronal in identity, rather than strictly glandular. Additional medusa-specific cell types include neuronal sub-types and the striated muscle. We explore the cnidocytes, neurons, and striated muscle populations further below.

## Cnidocytes highlight conserved features of cnidogenesis between anthozoans and scyphozoans

Recently we reported that within *Nematostella*, specification of the distinct cnidocyte types is marked by a diverging transcriptomic profile corresponding to the formation of the different capsule types, which then switch to converges upon a neural-like expression profile, demarcated by up-regulation of GFI1B (23). Notably, we find a similar forked trajectory within the cnidocyte population of *Aurelia*. (**Fig. 3A**). Within our dissociated tissue samples, we consistently find two distinct cnidocyte phenotypes that differ greatly in size (**Fig. S17C**). We assume these two recovered trajectories correspond to these large and small varieties, but validation of this assumption is lacking. One sub-lineage is under-represented in the library derived from the central gastrodermis-enriched territory. We identify a temporal sequence of transcription factor expression that follows the specification trajectories of these cnidocyte

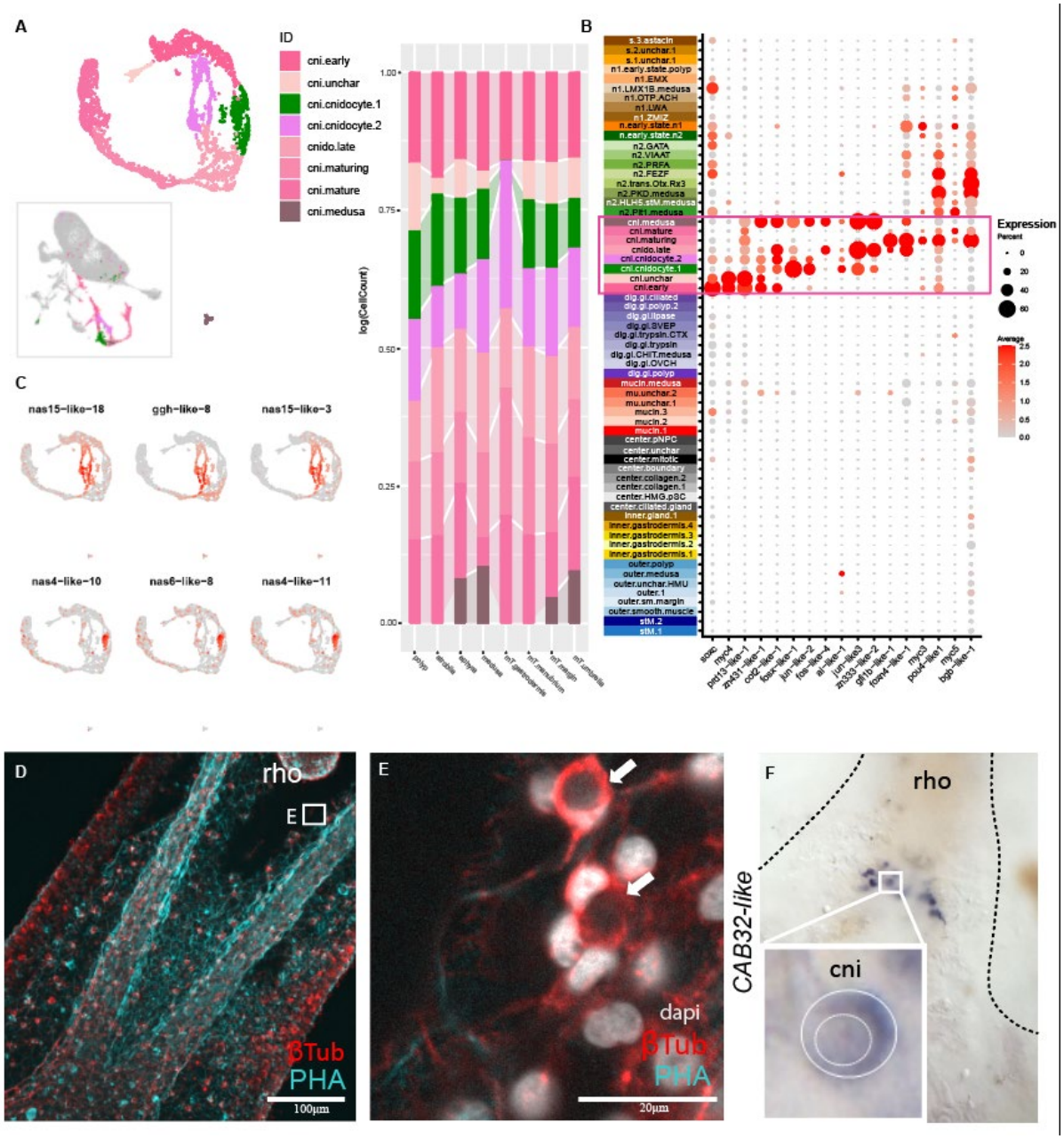



**Fig. 3. Cnidocyte specification pathways indicate two principle capsule types.** (**A**) Cnidocyte trajectory and distribution. Inset: location of subset within the full dataset. (**B**) Dotplot showing sequential transcription factor activation along the trajectory. (**C**) The two principle sub-types have distinct toxin profiles. (**D**) Overview of the ephyra lappet stained with Phalloidin and a βTubulin antibody. (**E**) A close-up shows cnidocytes (arrows) labelled by the same βTubulin antibody located in the distal region of the lappet. (**F**) An in situ hybridization for the cnidocyte marker *CAB32-like* visualises cnidocytes at the corresponding region in the ephyra lappet.

subtypes (**Fig. 3B**). Like what has been described for the anthozoan *Nematostella*, we find the earliest cell states express a high mobility group (HMG) protein (*soxC*), a myc ortholog (*myc4*), prdm13 ortholog (*prdm13a*), and a C2H2-type zinc finger protein (ZN431-like-1). Similar to *Nematostella*, the end of the separate specification trajectories is marked by activation of JUN/FOS orthologs, and expression of a *gfi1b* zinc finger ortholog marks the

switch to a more neural-like transcriptomic profile and activation of additional myc orthologs. The two principal specification trajectories also have distinct toxin profiles (**Fig. 3C**), and one trajectory expresses *arx*, also paralleling what is described in the anthozoan *Nematostella*. Additionally, the terminal mature transcriptomic signature in *Aurelia* shares expression profiles with some class 2 neurons, including *pou4* and *bgb* (**Fig. 3B)**.

## Aurelia neural complement reveals two neural classes with similarities to anthozoan neurons.

We analysed both neural populations together, to account for any unstable cluster boundaries between the two populations. Within the combined neural population we anyhow distinguish two principle sub-populations of neurons, classified here as class 1 and class 2 subtypes (**Fig.4A,B**). Class 1 neurons (orange) are insulinoma (*ins1*) positive (**Fig. 4C**). This population also includes putatively non-neural secretory cell types ("s"), which are enriched in biological functions associated with digestion (**Suppl. Fig. S19 and Data S1**). These data suggest a close relationship between neurons and glands, similar to that which has been suggested in other cnidarians (35). In *Aurelia*, class 1 neurons predominate in the polyp stage (**Fig4B:** shades of brown/orange), and includes only one subtype that is medusa-specific (n1.LMX1B). The candidate specification factor for the n1.LMX1B neuron is the LIM homeobox *lmx1b*. This homeodomain gene is also expressed in some class 2 neurons, but in this case occurs in the absence of *ashA* and *pou4* expression characteristic of class 2 neurons (**Fig.4C**, **Suppl. Data S1**). Class 2 neurons (greens) all express the bHLH transcription factor *ashA* (**Fig. 4C,D, S3A**), and elevated levels of specific alpha- and beta- tubulins (TBB-like-1&2 and TBA1-like3; **Fig. 4D**). Anti-tubulin antibody staining highlights concentrations of the motor neural network centrally located within the lappets of the ephyra (**Fig. 4E**), where *in situ* hybridization detects mRNA expression of *ashA* (**Fig. 4K**). Morphologically, this is the territory that includes the radial canals and overlying muscle fields of the ephyra. FMRFamide and b-Tubulin antibodies stain independent portions of the medusa nerve net (**Fig. 4F**). Whereas b-Tubulin positive cells are found near the striated muscle field and the antibody also labels ganglion cells projecting out of the rhopalia, FMRFamide positive neurons are more scattered across the ectoderm (**Fig. 4E:H**). We find a small number of double positive cells located at the periphery of the striated muscle fields. Within the polyp, only the FMRFamide class 2 neurons are abundant (n2.PRFA), and all other class 2 neural types are under-represented or absent (**Fig. 4B**).

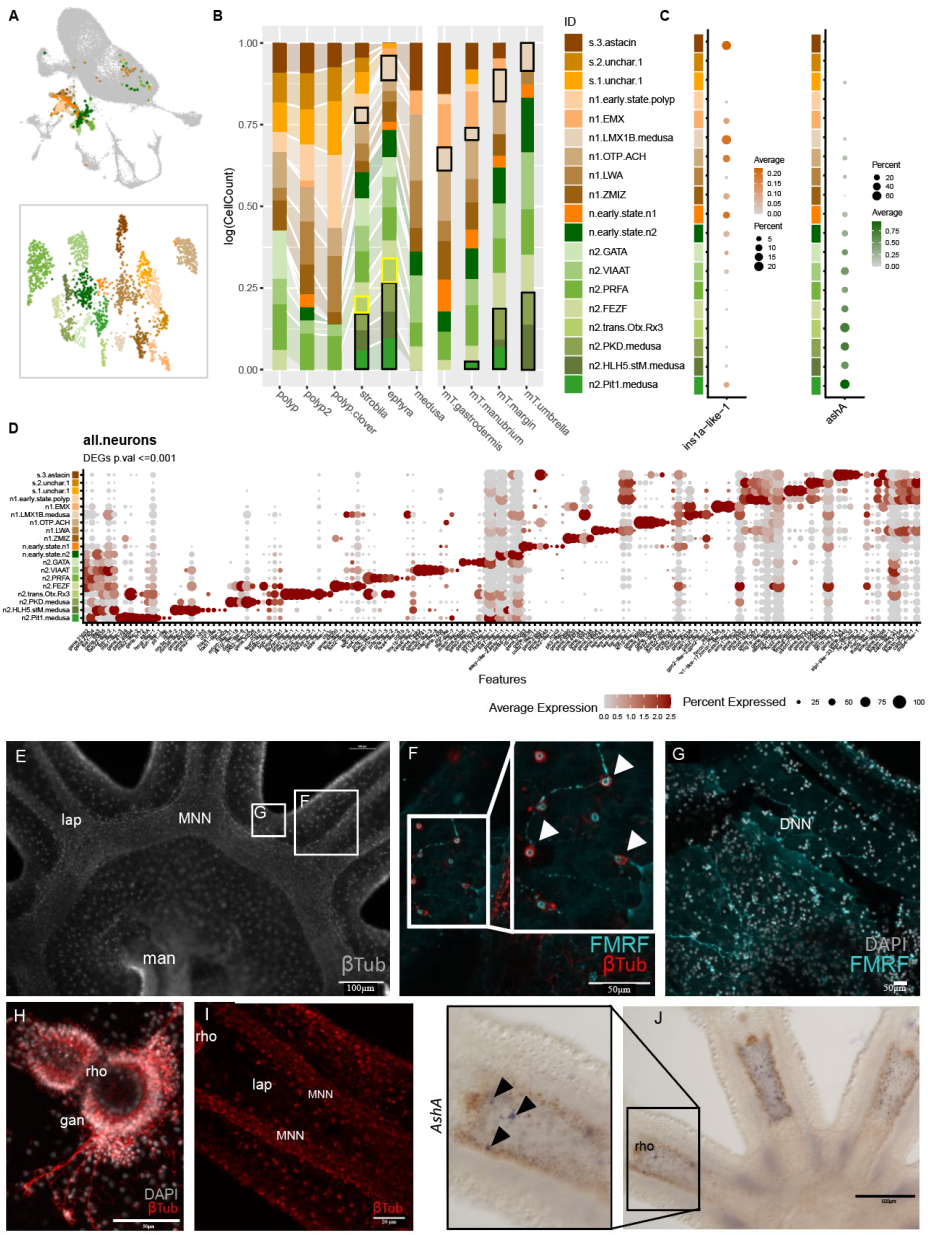



**Fig. 4 Neural cell complement includes two distinct classes, with medusa-specific subtypes in each.** (**A**) Distribution and characterization of neural complement, illustrated on an UMAP embedding of the full data (top) and neural subset (bottom) (**B**) Distribution histogram across all samples, medusa-specific cell clusters are highlighted with black outline. One transient cell state is present only in the metamorphic strobila/ephyra (yellow outline). (**C**) Two classes can be recognized by differential expression of transcription factors. Class 1 cells express *insulinoma* (*ins1a*) (orange left) and class 2 express *achaete-scute* (*ashA*) (green right). (**D**) Upregulated gene sets indicate both common class-markers and highlight unique neural sub-types. Class 2 neurons share expression of specific tubulins and calcium-binding genes whereas class 1 neurons share expression of sequestosome-like orthologs. (**E**) An anti-βTubulin antibody stains the motor nerve net of the ephyra. (**F**) Co-staining of anti-βTubulin and anti FMRFamide visualises cells that are positive for both antibodies. (**G**) FMRFamide positive cells are scattered at the periphery of the ephyra lappet. (**H**) Network of the rhopalia nervous system positive for βTubulin. (**I**) An overview of the ephyra lappet displays condensed regions of the βTubulin positive motor nerve net aligning with the radial canal of the lappet. (**J**) Single cells are stained with an anti-*ashA* in situ probe in proximity of the radial canal within the ephyra lappet. Lap = lappet, MNN = motor nerve net, DNN= diffuse nerve net, gan = ganglion, rho = rhopalia.

*Ins1*-negative class 2 neurons express a *pou4/brn3* ortholog, which is shared also with the cnidocytes (**Fig3C**). Class 2 neurons (**Fig4A,** shades of green) increase in abundance during the transition to the medusa form, including the generation of three additional medusa-specific class 2 neurons. We presume that these correspond to sensory cells of the rhopalium, and putative motor neurons that innervate the striated swimming musculature, as these tissues are medusa-specific. During the period of rhopalia development in the ephyra, the population "n2.pit1.medusa" appears. Candidate specification markers include the POU-family *pit1*, bHLH-family *ptf1a*, and a distaless ortholog *dlxb.3*. *Pit1* has previously been reported to be expressed in putative sensory neurons of the developing rhopalia (19). Also appearing at this time is another putative sensory neuron (n2.PKD), identifiable by enrichment of polycystin orthologs and *otx1d* expression (**Fig. 6D**). We also detect a possibly transient, perhaps very rare neuronal subtype n2.otx.Rx3 that is detected only in strobila and ephyra, but not in the medusa. This population is identifiable by *otx1a* and *otx1b* expression and appears to be an FMRFamide positive cell-type. This could correspond to a set of FMRFamide positive sensory cells in the oral-distal portion of the ephyra lappets (19). One of the medusa-specific neurons, n2.HLH5, also expresses low levels of the striated muscle signature. This could be explained by either a common progenitor of these two cell types, or by cross contamination of the transcriptomic signature by inefficient separation of the striated muscle from its motor neuron innervation. The candidate specification marker for this neuron is an atonal-like bHLH family transcription factor, *hlh5* (**Fig.4D**). We detect an expansion of this gene family, with detectable expression for 5 gene models (**Fig.S3A**). Two are expressed in all class 2 neurons, while others show expression also in cnidocytes and mucin gland (**Fig.S19C**). Only *hlh5* (n2.hlh5 neuron) and *atoh8-like-1* (mucin.medusa) show cell-type expression in medusa-specific cells. This suggests these two medusa cell types may be novel cell types that arose after gene duplication, as has been proposed for anthozoan muscle (36).

## Striated and smooth muscle share partial transcriptomic profiles.

The striated musculature is one the key features of the medusa stage, and striated muscle fibres have not been described within the polyp. Transcriptomic profiles characterised by a clear upregulation of muscle-related structural genes like myosin, actin, and actin-binding proteins were expected to be of either smooth or striated muscle nature. One of the main partitions in the dataset appears at the strobila stage and is defined by a large set of muscle-related genes, has a large contribution from umbrella tissue scRNAseq library, and is missing from the manubrium and gastrodermis tissue libraries (**Fig2B:** blues). We therefore conclude this cell cluster to be representative of the striated musculature. Candidate specification markers include two *otx* orthologs [*otx1a, otx2*], which have previously been shown to express

within the striated muscle field (19), and the single-copy ortholog of the PaTH (Paraxis, Twist Hand-related) bHLH family of proteins *cpath* (36) **Fig. 2C**). Within sea anemones, expansion of this gene family has been demonstrated to be linked to the expansion of the muscle repertoire in the starlet sea anemone *Nematostella vectensis* (36). Within the outer layer population, we can identify an additional sub-population of cells with a clear muscle signature. UMAP topology suggests intermediate cell states along a trajectory leading from the smooth muscle in the outer cell layer towards the striated musculature of the developing medusa. To investigate the dynamics of the muscle, we further analysed the subset including the outer cell layer and the striated muscle population together (muscle subset).

