The Mediterranean mussel Mytilus galloprovincialis : a novel model for developmental studies in mollusks

Angelica Miglioli, Marion Tredez, Manon Boosten, Camille Sant, João Carvalho, Philippe Dru, Laura Canesi, Michael Schubert, Rémi Dumollard

The Mediterranean mussel, *Mytilus galloprovincialis*, a novel model for developmental studies in mollusks

Angelica Miglioli[1$†], Marion Tredez[1$], Manon Boosten[1,2], Camille Sant[1,2], João E. Carvalho[1], Philippe Dru[1], Laura Canesi[3], Michael Schubert[1*], and Rémi Dumollard[1*]

1- Laboratoire de Biologie du Développement de Villefranche-sur-Mer (LBDV), Institut de la Mer de Villefranche (IMEV), Sorbonne Université, CNRS, Villefranche-sur-Mer, France

2- Laboratoire d'Océanologie de Villefranche (LOV), Institut de la Mer de Villefranche (IMEV), Sorbonne Université, CNRS, Villefranche-sur-Mer, France

3- Università degli Studi di Genova, Dipartimento di Scienze della Terra dell Ambiente e della Vita (DISTAV), Genova, Italy

[$] Co-First authors

[*] Co-Last authors

[†] Corresponding author: angelica.miglioli@imev-mer.fr

Keywords: bivalve mollusks, embryonic and larval development, trochophore, shell field, transcriptome, hybridization chain reaction

Summary statement: Resources and techniques are described to establish the Mediterranean mussel *Mytilus galloprovincialis* as a novel model for developmental studies in mollusks and animal evolution.

Running title: A novel developmental model: the Mediterranean mussel

Abstract

A model organism in developmental biology is defined by its experimental amenability and by resources created for the model system by the scientific community. For the most powerful invertebrate models, the combination of both has already yielded a thorough understanding of developmental processes. However, the number of developmental model systems is still limited, and their phylogenetic distribution heavily biased. Members of one of the largest animal lineages, the Spiralia, for example, have long been neglected. In order to remedy this shortcoming, we produced a detailed developmental transcriptome for the bivalve mollusk *Mytilus galloprovincialis*, and expanded the list of experimental protocols available for this species. Our high-quality transcriptome allowed us to identify transcriptomic signatures of developmental progression and to perform a first comparison with another bivalve mollusk, the Pacific oyster *Crassostrea gigas*. To allow co-labelling studies, we optimized and combined protocols for immunohistochemistry and hybridization chain reaction to create high-resolution co-expression maps of developmental genes. The resources and protocols described here represent an enormous boost for the establishment of *Mytilus galloprovincialis* as an alternative model system in developmental biology.

1. Introduction

Historically, our understanding of the origin, mechanisms, and evolution of animal development has been predominantly derived from the study of three groups: insects, nematodes, and vertebrates (Müller and Grossniklaus, 2010). This resulted in a compendium of available tools, resources, and experimental protocols for these organisms that has turned them into extremely amenable and efficient models for the scientific community (Müller and Grossniklaus, 2010; Nigon and Félix, 2017; Stephenson and Metcalfe, 2013). In comparison, members of the Spiralia, a species-rich and morphologically diverse clade of protostomes characterized by spiral cleavage during early development, have so far only been scantly investigated and still lack powerful experimental models (Martín-Durán and Marlétaz, 2020; Martín-Zamora et al., 2023a). As almost half of all bilaterian phyla are spiralians, members of the spiralians are of particular interest for understanding the evolution of bilaterian body plans (Henry, 2014; Martín-Durán and Marlétaz, 2020; Martín-Zamora et al., 2023a, 2023b; Piovani et al., 2023). Similarly, morphological similarities between spiralian larvae have been used as taxonomic characters and comparisons between larval types have yielded highly influential hypotheses on animal evolution (Carrier et al., 2018; Martín-Zamora et al., 2023b; Piovani et al., 2023; Wu et al., 2019). The trochophore larva, for instance, is widely distributed amongst spiralians, and it has been suggested that it may have already been present in the last common ancestor of all bilaterian animals (Carrier et al., 2018; De Robertis and Tejeda-Muñoz, 2022; Martín-Zamora et al., 2023b; Nielsen, 2005, 2018). However, despite decades of studies, our understanding of the evolutionary origin of the trochophore larva as well as of the biological processes underpinning its ontogeny and transition to the adult body remains fragmentary (Carrier et al., 2018; De Robertis and Tejeda-Muñoz, 2022; Piovani et al., 2023; Wada et al., 2020; Wang et al., 2020; Xu et al., 2016).