We first sought to confirm the morphological location of the muscle fields with phalloidin staining of active filaments in *Aurelia* polyps and ephyrae. We find that polyp muscles split into three muscle fields, the longitudinal tentacle muscles (**Fig. 5F**), the radial muscles of the oral disc (**Fig. 5E**) and the longitudinal retractor muscles that run along the body column (**Fig. 5H**). These data largely confirm previous studies that have investigated the various muscle fields of *Aurelia* jellyfish on the ultrastructural level based on electron microscopy (37,38). During strobilation, the aboral part of the polyp remains, which after detachment of the entire stack of ephyrae regenerates into a functional polyp again (**Fig. 6A,B,G**). In the freshly detached ephyra, we found two muscle fields (**Fig. 7B-D**), both consisting of striated muscles as has been previously described (38). One muscle field consists of radial striated muscles that extend into the ephyra lappets (**Fig. 7C**), the other field of circumferential striated muscles that run around the oral disc (**Fig. 7D,E**). However, some ephyrae still bear leftover fragments of the polyp retractor muscles (**Fig. 6S**). As medusa development progresses, these two muscle fields will merge into one umbrella muscle (39) . At the periphery of these striated muscle fields in the ephyra, at the tips of the lappets, we observe alternated smooth and striated muscle fibres (**Fig.7:F**) in both, freshly detached and metaephyrae (an intermediate stage between ephyra and juvenile medusa) (40) . Further, we find mixed smooth and striated muscle fibre composition at the base of the developing manubrium in metaephyrae (**Fig. 7G-I**). In some cases zones of striation can be observed within otherwise smooth muscle fibres (**Fig. 7H,I**).

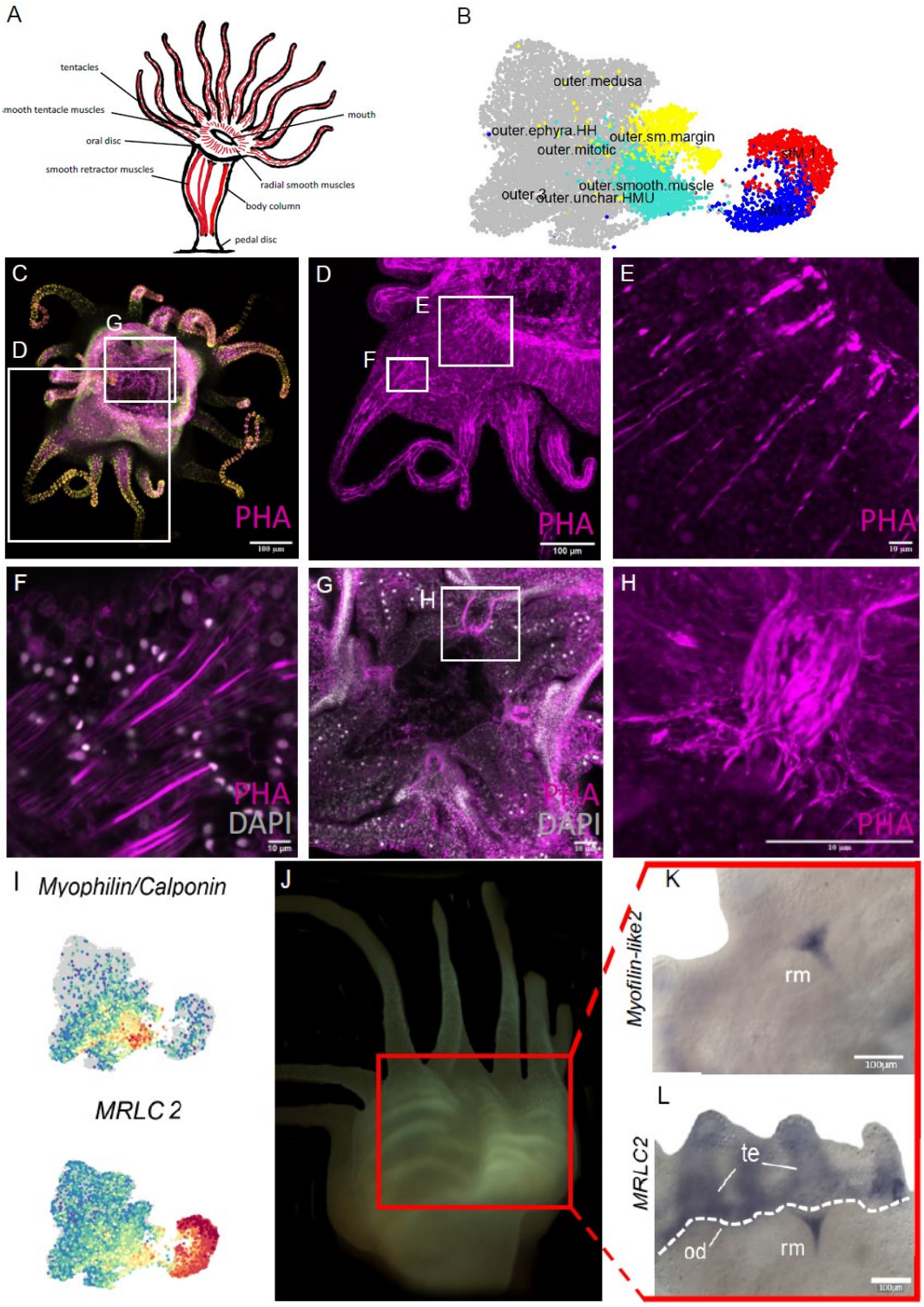



**Fig.5. Muscle territories in the *Aurelia* polyp.** (**A**) Schematic representation of the different smooth muscle territories in the polyp from a lateral view. (**B**) A reference UMAP of the muscle subset with highlighted smooth (yellow, turquoise) and striated (blue, red) muscle cell clusters. (**C-H**) Phalloidin stains actin-rich muscle fields in the animals. (**I**) Gene expression plots illustrating muscle gene expression across muscle subset. (**J-L**) In situ hybridization for two marker genes confirms the spatial location of the retractor muscles, tentacle muscles, and muscles at the oral disc of the polyp. te = tentacle, od = oral disc, rm = retractor muscle.

Next, we wished to confirm the identity of smooth and striated muscle cell types detected in the scRNAseq data by in situ hybridization (ISH) (**Figs. 5,6,7**). *myofilin-like2* is restricted to smooth retractor muscles of polyps (**Fig. 5K**) and expression is still visible in fragmented retractor muscles in the late strobila, consistent with the single-cell expression profile (**Fig. 6D**). *stmyhc5* reads are restricted to the striated muscle cluster (**Fig. 7P**) and are expressed in the striated muscle fields of the meta ephyra **(Fig. 7M**). *mrlc2-like* and *stmhc1* reads are detected in both muscle types (**Fig. 5L, Fig 6C,D,F-R, Fig. 7J-L,N,O**), and ISH indicates that the expression is localised to the fields of striated muscles in ephyrae (**Fig. 7J,K,N**), as well as the smooth muscle populations in polyps including longitudinal tentacle muscles, radial muscles of oral disc and retractor muscles of the body column (**Fig. 5L**), and the muscles of the manubrium in the meta ephyra (**Fig. 7L,O**). However, *stmhc1* expression is barely detectable in smooth muscle scRNAseq profiles (**Fig. 6E**) and ISH showed that its expression is detectable in developing smooth and striated muscles in the strobila (**Fig. 6C,F-M**), ephyra or meta ephyra (**Fig. 6N-R; Fig. 7N,O**) as well as muscles in regenerating polyps (**Fig. 6F-H**). Having confirmed the spatial gene expression patterns of different muscle-cluster marker genes, we generated gene lists specific to either smooth-, striated-, or both muscle types (**Supp. Data S2**).

In general, *Aurelia* muscles show a clear molecular signature specifically enriched in muscle related structural genes. While smooth and striated muscles are clearly distinguishable by the differences in their molecular content, they do also share a set of genes (**Supp. Data S2.1**). However, a large fraction of both muscle profiles consists of genes with unknown affinities (**Supp. Data S2**). *Aurelia* striated muscles express many cluster-specific structural gene isoforms affiliated with a putative contractile core complex in contrast to *Aurelia* smooth muscles, which share expression of most of these genes with the striated muscles (**Supp. Data S2**). In line with previous findings, the striated muscles express a core set of genes shared with bilaterian smooth muscles, for instance *calmodulin*, *myosin light chain kinase (mylk)* or *tropomyosin* ((41,42) (**Fig.6:F; Supp. Data S2.2**).

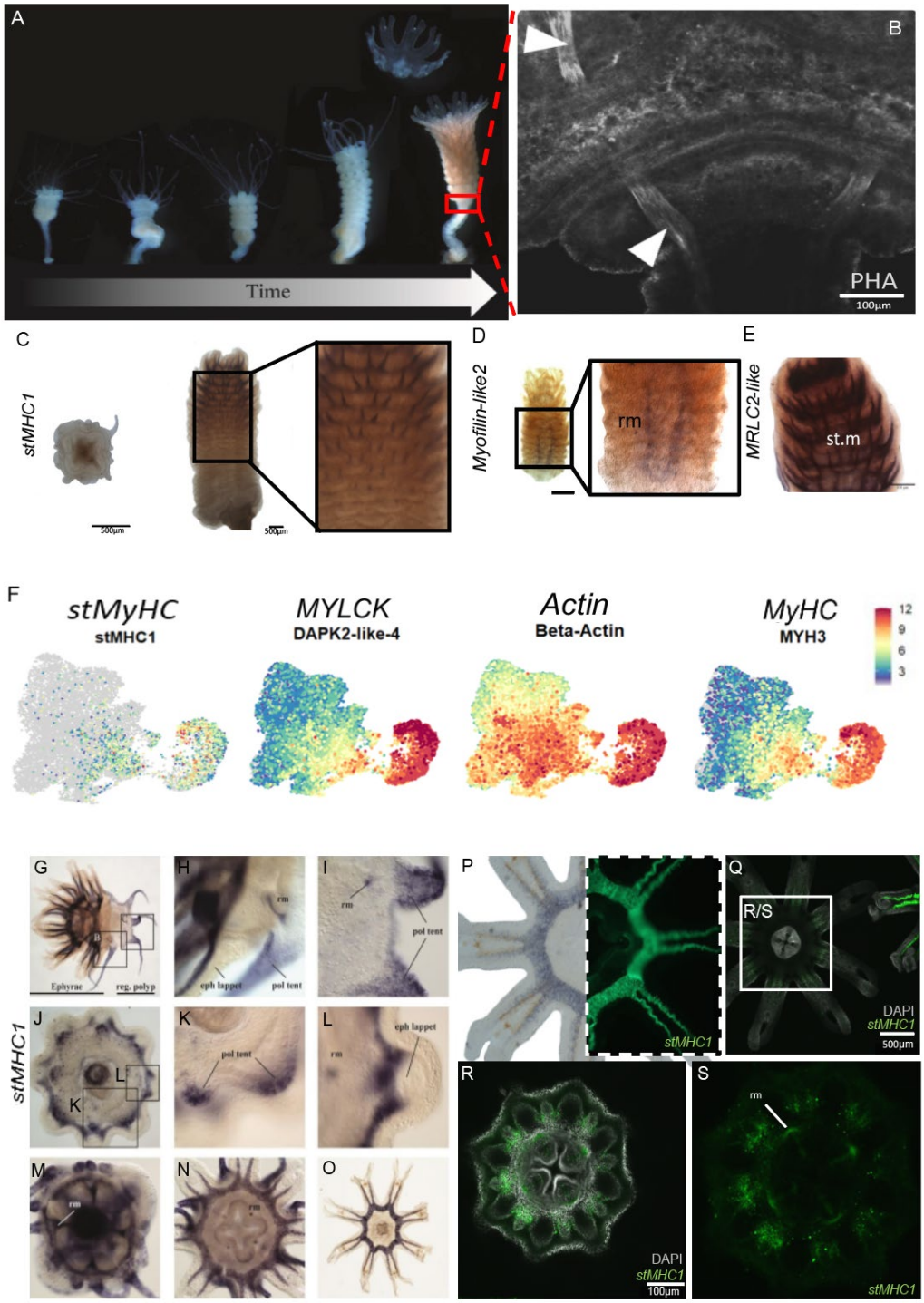



**Fig. 6. Establishment of striated muscle field during strobilation.** (**A**) Consecutive stages of strobilation from polyp (scyphistoma) to strobila and ephyra. (**B**) Foot region of the strobila stained with phalloidin showing polyp retractor muscles (arrowheads). (**C,D,E**) ISH for marker genes that label either smooth or striated muscles in the strobila. (**F**) shows gene expression plots for genes that are specific to smooth and striated muscle populations. (**G-O**) gene expression patterns of *stMHC1* in the advanced strobila. (**P-S**) gene expression pattern for the same gene in the detached ephyra. st.m = striated muscles, rm = retractor muscles.

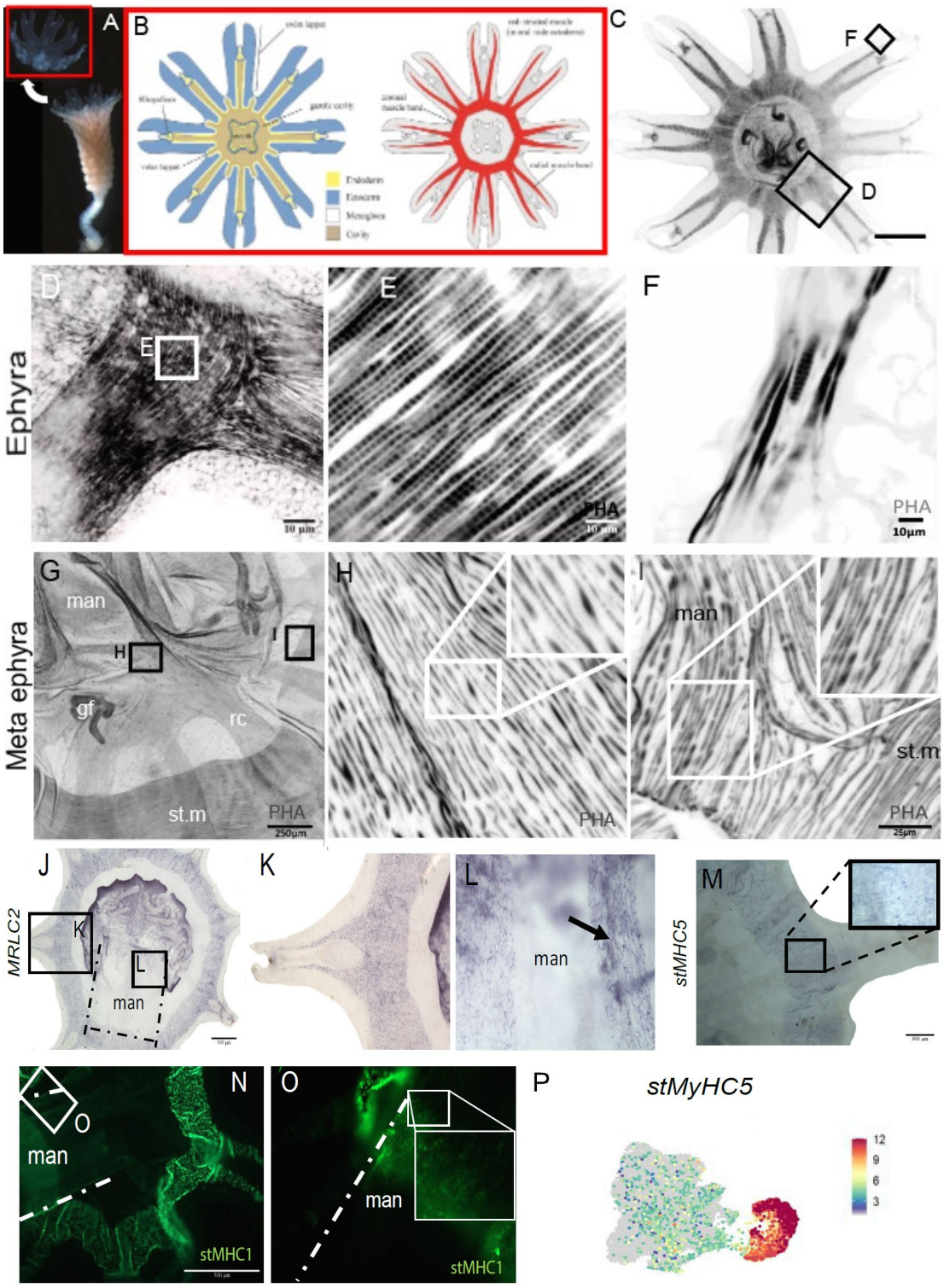



**Fig.7: Muscle territories in the ephyra.** (**A**) Ephyra detachment from the strobila. (**B**) Schematic of the ephyra body plan illustrating tissue anatomy (left) and striated muscle field (right). (**C-I**) Phalloidin labels the different muscle fields in the ephyra that are composed of either smooth, striated or a mixture of both muscle types. (**J-O**) depicts in situ hybridization for striated- (*stMHC5*), as well as all muscle marker genes (*stMHC1, MRLC2*) identified from single-cell RNAseq data. (**P**) An expression plot of striated muscle specific gene *stMHC5*. man = manubrium, gf = gastric filaments, rc = radial canal, st.m = striated muscle.

## Gene signatures suggest presence of a multipotent cell state.

In other cnidarian lineages, distinct populations of stem cells have been found either morphologically or molecularly (23,35,43). We observe cellular trajectories arising from the *foxA* positive partition towards differentiated cell types that in other cnidarians arise from the dedicated stem cell populations (**Fig. 8**). Within this partition, we identify a subset of cells that are in active mitosis, enriched in expression of numerous histones and other genes indicative of mitotic cellular processes (**Fig. 8A; S23, Data S1**). From these data, we identify two putative chromosomal HMG proteins that appear to play a role in mitosis: *hmgl-like* and *hmgb3-like*, and a third, *soxf2* that is upregulated together with a *nanos* ortholog (**Fig. 8C,D, DataS1**). These cells are also enriched for transcription factors that we find expressed in putatively early neural states (*pitx2 and rx3-like1*). We can identify trajectories emerging from this population that correspond to different cell populations. For example, we detect ciliated cells (*enkur* and *foxJ* expression) found throughout all stages and concentrated within the gastodermis-enriched sample (**Fig. S22, dark blue**), which shares gene expression profiles with subsets of both the digestive gland population and the class 1 neurons, including transcription factors *arnt* and *klf15-like-1* (**Fig. 8C,D**). Other branches are specific to the medusa stage, and express markers of class 2 neurons (*ashA*), cnidocytes (*prd13*), striated musculature (*cpath*), the intermediate zone marker *twist*, or putative germ cell markers (*piwi*- and *foxO*- orthologs) (**Fig. 8C,D**).

These putative stem cells derive from a population of cells that show intermediate properties between the inner and outer cell layer, and include *twist* expressing cells that line the boundary between the two tissues in the ephyra (31). We used Monocle3 to calculate putative trajectories through the entire data subset (**Fig. 8A**)(44,45). We ran the analysis from UMAPs generated from 12 random start seeds to estimate stability of the trajectory reconstructions (**Fig. S23C**). As has been noted in anthozoans, trajectory origins indicated by expression data are not always readily detected by this method. This is particularly true here for cnidocytes and mucin gland cells (red trajectories). Nonetheless, class 1 neurons and digestive glands consistently fall within the same trajectory, together with the class 2 neurons. Striated musculature connects with the smooth musculature of the outer cell layer, as expected from the anatomical data (**Fig.7F**). We compared the trajectory inference with cell type similarity estimates by constructing a neighbour-joining tree (**Fig. S23D**). Here we also recover clustering of the digestive glands and class 1 neuro-secretory cells. Cnidocytes and class 2 neurons also form a well supported cell-type clade together with the outer epithelial layer. The cluster boundaries between the intermediate layer and either epithelial layer are not very robust, thus it is possible that cells from both layers contribute to the multipotent stem cells.

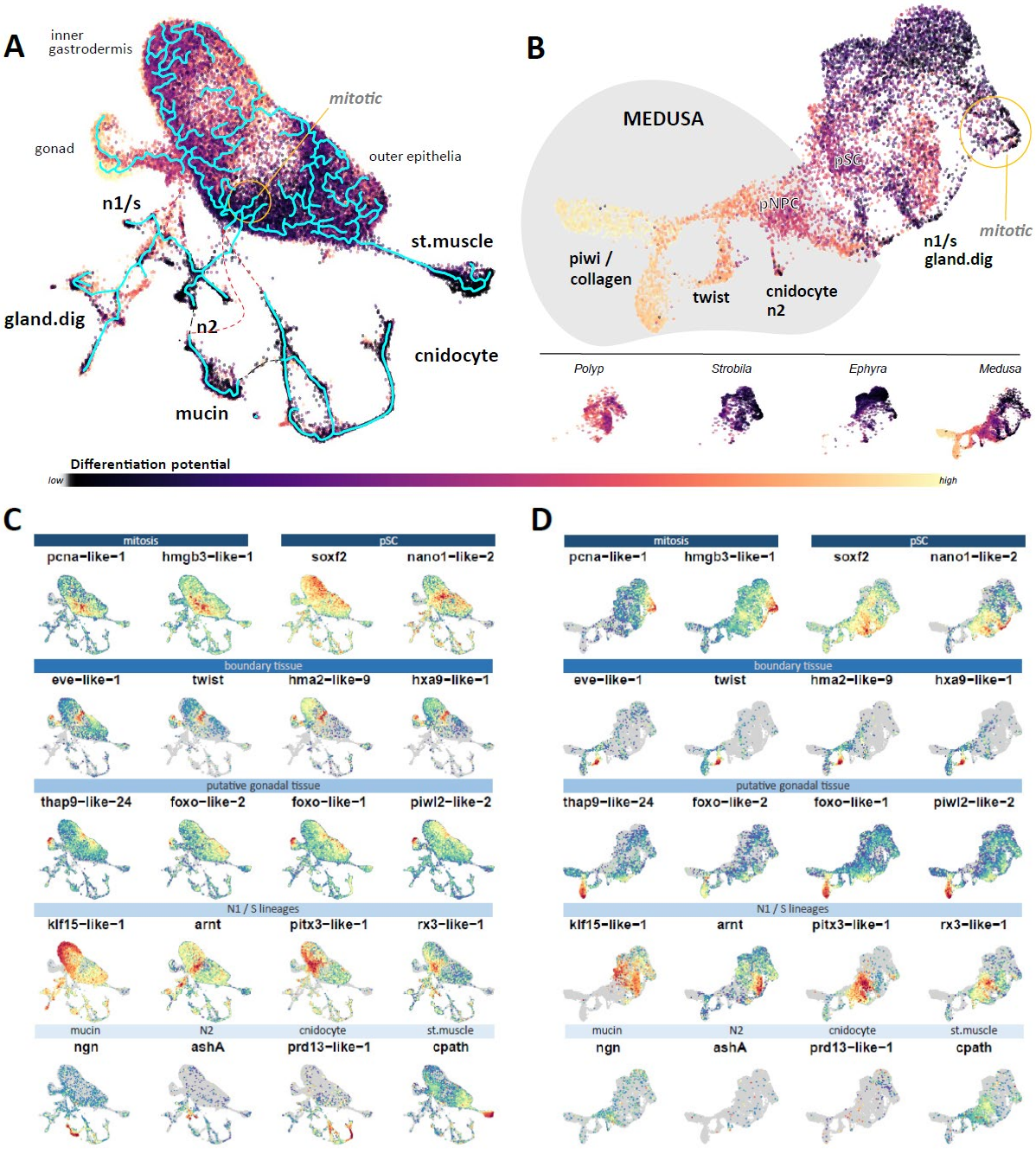



**Fig.8 Putative multipotent stem cell cluster demonstrates early initiation of many differentiation lineages**. (**A**,**B**) Differentiation potential as inferred with cytoTRACE indicates an increase in potential in the medusa-specific collagen/PIWI-expressing branch. (**A**) full dataset from all stages. MONOCLE3 estimation of differentiation trajectories (cyan). (**B**) intermediate cell cluster. Bottom: contributions from separate life cycle stages. Mitotic centers are highlighted in yellow. (**C**,**D**) Expression of specific differentiation markers suggests specific lineage relationships between cell types in the context of the entire dataset (**C**) and within the central cluster which includes the putative stem cell population (**D**).

## Discussion

**Single cell transcriptomic atlas reveals distinct cellular subtypes associated with the formation of medusa in the moon jelly.**

The life cycle of the sea anemone *Nematostella vectensis* has been characterized at a single cell level by recent studies (23,46). However, the medusozoans typically have a more complex, metagenic life cycle with a transition of the asexually reproducing polyp to the sexually active medusa. The formation of these two distinct adult life forms has fascinated researchers for more than a century ((8,31,47,48)). Within medusozoans, hydrozoans, scyphozoans and cubozoans form medusae in drastically different ways, raising the question whether they are homologous or convergent ((31,49)). While recent single cell transcriptome studies analysed the freshwater polyp Hydra (35), which has lost the medusa stage, or of the medusa stage of *Clytia hemisphaerica* (43), at present there is no understanding of the molecular and cellular changes during the metagenic transition from the polyp to the medusa. Here we present the first single cell characterisation of the metagenesis in a scyphozoan, the moon jelly *Aurelia coerulea* (formerly sp.1). We document an overall increase in gene detection of approximately 20% during formation of the medusa. Other studies have looked for differential gene use from bulk RNA extracts from medusa versus polyp stages, and have found no evidence for increased taxon-specific genes in either stage (27), or even composition of reconstructed proteome or transcriptomes from polyps compared to medusa (50). While initially this would seem to contrast our finding of an overall increase in genome usage in the medusa compared to the polyp, these results are not inconsistent. While we do not find any polyp-specific cell subtypes within any of the recovered cell populations, we document a few distinct cell types associated with the generation of the medusa stage that are detectable already in the strobila but absent from the polyp. Apart from the striated musculature, which is an abundant component of the medusa, these medusa-specific cell types comprise a small fraction of the overall tissue composition and share transcriptomic similarity with related cell types that are present across the life cycle. Comparisons of differential gene expression clearly demonstrate changes in the overall gene representation across the life cycle (51), in agreement with the current results.

Of the thirteen cell clusters that are found exclusively in the medusa stage, five of these are neurons. We can tentatively associate these neural subtypes with components of the nervous system known to arise in association with the formation medusa-specific tissues, including the rhopalia, motor nerve net, and striated muscle (19,25,31). We note extensive parallels in the complement of neural subtypes with the organisation described in the sea anemone, in particular the partitioning into two neural classes. This division is similar to anthozoans, where

Steger et al (36) identify insulinoma positive and negative neural trajectories in *Nematostella*. In *Nematostella*, the INS-positive neuron population is more abundant and possesses a greater number of subtypes, including the sensory/secretory "S"- class of largely uncharacterized cell types. It is this subset that predominates the polyp stage of the lifecycle in *Aurelia*. The digestive gland cell populations in both species express orthologs of the achaete-scute family (*ashC*) (23). AshC is an ortholog of the ascl345 family in vertebrates, which is expressed in the skin (ASCL4), salivary glands (ASCL5), and teeth (ASCL6) [UNIPROT database (52)]. Orthologs of *ashB* and *ashD* are expressed in the neurons in both, *Nematostella* and *Aurelia* (23), and both are orthologs of the Achaete-scute complex, notable regulators of neurogenesis in bilaterians (53). AshA does not have a mouse or human ortholog, but is present within nematodes, where it is a key regulator of a nociceptive neuron identity (54). In *Aurelia* the *ashA* ortholog is restricted to the class 2 neurons, while in *Nematostella ashA* expression is found in both types (23,55) and *ashD* is restricted to INS-positive (n1) neurons in both species. The INS-negative populations in both lineages also express *pou4* orthologs, found also within the cnidocyte lineages and thus further supports claims of a close relationship between cnidocytes and *insulinoma* negative / *pou4* positive n2 neurons (23,56). Further, we demonstrate a similar specification trajectory in the formation of cnidocytes as has been described for the sea anemone *Nematostella vectensis* (23,24,57). This strongly suggests that the radiation of achaete-scute genes not only preceded the split of cnidarians and bilaterians, but also that the paralogs were implemented in the diversification of the neuro-glandular lineages across the Metazoa.

**Transcription factor family expansions may be related to the generation of novel medusa cell types.**

Recently, we demonstrated an anthozoan expansion of an orphan bHLH family that correlated with the expansion of muscle cell types (23,36). Expansions of other transcription factor families in a similar manner could be related to generation of novel cell types (24). The medusa-specific cell types identified here also demonstrate cell-type specific expression of expanded transcription factor families. For example, we find *Aurelia distaless* (*dlx*) paralogs expressed specifically in the putative sensory neurons of the rhopalia [n2.pitx.medusa], an organ found within scyphozoan and cubozoan lineages (19). *Myc* genes appear to be expanded within the cnidarians, at least in *Nematostella* [6 copies (29)] and scyphozoans [7 or 8 copies; **Fig. S3**]. While four orthologs are reported from the hydrozoan *Hydra* (58), only 2 have been investigated (59–61). Similarly, only 2 orthologs are reported from another hydrozoan (62). Interestingly, of the two *hydra myc* genes, one is expressed in the cnidocyte lineage, while the other is expressed in the interstitial stem cell lineage (35). The myc

expansion in *Nematostella* and *Aurelia* are used in parallel ways during cnidocyte development, and orthologous pairs of *Nematostella* and *Aurelia* myc genes are similarly expressed in the early and mature phases of cnidocyte differentiation (**Fig. S3**). In both lineages, the basal-most *myc* paralog is expressed within the putative stem cells. Within the atonal/neurogenin bHLH TF clade we find two expansions, one shared with *Nematostella* and thus likely representing a cnidarian expansion, and the other specific to *Aurelia* (**Fig. S3**). Within the *Aurelia* branch we find paralogs that define pou4-positive/insm-negative class 2 neurons, but also others that have expression restricted to the medusa-specific putative motor neuron population n2.hlh. These neurons also express paralogs of the expanded *otx* family. Of interest, expression of one of these *otx*-family paralogs, *otx1b*, is shared with the striated muscle, while the *otx2* paralog is exclusively expressed in the striated muscle. These data support the hypothesis connecting gene duplication with cell type expansion (41,63–67), suggesting that at least some of the medusa-specific cell types identified here may represent synapomorphies of the Medusazoa, or perhaps even the Scyphozoa. *Otx* was also found to be expressed in striated muscles of the hydrozoan *Podocoryne carnea*, suggesting that its role in striated muscle formation is ancestral to the medusozoans (66). Of note, in bilaterians, o*tx* genes mark mostly anterior parts of the body, in particular anterior brain regions in a conserved manner. While it has been implicated in repressing muscle development in ascidians (68), to our knowledge has never been implicated in positive regulation of muscle development. This supports the view of an independent evolution of striated muscles in cnidarians and bilaterians (41).

**Shared core contractile machinery between smooth and striated muscles indicates a gradual maturation of smooth fibres into the striated phenotype.**

Striated muscles enabled jellyfish to evolve a unique swimming behaviour and conquer the oceans. Still, we don't know much about the origin and evolution of striated muscle cell types. To gain more understanding on the formation of this type of muscle in true jellyfish requires to identify their molecular identities in the first place. In addition, understanding its developmental and evolutionary origin can also aid in our understanding of the evolution of the tri-phasic medusozoan life cycle. We distinguish muscle cell transcriptomes across different stages of the *Aurelia* life cycle, including both striated and smooth types. The striated muscle population is the largest medusa-specific cell population with a distinct transcriptomic profile. The smooth muscle transcriptomic profile shares features with the outer epithelial layer and is recovered as a cell state within this larger population. While a small set of genes are shared across the two muscle phenotypes (eg *stmhc1* and *mrlc2-like*), others are more specific to either phenotype (eg. *stmyhc5* in striated muscle; *myofilin-like-2* in smooth muscle), which were

verified by in situ hybridization (**Figs. 5,6,7**). Thus, a large fraction of muscle-related structural genes is most likely part of a core contractile machinery that is shared between smooth and striated muscles (**Data S2.1**). Interestingly, a functional study done by Tanaka et al. (42) linked some of the genes expressed in *Aurelia* striated muscles to a bilaterian smooth muscle contractile apparatus. We were also able to identify the genes used in the study by Tanaka et al. (42) from our scRNAseq data (**Data S2.1**). This core contractile complex is regulated via phosphorylation of *myosin regulatory light chain* (*mrlc*) by *myosin light chain kinase* (*mlck*) in the presence of Ca2+ which is known to be specific for bilaterian smooth muscle contraction (42,69,70). *mrlc's* are also present in bilaterian striated muscles but appear to be not mediating Ca2+ binding via phosphorylation and thereby triggering contraction (71,72). Main players in this smooth contraction mechanism are *calmodulin*, *mrlc*, *melc* or *mlck*, which are expressed in both *Aurelia* muscle types identified from our single-cell data (**Data S2.1**). Furthermore, we found that at least one of our identified muscle markers (*stMHC1*) seems to be exclusively expressed in either regenerating smooth or developing striated muscles, but not in mature smooth muscles of polyps (**Fig. 6C,G-O**). We further find smooth and striated muscle fibres at the periphery of the ephyra striated muscle fields **(Fig. 7F)** and we could find expression of *stHMC1* at the base of the developing manubrium in the metaephyra **(Fig. 7N,O),** which also showed a mix of both smooth and striated muscle phenotypes (**Fig. 7G-I**). Strikingly, we also observe individual smooth fibres located at the base of the developing manubrium that show striation at some spots along the fibre (**Fig. 7H,I**), raising the possibility that smooth and striated fibres can be found in the same cell. In line with that, the ISH expression data together with the juxtaposition of both fibre types within the ephyra (**Fig. 7F,G-I**) are suggestive of a gradual maturation of smooth fibres into the striated phenotype, at least during metamorphosis from the polyp to the medusa form. Taking these results together we hypothesise that during strobilation striated muscles might originate from smooth muscle precursor cells already present in polyps. However, future work is required to further provide experimental evidence for this scenario.