Within spiralians, mollusks are the largest and anatomically most diverse phylum (Davison and Neiman, 2021; Yang et al., 2020). Yet, only a limited number of species is currently available for studying the mechanisms of their development (Davison and Neiman, 2021; Goulding and Lambert, 2016; Lyons and Henry, 2022). This limitation is mainly due to the difficulty of spawning and rearing mollusks in a laboratory setting and to the related complexity of setting up experimental protocols with only limited access to live material (Davison and Neiman, 2021; Yang et al., 2020). Several mollusk species have nonetheless been highlighted in terms of their potential as models systems in diverse fields of biology, including bivalves of the order Mytilida (Davison and Neiman, 2021).

Based on this observation, the current work provides an ensemble of updated protocols and novel resources to significantly facilitate access to and work with embryos and larvae of a marine mytilid bivalve, the Mediterranean mussel *Mytilus galloprovincialis*.

M. galloprovincialis is an established model in environmental studies, and a species of great ecological and economic interest with a worldwide distribution (Brzozowska et al., 2012; Daguin and Borsa, 2000; FAO, 2022). Large-scale larval cultures of *M. galloprovincialis* are easily set up in the laboratory, following standardized culture protocols by the International Standardization Organization (ISO) and the American Society for Testing Materials (ASTM) (ASTM E724-21, 2021; ISO 17244:2015, 2020). *M. galloprovincialis* is also the first metazoan with a published pangenome (Gerdol et al., 2020). The *M. galloprovincialis* genome is thus characterized by accessory genomic regions encoding up to 20,000 dispensable genes subject to presence-absence variation (Gerdol et al., 2020). This genomic architecture resembles to the open pan-genomes found, for example, in prokaryotes and non-metazoan eukaryotes (Gerdol et al., 2020). This feature of the *M. galloprovincialis* genome is of significant importance for researchers interested in genome evolution. The presence of detailed culture protocols and of a well-annotated genome make *M. galloprovincialis* a promising model organism for developmental studies. In addition, basic protocols for pharmacological treatments, quantification of shell biogenesis, quantitative polymerase chain reaction (qPCR) as well as immunohistochemistry and *in situ* hybridization have recently been described for developing *M. galloprovincialis* embryos and larvae (Miglioli et al., 2019, 2021a, 2021b).

In this study, we compiled a detailed developmental transcriptome and validated it with a standardized qPCR protocol. We further adapted, to *M. galloprovincialis* embryos and larvae, the previously described protocol for multi-color, fluorescent whole-mount *in situ* hybridization using the hybridization chain reaction (HCR) (Choi et al., 2014; Piovani et al., 2023) and successfully coupled this approach with immunohistochemistry. The developmental transcriptome of *M. galloprovincialis* allowed us to define major developmental periods as well as comparisons with those of another bivalve mollusk, the Pacific oyster *Crassostrea gigas*. The novel *in situ* hybridization and immunohistochemistry protocols yielded detailed information on the expression of genes marking developmental progression, for example, in the early shell field, apical sensory cells, muscles, ciliated epithelia, and peripheral mantle cells. Taken together, the results of this work represent a valuable resource for future work with *M. galloprovincialis* embryos and larvae, setting the stage for establishing this species as an emerging model system for studying the developmental biology of mollusks.