**Putative multipotent stem cells in Aurelia do not share the typical molecular characteristics observed in other cnidarians.**

Recent discussions regarding the utility of UMAP embeddings in data inference has highlighted the necessity to use caution when making conclusions regarding cell-cell relationships (73). Nonetheless, in cases where the sampling resolution is sufficient, accurate cell lineage trajectories can be recovered from single cell transcriptomic data using these types of embeddings (74–78). Predicted lineage relationships have been validated with transgenic lines in the sea anemone *Nematostella vectensis* (23), and by *in situ* hybridization in *Hydra*

(35), illustrating the value of such reconstructions when used with caution. We replicated the trajectory inference from such an embedding by running the algorithm multiple times and indeed we find variation in the predicted cell relationships that would be expected from this approximation. Nonetheless, the relationships that are recovered are remarkably similar to those described in other cnidarian lineages (23,35,43). However, at the level of the putative multipotent stem cell, *Aurelia* does not share the typical characteristics. In both anthozoans and hydrozoans, multipotent neuroglandular progenitor cells are clearly identifiable molecularly (23,35,43). In the hydrozoans, interstitial stem cells have been known for decades (79), while in the sea anemone their precise morphological location within the animal remains cryptic. Nonetheless, the neuroglandular progenitor cells share expression of *myc, ngn, soxB*, and *soxC* orthologs. In contrast, here we find that the direct orthologs for the anthozoan counterparts are not over-expressed within the putative stem cell territory but rather show cell-type specific expression. Putative multipotent stem cells identified here do not harbour a strikingly unique transcriptomic signature at all. While we interpret these cells as the putative origin of all neuro-glandular cell types in an analogous manner as described in other cnidarians, this interpretation could be erroneous. Support for our interpretation includes expression profiles of cell lineage - specific transcription factors (**Fig. 6**), and over-abundance of *soxf2* within cells at this node. The *soxB(2)* ortholog (*soxB1b*), which in *Nematostella* and bilaterians is a conserved upstream regulator of neurogenesis, shows rather broad expression across nearly all catalogued cell states, with the exception of the inner cell layer. In summary, the nature and localization of multipotent stem cells in the scyphozoan remains enigmatic at this point.

**Limitations of the study**

Aurelia lab strains have been commonly used as part of developmental, evolutionary as well as ecological studies in the last decades. However, *Aurelia* is not generally considered a model organism largely due to the inability to close the animal's life cycle in a laboratory setting. Access to early developmental stages is very limited and gametes or larvae can only be harvested if animals are captured from their natural environment, and consequently gene manipulation methodologies are limited. Further, cell dissociations can only be achieved through enzyme-based approaches due to large amounts of acellular mesoglea that is difficult to remove from cell dissociations. Immune response to the presence of degraded extracellular matrix components could elicit a transcriptomic stress response, broadly influencing gene expression values across the dataset.

## Material and Methods

*Aurelia coerulea* (formerly sp.1) culture:

The lab strain was obtained from the collection of Dr. Harms (University of Hamburg approx. 25 years ago). Genomes of several *Aurelia* strains have recently been published and so sequenced reads were first mapped to all available *Aurelia* genomes, with gene annotations extended by 1000bp towards the 3' end, or until reaching the next gene model in the same orientation (15,27,80). We were able to align 90% of the reads from our libraries to the genome of the San Diego aquarium strain (27), compared to only 40% to the atlantic strain or 84% to the pacific/Roscoff strain (15). CO1 gene trees cluster the Roscoff and SanDiego aquarium strain together with *Aurelia sp.1*, which is now identified as *Aurelia coerulea* (81). *Aurelia* scyphistomae were grown on small petri dishes placed in 80ml glass containers filled with ASW (artificial sea water) with the salinity adjusted to 35‰. Polyp cultures were fed with *Artemia* nauplii once a week. Containers were cleaned from algae approximately once a month. Strobilation of polyps was induced through either a shift in temperature of about 2-5°C or an 24h treatment with 10µM Indomethacin (28). Freshly liberated ephyrae were transferred to a 5L beaker. Water motion was induced via plastic propellers attached to a small 12V engine. Ephyrae were fed once a day with freshly hatched artemia nauplii. Small jellyfish were transferred to a 30L Kreisel Tank and daily fed with *Artemia* nauplii, which were incubated in Selco S.presso (INVE) and diluted isochrysis extract overnight to enrich live food with nutrients. Kreisel Tanks were cleaned once a week.

Single-cell dissociations:

Magnesium-calcium-free artificial seawater (MCF-ASW) was prepared according to (82). Either scyphistomae, strobilae, ephyrae and juvenile medusae were collected in to 1.5ml tubes, washed 2x with FSASW (filter sterile artificial sea water) and incubated in MCF-ASW afterwards, followed by an incubation in 300µl Collagenase/Dispase Blend I (Merck) dissolved in 35‰ FSASW at RT. ½ h after incubation started mechanical force was used to disrupt the rest of the tissues and obtain single-cell suspensions. The status of the suspension was monitored using a compound microscope. To stop the digestion, the suspensions were put on ice and 0.1% BSA was added (bovine serum albumin). The suspensions were transferred to a pre-coated (0.2% PTW) and extensively washed 1.5ml tube. Next, cells were pelleted for 5 minutes and 0.4rcf using a pre-cooled centrifuge (10°C). After the cells pelleted, the supernatant was removed. Further, tubes were filled with 100µl of BSA (0.1% final concentration) and 100ml MCF-ASW. Further the pellet was resuspended and mixed properly

by gently pipetting. After resuspension, the samples were filtered through a Flowmi-filter (40µm). Next, fractions of the suspensions were stained with a seawater buffer containing 10µM Fluorescein and Propidium iodide. Cell quantities and viabilities were validated via a Nexelcom cell counter. Finally, single-cell suspensions were loaded into microfluidic chips (10X Chromium microfluidic chips).

Library preparation

Single-cell libraries from *Aurelia* polyps, late strobilae, ephyrae, juvenile medusae as well as tissues from larger jellyfish were generated from either whole animal single-cell suspensions or suspensions prepared from dissected tissues. Isolated single-cells were encapsulated in oil droplets using a 10X Genomics single-cell controller, using version 2 reagents for the polyp and medusa, and version 3 reagents for the strobila and ephyra as well as dissected tissue libraries. Viable cell suspensions were diluted to a concentration of 1000 cells/ml, loaded with the target cell recovery of 5000-10.000 cells and processed into a sequencing library following the manufacturer protocol. The resultant single-cell libraries were sequenced with the Illumina platform according to the 10X genomics specifications for paired end sequencing.

Transcriptome mapping:

The best matching genome mapping had only 40% of reads assigned to an interpretable gene model. To improve data recovery the following approach was used: Bulk RNAseq libraries generated in house were assembled with the Trinity assembler v2.8.4 using standard parameters. Additionally, the same libraries were aligned to the San Diego reference genome (27) using STAR v2.3.7a with standard parameters. Single cell RNAseq libraries were also aligned to the same reference genome using cellranger v6.1.2. Final bam files were merged, and gene models were called with stringtie v2.2.1 using standard parameters. This way, three transcriptomes were available: reference guided transcriptome (stringtie), ab initio transcriptome (San Diego transcriptome) and a denovo transcriptome produced by trinity. All three transcriptomes were merged, and redundant transcripts were removed from the assembly using two rounds of clustering with cd-hit v4.8.1. In the first round, all transcripts that were completely contained in another larger transcripts were removed. In the second round, CDS of representative transcripts were predicted with transdecoder v5.5.0 prior to clustering. Redundant CDS were removed following the following criteria: 1) at least 90% of the CDS was completely contained in another CDS; 2) minimum sequence identity of 99.9%. Finally, the quality of the transcriptome assembly was assessed by BUSCO completeness percentage, BUSCO duplication percentage and RNAseq mapping efficiency.

Functional annotation

Final transcripts were functionally characterised using different databases including Pfam, Uniprot, NCBI NR, InterproSites, TMHMM, SignalP and eggnogmapper. Putative names were assigned based on sequence homology to Uniprot best blast hit or NVBI NR best blast hit.

Single-cell analysis

Raw sequence reads were first processed through the cell ranger pipeline with forced recovery of 7000 cells and mapped using a custom mapping tool. Genes are named here according to their best blast hit whenever possible. UMI-collapsed count matrices were imported into R for further processing with the Seurat package (vs. 4.3) (30,83,84) [GenerateAaAllData.R https://github.com/technau/AureliaAtlas]. Expression matrices from individual samples were first filtered to include only samples with at least 750 unique reads. Samples were further filtered for minimal genes (sample range: 350-500), and putative doublets were removed by eliminating outliers with high count values (sample range: 4000-40000). The data was then log normalized with a scale factor of 5000 reads (Seurat::NormalizeData), and top 2000 variable genes were identified for each sample independently (Seurat::FindMarkers). To avoid clustering according to cell cycle state, the resultant variable genes were filtered to exclude those involved in regulation of the cell cycle. Scaled gene expression of this combined set was used as input to the RunPCA function prior to data integration. Samples were then integrated with a reciprocal PCR approach, using the Seurat::FindIntegrationAnchors and Seurat::IntegrateData functions.

The resultant integrated dataset was then used for analyses of cell-state composition following a standard Seurat workflow. Briefly, the integrated data was scaled and used to identify the first 22 principal components with standard deviation of >2. These 22 components were used to generate a nearest neighbour graph (Seurat::FindNeighbors with annoy.metric = ‘cosine’, k.param = 10) from which cells were clustered with a resolution of 0.1. To visualise the clusters, all cells were projected in a two-dimensional space (UMAP). Each population was then re-analyzed separately, and distinct cell states and/or cell types from each population were annotated. Cell cluster identity was assigned semi-automatically as described in (23). Gene sets used to annotate the single libraries are found in Supplemental Data S3. These annotations were then imported into the full dataset for further exploration. To explore data from the perspective of the life history stage, samples of a similar stage were grouped together (polyp = 3, strobila = 1, ephyra = 1, medusa = 5). Upregulated marker genes for each cluster (or stage of the life cycle) were calculated with the Seurat 'FindAllMarkers' function, returning markers that show a fold-change equal or greater than 0.6, detected in at least 30% of the

cluster, and having a p-value threshold of >=0.0001. For presentation purposes, smoothed gene expression plots were generated by first imputing missing data, while retaining biological zeros with the SeuratWrappers::RunALRA function (85). Scripts for running all analyses can be found on our GitHub page: Aurelia51k.Atlas.LifeHistoryStages_clean.RMD; Aurelia51k.Atlas.Alldata_clean.RMD; Aurelia51k.Atlas.CellComplement_clean.RMD: https://github.com/technau/AureliaAtlas.

Immunohistochemisty and phalloidin:

Animals from different life-stages were relaxed in the fridge at 4°C followed by a fixation step via 4%PFA at RT for 2h and washed two times in 0.2% PTW after fixation with a minimum washing time of 1/2h. Further, they were blocked using 3% (w/v) BSA (bovine serum albumin) for 2 minutes and incubated in a 1:50 PHA Alexa568/0.2% PTW and 1:1000 DAPI solution for 2h at RT in the dark. After staining they were washed 3x in 0.2% PTW, incubated in vectashield (mounting medium) at RT for 1/2h, mounted on glass slides and imaged using a confocal microscope (Leica SP8). In case of immunostaining, both the sample and the primary antibodies (βTubulin (Sigma Sc-5274), Acetylcholine (Merck/Millipore AB5522), or FMRFamide (Millipore AB15348)) were blocked (1:300) in 0.1% FCS (fetal calf serum) for 1h at RT (sample) or at 4°C (AB). Afterwards the samples were incubated in the pre-blocked antibodies for 24h at 4°C and washed 2x with 0.2% PTW (1/2h/wash). Further secondary antibodies (Goat α rabbit, life technologies; Goat α mouse/rat, Invitrogen; Alexa488) were blocked for 1h in FCS (1:1000) at 4°C and washed. The sample was incubated in the secondary antibodies for 20h (4°C). After incubation treatment followed the same steps as after PHA and DAPI stains.

Cloning of genes and preparation of in situ hybridization probes.

RNA was extracted from single-cells, starting with a digestion of animal tissues form different life-stages in 500ml of Animal Free Collagenase/Dispase Blend I (Merck) dissolved in 35‰ FSASW (filter sterile artificial sea water) prepared with DEPC treated water. Single cells were pelleted with 0.4rcf at 4°C for 5 minutes after digestion. The pellet was resuspended, and RNA extracted from single cells by following a standard Trizol extraction protocol (86). Marker genes sequences for each population of cells were identified from scRNAseq data. Primer sequences were designed by the use of the primer3 website and marker gene sequences amplified from cDNA obtained from extracted mRNA. Genes were cloned using a pJET1.2 blunt vector cloning kit according to the manufacturer. Plasmids were transformed into E.coli and bacteria grown on ampicillin treated agar plates afterwards. Minipreps were performed using a peqGold miniprep kit (VWR). WMISH probes were generated via a SP6 or T7

transcription kit (NEB) and prospectively labelled with either Fluorescein or DIG-UTP. WMISH or FISH were performed by the use of a standard protocol with Hybe buffers based on Urea (87), but slightly adapted to the needs of *Aurelia* tissues (31). Commercial fluorescein/POD and DIG/POD antibodies and buffers were used according to the manufacturer's protocol (perkin elmer) for the performance of FISH.

## References


1. Pisco AO, Tojo B, McGeever A. Single-Cell Analysis for Whole-Organism Datasets. Annu Rev Biomed Data Sci. 2021 Jul 20;4(1):207–26.
2. Bump P, Lubeck L. Marine Invertebrates One Cell at A Time: Insights from Single-Cell Analysis. Integr Comp Biol. 2023 May 15;icad034.
3. Ryan JF, Burton PM, Mazza ME, Kwong GK, Mullikin JC, Finnerty JR. The cnidarian-bilaterian ancestor possessed at least 56 homeoboxes: Evidence from the starlet sea anemone, Nematostella vectensis. Genome Biol. 2006;7(7):R64.
4. Hejnol A, Obst M, Stamatakis A, Ott M, Rouse GW, Edgecombe GD, et al. Assessing the root of bilaterian animals with scalable phylogenomic methods. Proc R Soc B Biol Sci. 2009 Dec 22;276(1677):4261–70.
5. Philippe H, Brinkmann H, Copley RR, Moroz LL, Nakano H, Poustka AJ, et al. Acoelomorph flatworms are deuterostomes related to Xenoturbella. Nature. 2011 Feb;470(7333):255–8.
6. Galliot B. Hydra, a fruitful model system for 270 years. Int J Dev Biol. 2012;56(6-7–8):411–23.
7. Collins AG, Schuchert P, Marques AC, Jankowski T, Medina M, Schierwater B. Medusozoan Phylogeny and Character Evolution Clarified by New Large and Small Subunit rDNA Data and an Assessment of the Utility of Phylogenetic Mixture Models. Collins T, editor. Syst Biol. 2006 Feb 1;55(1):97–115.
8. Collins AG. Phylogeny of Medusozoa and the evolution of cnidarian life cycles: Phylogeny of Medusozoa. J Evol Biol. 2002 Apr 30;15(3):418–32.
9. Helm RR. Evolution and development of scyphozoan jellyfish. Biol Rev. 2018 May;93(2):1228–50.
10. Steele RE, David CN, Technau U. A genomic view of 500 million years of cnidarian evolution. Trends Genet. 2011 Jan;27(1):7–13.
11. Technau U, Steele RE. Evolutionary crossroads in developmental biology: Cnidaria. Development. 2011 Apr 15;138(8):1447–58.
12. Purcell J, Uye S, Lo W. Anthropogenic causes of jellyfish blooms and their direct consequences for humans: a review. Mar Ecol Prog Ser. 2007 Nov 22;350:153–74.
13. Shimomura O, Johnson FH, Saiga Y. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan,Aequorea. J Cell Comp Physiol. 1962 Jun;59(3):223–39.
14. Nong W, Cao J, Li Y, Qu Z, Sun J, Swale T, et al. Jellyfish genomes reveal distinct homeobox gene clusters and conservation of small RNA processing. Nat Commun. 2020 Jun 19;11(1):3051.
15. Khalturin K, Shinzato C, Khalturina M, Hamada M, Fujie M, Koyanagi R, et al. Medusozoan genomes inform the evolution of the jellyfish body plan. Nat Ecol Evol. 2019 Apr 15;3(5):811–22.
16. Helm RR. Evolution and development of scyphozoan jellyfish. Biol Rev. 2018 May;93(2):1228–50.
17. Helm RR, Tiozzo S, Lilley MKS, Lombard F, Dunn CW. Comparative muscle development of scyphozoan jellyfish with simple and complex life cycles. EvoDevo. 2015 Dec;6(1):11.
18. Piraino S, De Vito D, Schmich J, Bouillon J, Boero F. Reverse development in Cnidaria. Can J Zool. 2004 Nov 1;82(11):1748–54.

19. Nakanishi N, Yuan D, Hartenstein V, Jacobs DK. Evolutionary origin of rhopalia: insights from cellular-level analyses of Otx and POU expression patterns in the developing rhopalial nervous system: Otx and POU expression patterns in Aurelia. Evol Dev. 2010 Jul 2;12(4):404–15.
20. Spangenberg DB, Jernigan T, McCombs R, Lowe BT, Sampson M, Slusser J. Development studies of Aurelia (Jellyfish) ephyrae which developed during the SLS-1 mission. Adv Space Res. 1994 Aug;14(8):239–47.
21. Kakinuma, Y. An experimental study of the life cycle and organ differentiation of Aurelia aurita. Bull Mar Biol Stn Asamushi. 1975;15:3–6.
22. Fuchs B, Wang W, Graspeuntner S, Li Y, Insua S, Herbst EM, et al. Regulation of Polyp-to-Jellyfish Transition in Aurelia aurita. Curr Biol. 2014 Feb;24(3):263–73.
23. Steger J, Cole AG, Denner A, Lebedeva T, Genikhovich G, Ries A, et al. Single-cell transcriptomics identifies conserved regulators of neuroglandular lineages. Cell Rep. 2022 Sep;40(12):111370.
24. Babonis LS, Enjolras C, Ryan JF, Martindale MQ. A novel regulatory gene promotes novel cell fate by suppressing ancestral fate in the sea anemone *Nematostella vectensis*. Proc Natl Acad Sci. 2022 May 10;119(19):e2113701119.
25. Nakanishi N, Hartenstein V, Jacobs DK. Development of the rhopalial nervous system in Aurelia sp.1 (Cnidaria, Scyphozoa). Dev Genes Evol. 2009 Jun;219(6):301–17.
26. Gemmell BJ, Du Clos KT, Colin SP, Sutherland KR, Costello JH. The most efficient metazoan swimmer creates a ‘virtual wall’ to enhance performance. Proc R Soc B Biol Sci. 2021 Jan 13;288(1942):20202494.
27. Gold DA, Katsuki T, Li Y, Yan X, Regulski M, Ibberson D, et al. The genome of the jellyfish Aurelia and the evolution of animal complexity. Nat Ecol Evol. 2018 Dec 3;3(1):96–104.
28. Kuniyoshi H, Okumura I, Kuroda R, Tsujita N, Arakawa K, Shoji J, et al. Indomethacin Induction of Metamorphosis from the Asexual Stage to Sexual Stage in the Moon Jellyfish, *Aurelia aurita*. Biosci Biotechnol Biochem. 2012 Jul 23;76(7):1397–400.
29. Zimmermann B, Montenegro JD, Robb SMC, Fropf WJ, Weilguny L, He S, et al. Sea anemone genomes reveal ancestral metazoan chromosomal macrosynteny [Internet]. Genomics; 2020 Oct [cited 2023 Aug 1]. Available from: http://biorxiv.org/lookup/doi/10.1101/2020.10.30.359448
30. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive Integration of Single-Cell Data. Cell. 2019 Jun;177(7):1888-1902.e21.
31. Kraus JEM, Fredman D, Wang W, Khalturin K, Technau U. Adoption of conserved developmental genes in development and origin of the medusa body plan. EvoDevo. 2015 Dec;6(1):23.
32. Sur A, Renfro A, Bergmann PJ, Meyer NP. Investigating cellular and molecular mechanisms of neurogenesis in Capitella teleta sheds light on the ancestor of Annelida. BMC Evol Biol. 2020 Dec;20(1):84.
33. Lin HC, He Z, Ebert S, Schörnig M, Santel M, Nikolova MT, et al. NGN2 induces diverse neuron types from human pluripotency. Stem Cell Rep. 2021 Sep;16(9):2118–27.
34. Deryckere A, Styfhals R, Elagoz AM, Maes GE, Seuntjens E. Identification of neural progenitor cells and their progeny reveals long distance migration in the developing octopus brain. eLife. 2021 Aug 24;10:e69161.
35. Siebert S, Farrell JA, Cazet JF, Abeykoon Y, Primack AS, Schnitzler CE, et al. Stem cell differentiation trajectories in *Hydra* resolved at single-cell resolution. Science. 2019 Jul 26;365(6451):eaav9314.

36. Cole AG, Jahnel SM, Kaul S, Steger J, Hagauer J, Denner A, et al. Muscle cell-type diversification is driven by bHLH transcription factor expansion and extensive effector gene duplications. Nat Commun. 2023 Mar 29;14(1):1747.
37. Chia FS, Amerongen HM, Peteya DJ. Ultrastructure of the neuromuscular system of the polyp ofAurelia aurita L., 1758 (Cnidaria, scyphozoa). J Morphol. 1984 Apr;180(1):69–79.
38. Matsuno, A., Hisamatsu, T. Electron microscopy on striated muscles in the ephyra of Aurelia aurita. Mem Fac Sci Shimane Univ. 1982;(16):61–7.
39. Zimmerman KL, Jamshidi AD, Buckenberger A, Satterlie RA. Organization of the subumbrellar musculature in the ephyra, juvenile, and adult stages of *Aurelia aurita* Medusae. Invertebr Biol [Internet]. 2019 Sep [cited 2023 Aug 7];138(3). Available from: https://onlinelibrary.wiley.com/doi/10.1111/ivb.12260
40. Kienberger K, Riera-Buch M, Schönemann AM, Bartsch V, Halbauer R, Prieto L. First description of the life cycle of the jellyfish Rhizostoma luteum (Scyphozoa: Rhizostomeae). Steele RE, editor. PLOS ONE. 2018 Aug 22;13(8):e0202093.
41. Steinmetz PRH, Kraus JEM, Larroux C, Hammel JU, Amon-Hassenzahl A, Houliston E, et al. Independent evolution of striated muscles in cnidarians and bilaterians. Nature. 2012 Jul;487(7406):231–4.
42. Tanaka H, Ishimaru S, Nagatsuka Y, Ohashi K. Smooth muscle-like Ca2+-regulation of actin–myosin interaction in adult jellyfish striated muscle. Sci Rep. 2018 May 17;8(1):7776.
43. Chari T, Weissbourd B, Gehring J, Ferraioli A, Leclère L, Herl M, et al. Whole-animal multiplexed single-cell RNA-seq reveals transcriptional shifts across *Clytia* medusa cell types. Sci Adv. 2021 Nov 26;7(48):eabh1683.
44. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014 Apr;32(4):381–6.
45. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. Nat Methods. 2017 Oct;14(10):979–82.
46. Sebé-Pedrós A, Saudemont B, Chomsky E, Plessier F, Mailhé MP, Renno J, et al. Cnidarian Cell Type Diversity and Regulation Revealed by Whole-Organism Single-Cell RNA-Seq. Cell. 2018 May;173(6):1520-1534.e20.
47. Brooks W. The life history of the Hydromedusae: a discussion of the origin of the medusae, and of the significance of metagenesis. Soc Nat Hist. 1886;3(12).
48. Marques AC, Collins AG. Cladistic analysis of Medusozoa and cnidarian evolution. Invertebr Biol. 2005 May 11;123(1):23–42.
49. Salvini-Plawen LV. On the origin and evolution of the lower Metazoa. J Zool Syst Evol Res. 2009 Apr 27;16(1):40–87.
50. Li X, Ma X, Chen X, Wang T, Liu Q, Wang Y, et al. The medusa of Aurelia coerulea is similar to its polyp in molecular composition and different from the medusa of Stomolophus meleagris in toxicity. Toxicon. 2022 Apr;210:89–99.
51. Brekhman V, Malik A, Haas B, Sher N, Lotan T. Transcriptome profiling of the dynamic life cycle of the scypohozoan jellyfish Aurelia aurita. BMC Genomics. 2015 Dec;16(1):74.
52. The UniProt Consortium, Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, et al. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023 Jan 6;51(D1):D523–31.
53. Bertrand N, Castro DS, Guillemot F. Proneural genes and the specification of neural cell types. Nat Rev Neurosci. 2002 Jul;3(7):517–30.

54. Masoudi N, Tavazoie S, Glenwinkel L, Ryu L, Kim K, Hobert O. Masoudi N, Tavazoie S, Glenwinkel L, Ryu L, Kim K, Hobert O. Unconventional function of an achaete-scute homolog as a terminal selector of nociceptive neuron identity. PLOS Biology. 2018;16. Doi:10.1371/journal.pbio.2004979. PLOS Biol. (2018):16.
55. Layden MJ, Boekhout M, Martindale MQ. *Nematostella vectensis achaete-scute* homolog *NvashA* regulates embryonic ectodermal neurogenesis and represents an ancient component of the metazoan neural specification pathway. Development. 2012 Mar 1;139(5):1013–22.
56. Tournière O, Dolan D, Richards GS, Sunagar K, Columbus-Shenkar YY, Moran Y, et al. NvPOU4/Brain3 Functions as a Terminal Selector Gene in the Nervous System of the Cnidarian Nematostella vectensis. Cell Rep. 2020 Mar;30(13):4473-4489.e5.
57. Babonis LS, Enjolras C, Reft AJ, Foster BM, Hugosson F, Ryan JF, et al. Single-cell atavism reveals an ancient mechanism of cell type diversification in a sea anemone. Nat Commun. 2023 Feb 16;14(1):885.
58. Chapman JA, Kirkness EF, Simakov O, Hampson SE, Mitros T, Weinmaier T, et al. The dynamic genome of Hydra. Nature. 2010 Mar;464(7288):592–6.
59. Hartl M, Mitterstiller AM, Valovka T, Breuker K, Hobmayer B, Bister K. Stem cell-specific activation of an ancestral *myc* protooncogene with conserved basic functions in the early metazoan *Hydra*. Proc Natl Acad Sci. 2010 Mar 2;107(9):4051–6.
60. Hartl M, Glasauer S, Valovka T, Breuker K, Hobmayer B, Bister K. *Hydra myc2* , a unique pre-bilaterian member of the *myc* gene family, is activated in cell proliferation and gametogenesis. Biol Open. 2014 May 15;3(5):397–407.
61. Hartl M, Glasauer S, Gufler S, Raffeiner A, Puglisi K, Breuker K, et al. Differential regulation of *myc* homologs by Wnt/β-Catenin signaling in the early metazoan *Hydra*. FEBS J. 2019 Jun;286(12):2295–310.
62. Pascual-Torner M, Carrero D, Pérez-Silva JG, Álvarez-Puente D, Roiz-Valle D, Bretones G, et al. Comparative genomics of mortal and immortal cnidarians unveils novel keys behind rejuvenation. Proc Natl Acad Sci. 2022 Sep 6;119(36):e2118763119.
63. Oakley TH, Plachetzki DC, Rivera AS. Furcation, field-splitting, and the evolutionary origins of novelty in arthropod photoreceptors. Arthropod Struct Dev. 2007 Dec;36(4):386–400.
64. Arendt D, Musser JM, Baker CVH, Bergman A, Cepko C, Erwin DH, et al. The origin and evolution of cell types. Nat Rev Genet. 2016 Dec;17(12):744–57.
65. Kin K, Chen ZH, Forbes G, Schaap P. Evolution of a novel cell type in Dictyostelia required gene duplication of a cudA-like transcription factor. Curr Biol. 2022 Jan;32(2):428-437.e4.
66. Müller P, Yanze N, Schmid V, Spring J. The Homeobox Gene Otx of the Jellyfish Podocoryne carnea: Role of a Head Gene in Striated Muscle and Evolution. Dev Biol. 1999 Dec;216(2):582–94.
67. Sawada H, Yokosawa H, Lambert CC. Sawada H, Yokosawa H, Lambert CC. The Biology of Ascidians. Springer. :215–8; chapter 244.
68. Utsumi N, Saiga H. Comparison of the Structure and Expression of otx Genes between Ciona intestinalis and Halocynthia roretzi. In: Sawada H, Yokosawa H, Lambert CC, editors. The Biology of Ascidians [Internet]. Tokyo: Springer Japan; 2001 [cited 2023 Aug 7]. p. 215–8. Available from: http://link.springer.com/10.1007/978-4-431-66982-1_34
69. Wang CLA. Caldesmon and Smooth-Muscle Regulation. Cell Biochem Biophys. 2001;35(3):275–88.

70. Webb RC. SMOOTH MUSCLE CONTRACTION AND RELAXATION. Adv Physiol Educ. 2003 Dec;27(4):201–6.
71. Brunet T, Fischer AH, Steinmetz PR, Lauri A, Bertucci P, Arendt D. The evolutionary origin of bilaterian smooth and striated myocytes. eLife. 2016 Dec 1;5:e19607.
72. Sitbon YH, Yadav S, Kazmierczak K, Szczesna-Cordary D. Insights into myosin regulatory and essential light chains: a focus on their roles in cardiac and skeletal muscle function, development and disease. J Muscle Res Cell Motil. 2020 Dec;41(4):313–27.
73. Chari T, Pachter L. The Specious Art of Single-Cell Genomics [Internet]. Genomics; 2021 Aug [cited 2023 Aug 7]. Available from: http://biorxiv.org/lookup/doi/10.1101/2021.08.25.457696
74. Mahadevaiah SK, Sangrithi MN, Hirota T, Turner JMA. A single-cell transcriptome atlas of marsupial embryogenesis and X inactivation. Nature. 2020 Oct 22;586(7830):612–7.
75. Qiu C, Cao J, Martin BK, Li T, Welsh IC, Srivatsan S, et al. Systematic reconstruction of cellular trajectories across mouse embryogenesis. Nat Genet. 2022 Mar;54(3):328–41.
76. Mittnenzweig M, Mayshar Y, Cheng S, Ben-Yair R, Hadas R, Rais Y, et al. A single-embryo, single-cell time-resolved model for mouse gastrulation. Cell. 2021 May;184(11):2825-2842.e22.
77. Salamanca-Díaz DA, Schulreich SM, Cole AG, Wanninger A. Single-Cell RNA Sequencing Atlas From a Bivalve Larva Enhances Classical Cell Lineage Studies. Front Ecol Evol. 2022 Jan 26;9:783984.
78. Cole AG, Hashimshony T, Du Z, Yanai I. Gene regulatory patterning codes in early cell fate specification of the C. elegans embryo [Internet]. elife; 2023 May [cited 2023 Aug 7]. Available from: https://elifesciences.org/reviewed-preprints/87099
79. Bosch TCG. Hydra and the evolution of stem cells. BioEssays. 2009 Apr;31(4):478–86.
80. Levin M, Anavy L, Cole AG, Winter E, Mostov N, Khair S, et al. The mid-developmental transition and the evolution of animal body plans. Nature. 2016 Mar 31;531(7596):637–41.
81. Scorrano S, Aglieri G, Boero F, Dawson MN, Piraino S. Unmasking *Aurelia* species in the Mediterranean Sea: an integrative morphometric and molecular approach. Zool J Linn Soc [Internet]. 2016 Oct [cited 2023 Jun 20]; Available from: http://doi.wiley.com/10.1111/zoj.12494
82. Paganos P, Voronov D, Musser JM, Arendt D, Arnone MI. Single-cell RNA sequencing of the Strongylocentrotus purpuratus larva reveals the blueprint of major cell types and nervous system of a non-chordate deuterostome. eLife. 2021 Nov 25;10:e70416.
83. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021 Jun;184(13):3573-3587.e29.
84. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018 May;36(5):411–20.
85. Linderman GC, Zhao J, Roulis M, Bielecki P, Flavell RA, Nadler B, et al. Zero-preserving imputation of single-cell RNA-seq data. Nat Commun. 2022 Jan 11;13(1):192.
86. Rio DC, Ares M, Hannon GJ, Nilsen TW. Purification of RNA Using TRIzol (TRI Reagent). Cold Spring Harb Protoc. 2010 Jun;2010(6):pdb.prot5439.
87. Sinigaglia C, Thiel D, Hejnol A, Houliston E, Leclère L. A safer, urea-based in situ hybridization method improves detection of gene expression in diverse animal species. Dev Biol. 2018 Feb;434(1):15–23.

## Acknowledgments

The authors acknowledge the LISC (life science computer cluster) for providing infrastructure necessary for the bioinformatics, and CIUS (cell imaging and ultrastructure research) for technical support, equipment, and expertise. We thank the staff of the Tiergarten Schönbrunn (Vienna Zoo: Anton Weissenbacher, Roland Halbauer) and Haus des Meeres (Vienna Aquarium: Daniel Abed-Navandi) for sharing their expertise in cnidarian culture, and the UCSC Cell Browser (Maximilian Haeussler and Brittney Wick) for formatting and hosting these data for interactive exploration.

**Funding**: This work was supported by grants P34970 (UT) and P31018 (AGC) from the Austrian Science Fund FWF.

**Author contributions**:

Conceptualization: OL, AGC, UT
Methodology: OL, JDM, BZ, AGC
Investigation: OL, SMJ, BZ, KJ, JK, JDM, AGC
Supervision: AGC, UT
Writing—original draft: OL, AGC
Writing—review & editing: AGC, UT

**Competing interests**: Authors declare that they have no competing interests.

**Data and materials availability**: Raw sequence data will be deposited in the GEO database (https://www.ncbi.nlm.nih.gov/geo/), and the analysed data matrix will be available on the UCSC Cell Browser (https://cells.ucsc.edu/). Scripts for analysing the data and generating the figures in this manuscript can be found on our GitHub page (https://github.com/technau/AureliaAtlas).

## Figure Captions

**Fig. 1 Single cell transcriptomic atlas covering the polyp-medusa life cycle transition demonstrates shifts in genome usage between key life history stages.** (**A**) Ten single cell transcriptomic libraries were generated from cell dissociations of sequential life history stages (6) or tissue compartments from a sub-adult medusa (4). (**B**) Distribution of genome use across the life cycle. Boxplot representing 100 replicate sampling of 1000 cells over all expressed (min. 10 reads) genes from each stage. inset: total gene models recovered across sequencing libraries. (**C**) Random collection of diverse cell morphologies from whole animal tissue dissociations, representing partial diversity of cells captured from cell suspensions. Scale bars = 10µm. (**D**) Differentially expressed genes across the dataset. (**E**) Over-represented GO-terms represented in the up-regulated gene set highlights prostaglandin synthesis in polyps and over abundance of muscle in the medusa. (**F**) Stage-specific up-regulated transcription factors show differential activity across the life history stages.

**Fig. 2 Integrated cell atlas permits identification of cell states across the life cycle.** (**A**) Distribution of cell sub-types across all life cycle stages as a fraction of all cells captured (upper) and plotted as a 2d UMAP with cells coloured according to their population identity (lower). (**B**) Relative numbers of differentiated neuro-glandular cell types across each life history stage and medusa tissue. Medusa-specific cell states are highlighted in yellow. (**C**) Expression of population-specific genes indicative of each major partition. (**D**) Differentially expressed transcription factors across all cell states in the atlas. (**E**) A subset of *chitinase*-positive cells co-express Cholinesterase. *left: Chitinase* positive cells are found within the gastric filaments of the medusa stage. *center:* Cells within the digestive gland cluster (purple) co-express both genes. *right:* Acetylcholine positive immunostaining (red) colocalizes within the gastric filaments surrounding the mouth (mo) of a developing ephyra.

**Fig. 3. Cnidocyte specification pathways indicate two principle capsule types.** (**A**) Cnidocyte trajectory and distribution. Inset: location of subset within the full dataset. (**B**) Dotplot showing sequential transcription factor activation along the trajectory. (**C**) The two principle sub-types have distinct toxin profiles. (**D**) Overview of the ephyra lappet stained with Phalloidin and a βTubulin antibody. (**E**) A close-up shows cnidocytes (arrows) labelled by the same βTubulin antibody located in the distal region of the lappet. (**F**) An in situ hybridization for the cnidocyte marker *CAB32-like* visualises cnidocytes at the corresponding region in the ephyra lappet.

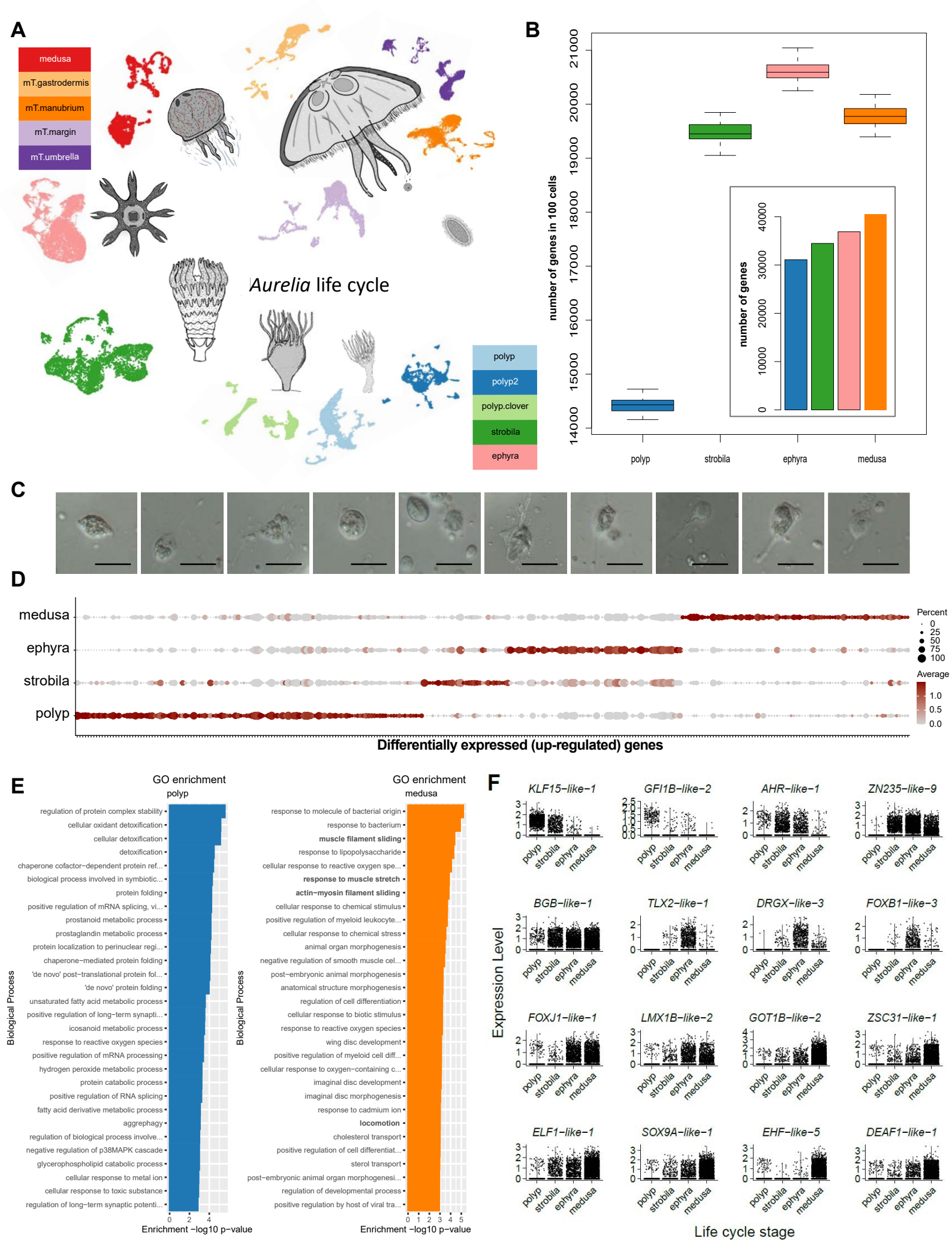
**Fig. 4 Neural cell complement includes two distinct classes, with medusa-specific subtypes in each.** (**A**) Distribution and characterization of neural complement, illustrated on an UMAP embedding of the full data (top) and neural subset (bottom) (**B**) Distribution histogram across all samples, medusa-specific cell clusters are highlighted with black outline. One transient cell state is present only in the metamorphic strobila/ephyra (yellow outline). (**C**) Two classes can be recognized by differential expression of transcription factors. Class 1 cells express *insulinoma* (*ins1a*) (orange left) and class 2 express *achaete-scute* (*ashA*) (green right). **(D**) Upregulated gene sets indicate both common class-markers and highlight unique neural sub-types. Class 2 neurons share expression of specific tubulins and calcium-binding genes whereas class 1 neurons share expression of sequestosome-like orthologs. (**E**) An anti-βTubulin antibody stains the motor nerve net of the ephyra. (**F**) Co-staining of anti-βTubulin and anti FMRFamide visualises cells that are positive for both antibodies. (**G**) FMRFamide positive cells are scattered at the periphery of the ephyra lappet. (**H**) Network of the rhopalia nervous system positive for βTubulin. (**I**) An overview of the ephyra lappet displays condensed regions of the βTubulin positive motor nerve net aligning with the radial canal of the lappet. (**J**) Single cells are stained with an anti-*ashA* in situ probe in proximity of the radial canal within the ephyra lappet. Lap = lappet, MNN = motor nerve net, DNN= diffuse nerve net, gan = ganglion, rho = rhopalia.

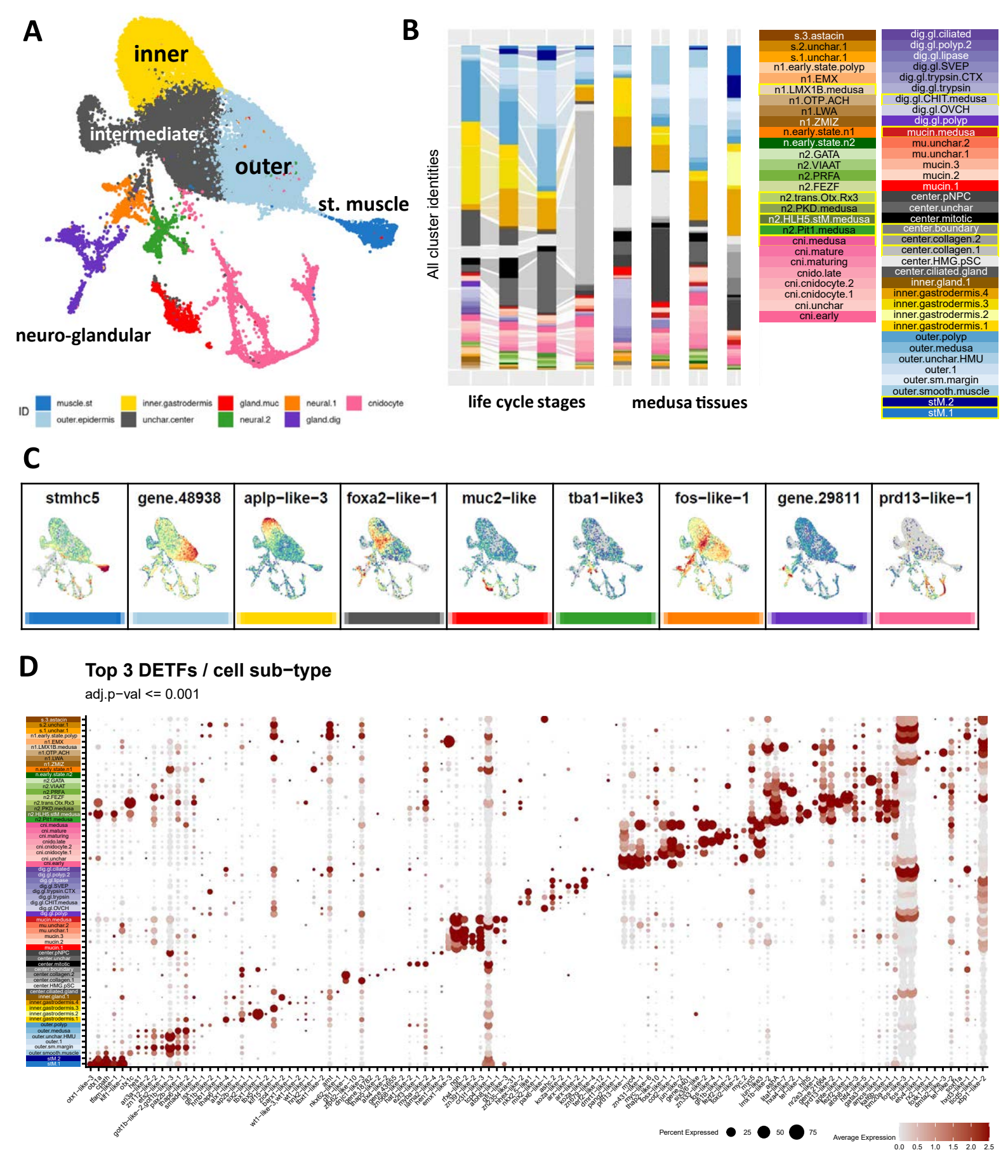
**Fig.5. Muscle territories in the *Aurelia* polyp.** (**A**) Schematic representation of the different smooth muscle territories in the polyp from a lateral view. (**B**) A reference UMAP of the muscle subset with highlighted smooth (yellow, turquoise) and striated (blue, red) muscle cell clusters. (**C-H**) Phalloidin stains actin-rich muscle fields in the animals. (**I**) Gene expression plots illustrating muscle gene expression across muscle subset. (**J-L**) In situ hybridization for two marker genes confirms the spatial location of the retractor muscles, tentacle muscles, and muscles at the oral disc of the polyp. te = tentacle, od = oral disc, rm = retractor muscle.

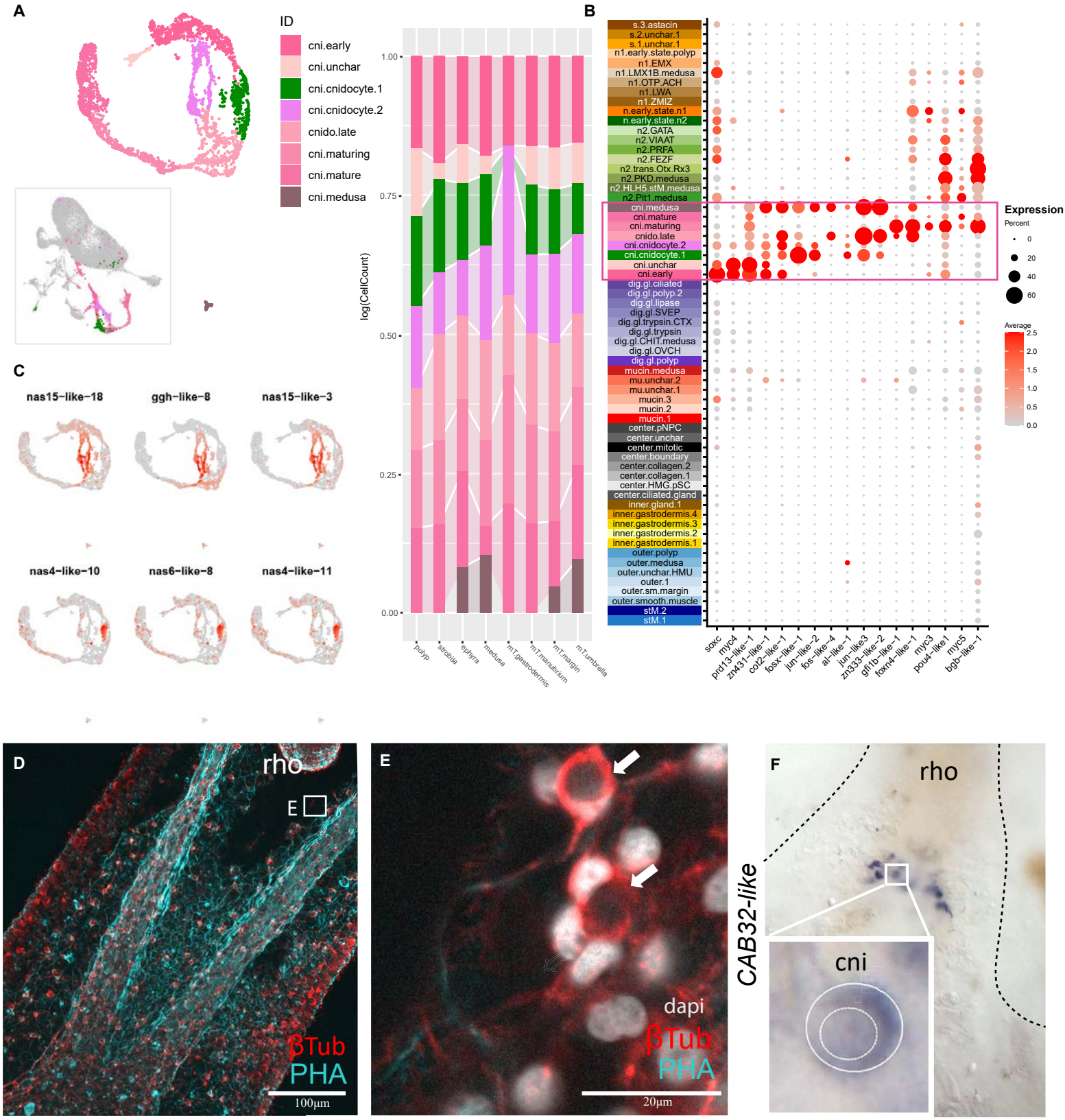
**Fig. 6. Establishment of striated muscle field during strobilation.** (**A**) Consecutive stages of strobilation from polyp (scyphistoma) to strobila and ephyra. (**B**) Foot region of the strobila stained with phalloidin showing polyp retractor muscles (arrowheads). (**C,D,E**) ISH for marker genes that label either smooth or striated muscles in the strobila. (**F**) shows gene expression plots for genes that are specific to smooth and striated muscle populations. (**G-O**) gene expression patterns of *stMHC1* in the advanced strobila. (**P-S**) gene expression pattern for the same gene in the detached ephyra. st.m = striated muscles, rm = retractor muscles.

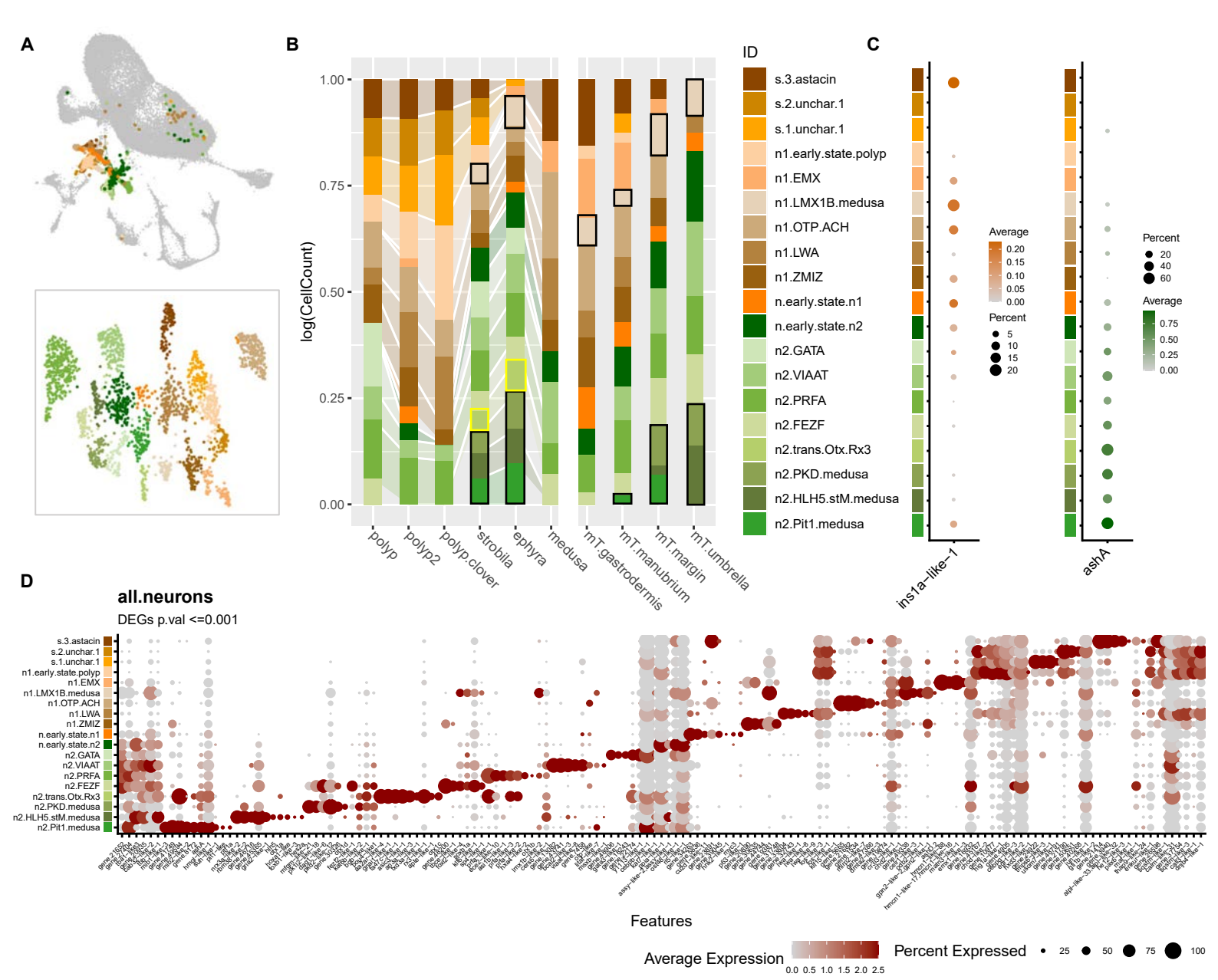
**Fig.7: Muscle territories in the ephyra.** (**A**) Ephyra detachment from the strobila. (**B**) Schematic of the ephyra body plan illustrating tissue anatomy (left) and striated muscle field (right). (**C-I**) Phalloidin labels the different muscle fields in the ephyra that are composed of either smooth, striated or a mixture of both muscle types. (**J-O**) depicts in situ hybridization for striated- (*stMHC5*), as well as all muscle marker genes (*stMHC1, MRLC2*) identified from single-cell RNAseq data. (**P**) An expression plot of striated muscle specific gene *stMHC5*. man = manubrium, gf = gastric filaments, rc = radial canal, st.m = striated muscle.

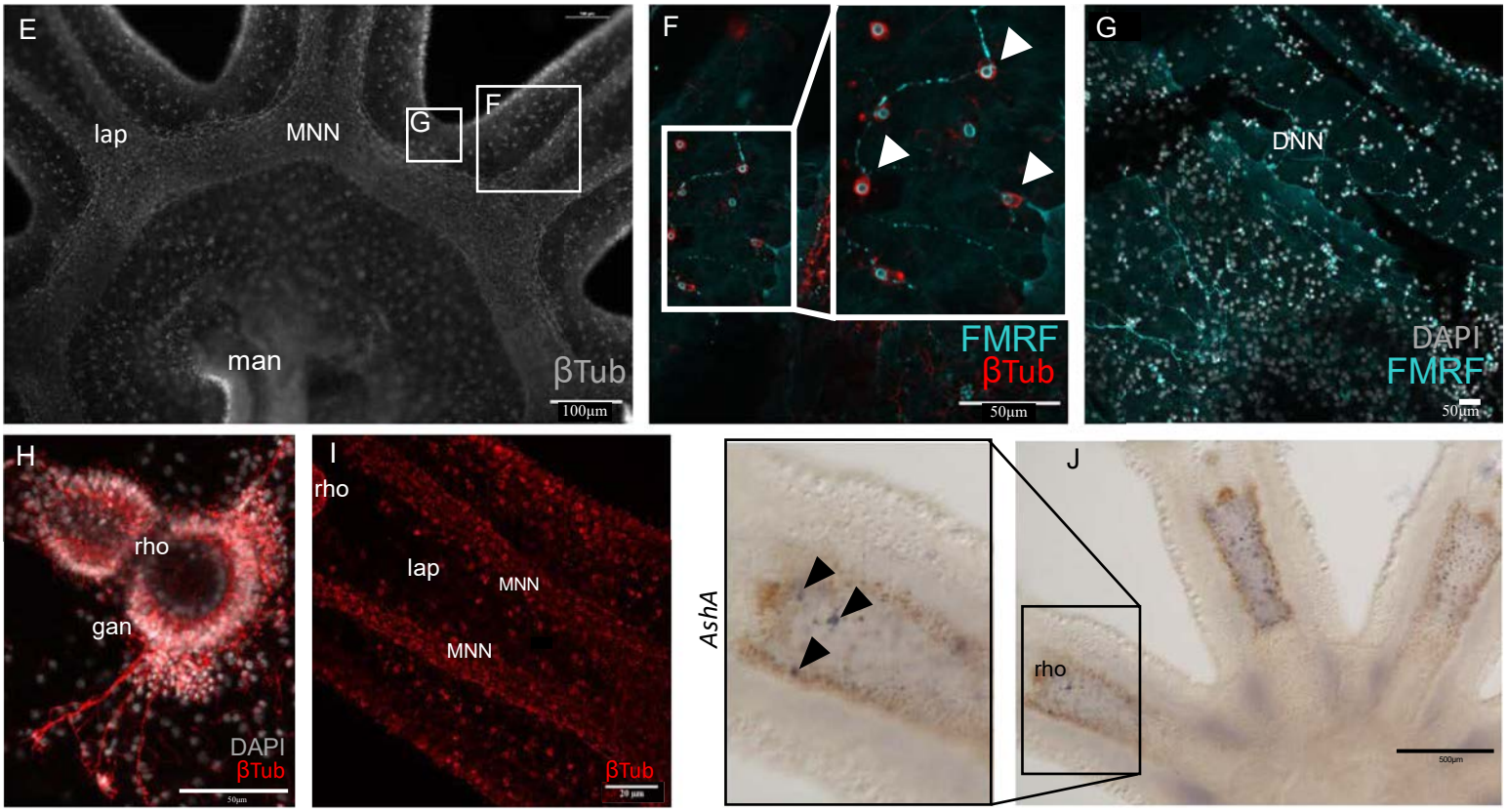
**Fig.8 Putative multipotent stem cell cluster demonstrates early initiation of many differentiation lineages**. (**A,B**) Differentiation potential as inferred with cytoTRACE indicates an increase in potential in the medusa-specific collagen/PIWI-expressing branch. (**A**) full dataset from all stages. MONOCLE3 estimation of differentiation trajectories (cyan). (**B**) intermediate cell cluster. Bottom: contributions from separate life cycle stages. Mitotic centers are highlighted in yellow. (**C,D**) Expression of specific differentiation markers suggests specific lineage relationships between cell types in the context of the entire dataset (**C**) and within the central cluster which includes the putative stem cell population (**D**).

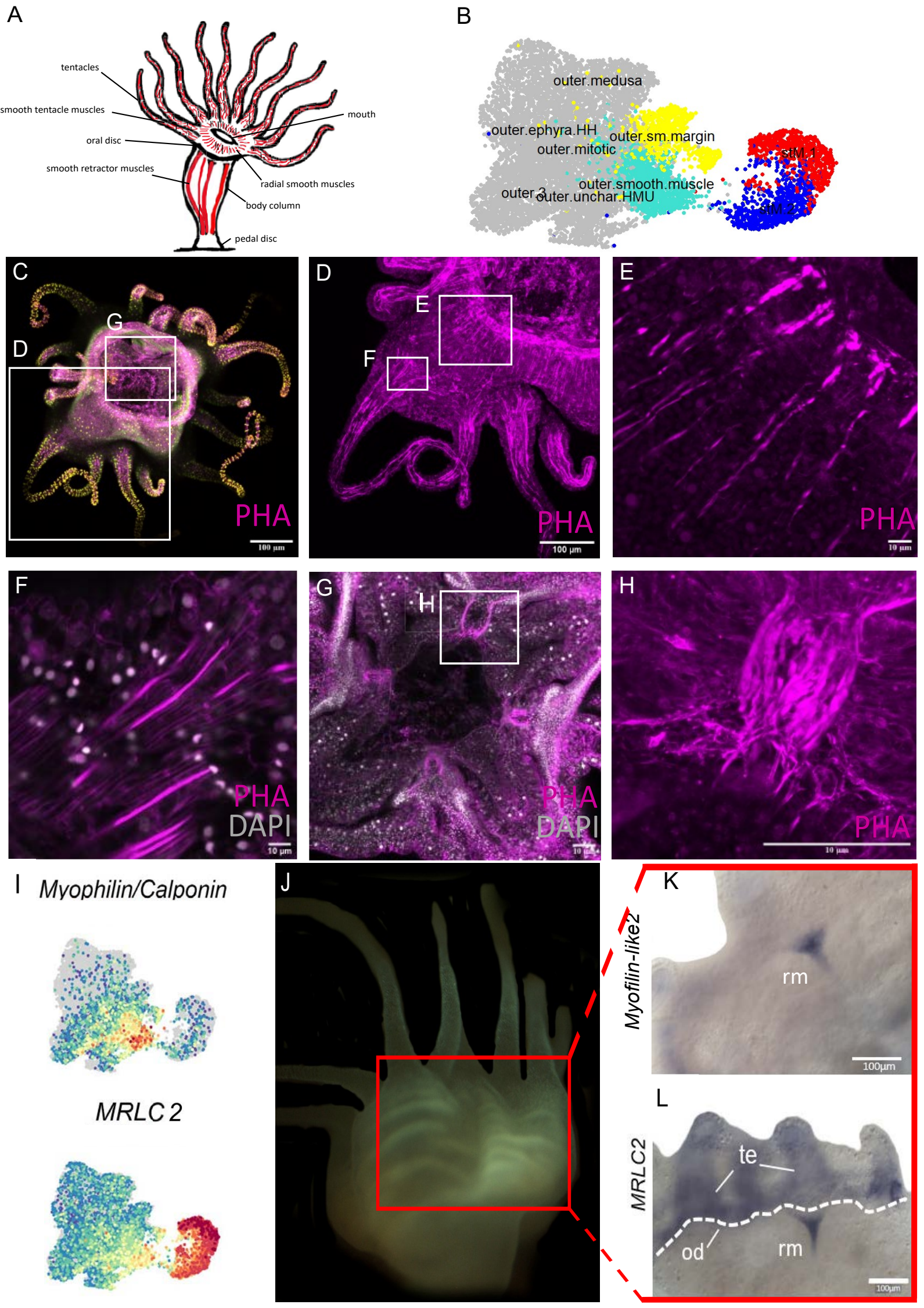

A
medusa
mT.gastrodermis
mT.manubrium
mT.margin
mT.umbrella
Aurelia life cycle
polyp
polyp2
polyp.clover
strobila
ephyra
B
number of genes in 100 cells
number of genes
polyp
strobila
ephyra
medusa
C
D
medusa
ephyra
strobila
polyp
Differentially expressed (up-regulated) genes
Percent
Average
E
GO enrichment
polyp
Biological Process
Enrichment −log10 p−value
GO enrichment
medusa
Biological Process
Enrichment −log10 p−value
F
KLF15−like−1
GFI1B−like−2
AHR−like−1
ZN235−like−9
BGB−like−1
TLX2−like−1
DRGX−like−3
FOXB1−like−3
FOXJ1−like−1
LMX1B−like−2
GOT1B−like−2
ZSC31−like−1
ELF1−like−1
SOX9A−like−1
EHF−like−5
DEAF1−like−1
Expression Level
polyp
strobila
ephyra
medusa
Life cycle stage

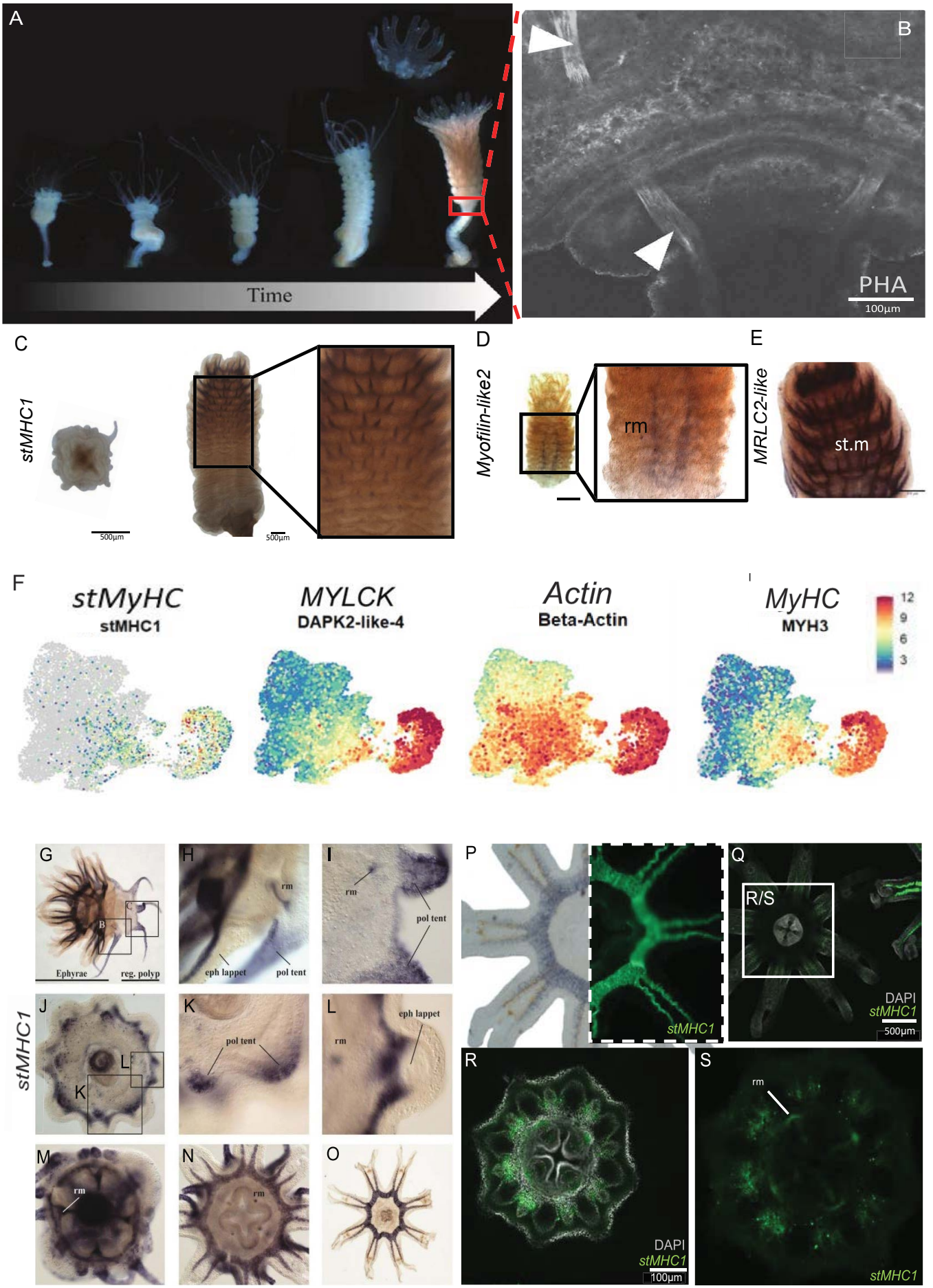

A
inner
intermediate
outer
st. muscle
neuro-glandular
ID
muscle.st
outer.epidermis
inner.gastrodermis
unchar.center
gland.muc
neural.2
neural.1
gland.dig
cnidocyte
B
All cluster identities
life cycle stages
medusa tissues
s.3.astacin
s.2.unchar.1
s.1.unchar.1
n1.early.state.polyp
n1.EMX
n1.LMX1B.medusa
n1.OTP.ACH
n1.LWA
n1.ZMIZ
n.early.state.n1
n.early.state.n2
n2.GATA
n2.VIAAT
n2.PRFA
n2.FEZF
n2.trans.Otx.Rx3
n2.PKD.medusa
n2.HLH5.stM.medusa
n2.Pit1.medusa
cni.medusa
cni.mature
cni.maturing
cnido.late
cni.cnidocyte.2
cni.cnidocyte.1
cni.unchar
cni.early
dig.gl.ciliated
dig.gl.polyp.2
dig.gl.lipase
dig.gl.SVEP
dig.gl.trypsin.CTX
dig.gl.trypsin
dig.gl.CHIT.medusa
dig.gl.OVCH
dig.gl.polyp
mucin.medusa
mu.unchar.2
mu.unchar.1
mucin.3
mucin.2
mucin.1
center.pNPC
center.unchar
center.mitotic
center.boundary
center.collagen.2
center.collagen.1
center.HMG.pSC
center.ciliated.gland
inner.gland.1
inner.gastrodermis.4
inner.gastrodermis.3
inner.gastrodermis.2
inner.gastrodermis.1
outer.polyp
outer.medusa
outer.unchar.HMU
outer.1
outer.sm.margin
outer.smooth.muscle
stM.2
stM.1
C
stmhc5
gene.48938
aplp-like-3
foxa2-like-1
muc2-like
tba1-like3
fos-like-1
gene.29811
prd13-like-1
D
Top 3 DETFs / cell sub-type
adj.p-val <= 0.001
Percent Expressed
25
50
75
Average Expression
0.0
0.5
1.0
1.5
2.0
2.5
E
500µm
Chitinase
Chitinase1
DAPI
ACH
acet.TUB
400µm
mo

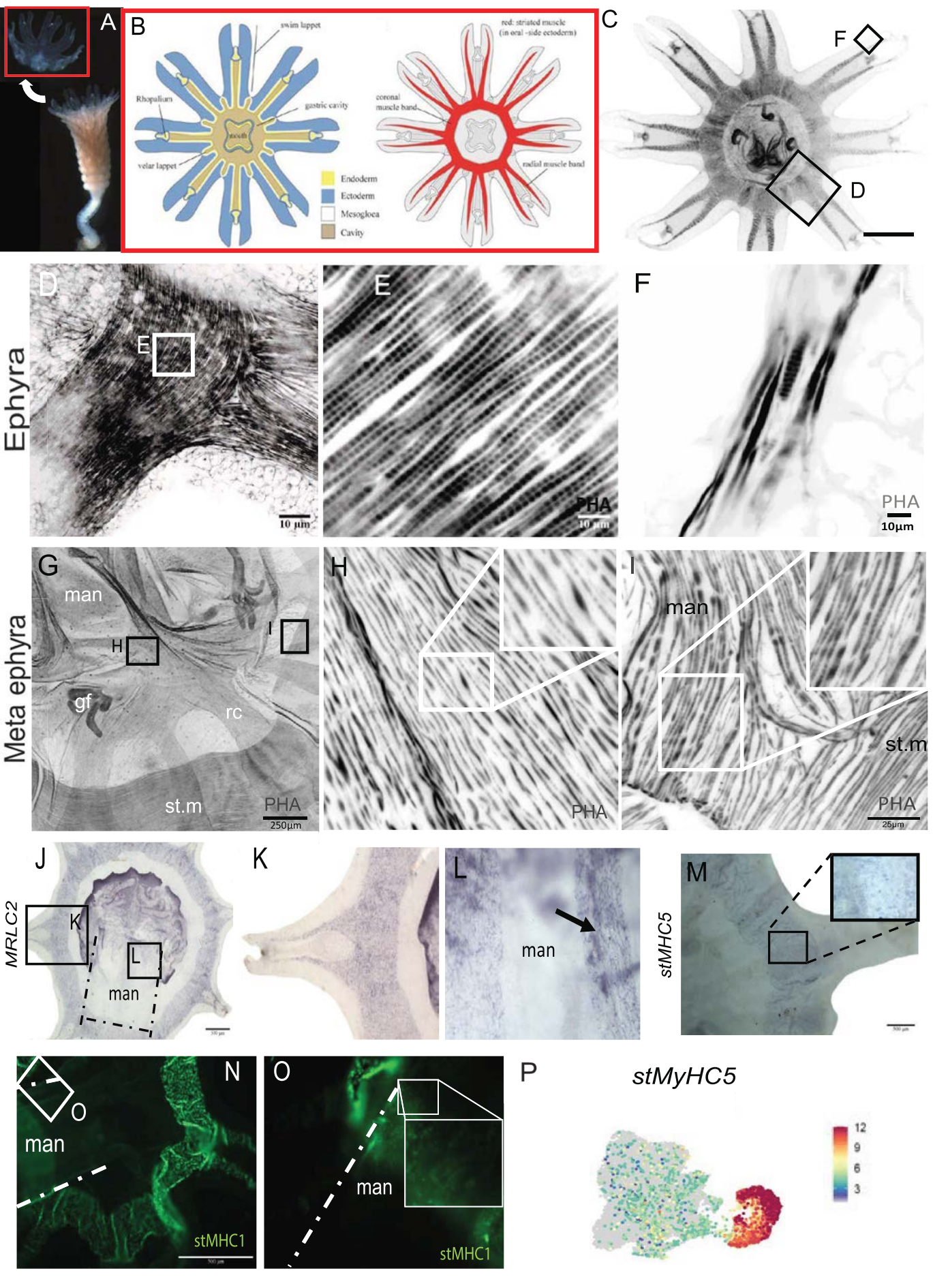

A
ID
cni.early
cni.unchar
cni.cnidocyte.1
cni.cnidocyte.2
cnido.late
cni.maturing
cni.mature
cni.medusa
log(CellCount)
1.00
0.75
0.50
0.25
0.00
polyp
strobila
ephyra
medusa
mT.gastrodermis
mT.manubrium
mT.margin
mT.umbrella
B
s.3.astacin
s.2.unchar.1
s.1.unchar.1
n1.early.state.polyp
n1.EMX
n1.LMX1B.medusa
n1.OTP.ACH
n1.LWA
n1.ZMIZ
n.early.state.n1
n.early.state.n2
n2.GATA
n2.VIAAT
n2.PRFA
n2.FEZF
n2.trans.Otx.Rx3
n2.PKD.medusa
n2.HLH5.stM.medusa
n2.Pit1.medusa
cni.medusa
cni.mature
cni.maturing
cnido.late
cni.cnidocyte.2
cni.cnidocyte.1
cni.unchar
cni.early
dig.gl.ciliated
dig.gl.polyp.2
dig.gl.lipase
dig.gl.SVEP
dig.gl.trypsin.CTX
dig.gl.trypsin
dig.gl.CHIT.medusa
dig.gl.OVCH
dig.gl.polyp
mucin.medusa
mu.unchar.2
mu.unchar.1
mucin.3
mucin.2
mucin.1
center.pNPC
center.unchar
center.mitotic
center.boundary
center.collagen.2
center.collagen.1
center.HMG.pSC
center.ciliated.gland
inner.gland.1
inner.gastrodermis.4
inner.gastrodermis.3
inner.gastrodermis.2
inner.gastrodermis.1
outer.polyp
outer.medusa
outer.unchar.HMU
outer.1
outer.sm.margin
outer.smooth.muscle
stM.2
stM.1
soxc
myc4
prd13-like-1
znf431-like-1
col2-like-1
fosx-like-1
jun-like-2
fos-like-4
al-like-1
jun-like3
zn333-like-2
gfi1b-like-1
foxn4-like-1
myc3
pou4-like1
myc5
bgb-like-1
Expression
Percent
0
20
40
60
Average
2.5
2.0
1.5
1.0
0.5
0.0
C
nas15-like-18
ggh-like-8
nas15-like-3
nas4-like-10
nas6-like-8
nas4-like-11
D
rho
E
βTub
PHA
100µm
E
dapi
βTub
PHA
20µm
F
rho
CAB32-like
cni

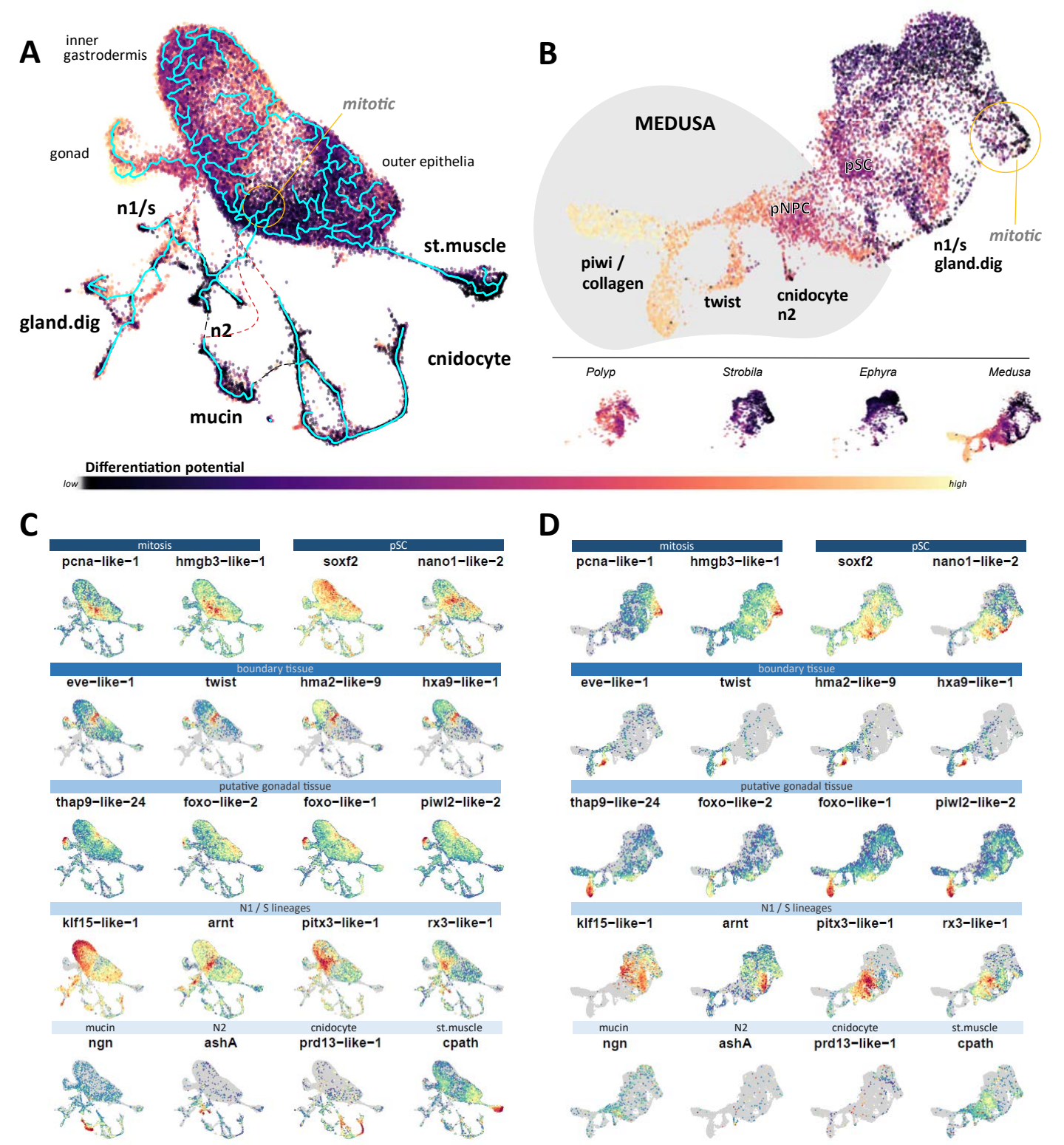

A
B
log(CellCount)
1.00
0.75
0.50
0.25
0.00
polyp
polyp2
polyp.clover
strobila
ephyra
medusa
mT.gastrodermis
mT.manubrium
mT.margin
mT.umbrella
ID
s.3.astacin
s.2.unchar.1
s.1.unchar.1
n1.early.state.polyp
n1.EMX
n1.LMX1B.medusa
n1.OTP.ACH
n1.LWA
n1.ZMIZ
n.early.state.n1
n.early.state.n2
n2.GATA
n2.VIAAT
n2.PRFA
n2.FEZF
n2.trans.Otx.Rx3
n2.PKD.medusa
n2.HLH5.stM.medusa
n2.Pit1.medusa
C
Average
0.20
0.15
0.10
0.05
0.00
Percent
5
10
15
20
ins1a-like-1
Percent
20
40
60
Average
0.75
0.50
0.25
0.00
ashA
D
all.neurons
DEGs p.val <=0.001
Features
Average Expression
0.0 0.5 1.0 1.5 2.0 2.5
Percent Expressed
25 50 75 100


E
lap
MNN
G
F
man
βTub
100µm
F
FMRF
βTub
50µm
G
DNN
DAPI
FMRF
50µm
H
rho
gan
DAPI
βTub
I
rho
lap
MNN
MNN
βTub
AshA
J
rho


A
tentacles
smooth tentacle muscles
oral disc
smooth retractor muscles
mouth
radial smooth muscles
body column
pedal disc
B
outer.medusa
outer.ephyra.HH
outer.mitotic
outer.sm.margin
stM.1
outer.smooth.muscle
outer.3
outer.unchar.HMU
stM.2
C
G
D
PHA
100 µm
D
E
F
PHA
100 µm
E
PHA
10 µm
F
PHA
DAPI
10 µm
G
H
PHA
DAPI
10 µm
H
PHA
10 µm
I
Myophilin/Calponin
MRLC 2
J
K
Myofilin-like2
rm
100µm
L
MRLC2
te
od
rm
100µm


A
Time
B
PHA
100µm
C
stMHC1
500µm
500µm
D
Myofilin-like2
rm
E
MRLC2-like
st.m
F
stMyHC
stMHC1
MYLCK
DAPK2-like-4
Actin
Beta-Actin
MyHC
MYH3
12
9
6
3
G
Ephyrae
reg. polyp
H
rm
eph lappet
pol tent
I
rm
pol tent
stMHC1
J
L
K
K
pol tent
L
eph lappet
rm
M
rm
N
rm
O
P
stMHC1
Q
R/S
DAPI
stMHC1
500µm
R
DAPI
stMHC1
100µm
S
rm
stMHC1


A
B
swim lappet
Rhopalium
gastric cavity
mouth
velar lappet
red: striated muscle
(in oral -side ectoderm)
coronal
muscle band
radial muscle band
Endoderm
Ectoderm
Mesogloea
Cavity
C
F
D
Ephyra
D
E
10 µm
E
PHA
10 µm
F
PHA
10µm
Meta ephyra
G
man
H
I
gf
rc
st.m
PHA
250µm
H
PHA
I
man
st.m
PHA
25µm
J
MRLC2
K
L
man
K
L
man
M
stMHC5
N
O
man
stMHC1
O
man
stMHC1
P
stMyHC5
12
9
6
3


A
inner gastrodermis
mitotic
gonad
outer epithelia
n1/s
st.muscle
gland.dig
n2
cnidocyte
mucin
B
MEDUSA
pSC
pNPC
mitotic
n1/s
gland.dig
piwi / collagen
twist
cnidocyte
n2
Polyp
Strobila
Ephyra
Medusa
Differentiation potential
low
high
C
mitosis
pSC
pcna−like−1
hmgb3−like−1
soxf2
nano1−like−2
boundary tissue
eve−like−1
twist
hma2−like−9
hxa9−like−1
putative gonadal tissue
thap9−like−24
foxo−like−2
foxo−like−1
piwl2−like−2
N1 / S lineages
klf15−like−1
arnt
pitx3−like−1
rx3−like−1
mucin
N2
cnidocyte
st.muscle
ngn
ashA
prd13−like−1
cpath
D
mitosis
pSC
pcna−like−1
hmgb3−like−1
soxf2
nano1−like−2
boundary tissue
eve−like−1
twist
hma2−like−9
hxa9−like−1
putative gonadal tissue
thap9−like−24
foxo−like−2
foxo−like−1
piwl2−like−2
N1 / S lineages
klf15−like−1
arnt
pitx3−like−1
rx3−like−1
mucin
N2
cnidocyte
st.muscle
ngn
ashA
prd13−like−1
cpath