2. Results

2.1. Description of *Mytilus galloprovincialis* transcriptome libraries

The developmental transcriptome of *M. galloprovincialis* embryos and larvae consists of 15 developmental stages (*i.e.*, a total of 30 libraries from two independent biological replicates). Samples were collected every 4 hours under standard culture conditions at 16°C (ASTM E724-21, 2021; ISO 17244:2015, 2020), from the unfertilized egg to 52 hours post fertilization (hpf), with an additional sampling time point at 72 hpf. The stages sampled for transcriptomic analysis were defined by morphology and acetylated α-tubulin immunoreactivity (Fig. 1A, Supplementary Figure 1). As previously described in other mytilids, the *M. galloprovincialis* zygote undergoes holoblastic, unequal cleavage (Kurita et al., 2009). The micromeres form at the animal pole of the embryo at 4 hpf (morula), and subsequent cleavage follows a spiral pattern (Kurita et al., 2009). By 8 hpf (blastula), the embryo reaches the blastula stage. By 12 hpf (gastrula), the macromeres at the vegetal pole start to invaginate, indicating the onset of gastrulation. At 16 hpf, (shell field - SF - gastrula), the stomodeum (or presumptive mouth) has moved anteriorly, and a rosette of dorsal ectodermal cells forms an invagination, indicating the onset of shell field ontogeny (Kniprath, 1980, 1981; Miglioli et al., 2019). At 20 hpf (trochophore 1), the columnar cells of the shell field thicken, creating a deeper invagination. At 24 hpf (trochophore 2), the larva has the stereotypical morphology of a trochophore, with a ventral stomodeum and a dorsal shell field that is still partially invaginated and secrets the organic matrix of the larval shell (Kniprath, 1980; Miglioli et al., 2019). By 28 hpf (trochophore 3) and 32 hpf (trochophore 4), the hinge region of the shell field is completely flat and the prototroch ciliary band surrounds the anterior hemisphere of the larval body. At 36 hpf (early veliger 1), a cavity appears on the dorsal side that connects the former stomodeum (that has now moved anteriorly) and the anal cavity on the posterior side. At 40 hpf (early veliger 2), the D-margin of the shell becomes clearly visible and only the velum (*i.e.*, the ventralized prototroch) of the larva is protruding. By 44 hpf (D-veliger 1), the velum is enclosed in the D-shaped shell. Thereafter, no major morphological changes are observable, except for the progressive outgrowth of the shell with respect to the mantle tissue along the D-border. In summary, the samples utilized for the transcriptome analysis include representatives of all major developmental stages of early bivalve development.

2.2. Transcriptome assembly and gene expression dynamics

More than 31 million paired-end clean reads per library were mapped on the published *M. galloprovincialis* genome (Gerdol et al., 2020), with an average input length of 200 bp (Supplementary Table 1). More than 85% of the reads were alignable, with 71% of them assigned to a unique locus (Supplementary Table 1). The output gene count matrix included 52,422 expressed genes (sum of gene raw counts in all libraries > 0) out of a total of 60,302 genes in all libraries (Supplementary Table 2). The minimum number of expressed genes (EGs) was found in unfertilized eggs (≅ 25,000 at 0 hpf), while the highest was found at the late D-veliger stage (≅ 40,000 at 72 hpf), indicating that at least 15,000 genes (corresponding to about 25% of the genome) are dynamically expressed during early development of *M. galloprovincialis* (Fig. 1B). The number of EGs increased drastically during the first 12 hours of development (*i.e.*, up to the gastrula), with the two replicates yielding similar results. Thereafter, a slow but steady increase of EGs was observed until the last sampling point. Principal component analysis (PCA) of the most variably expressed genes validated sample reproducibility (Fig. 1C). A majority of the variance was explained by the first principal component (78%), which was clearly correlated with developmental timing, lending support to the notion that developmental progression is driven by transcriptomic changes.

2.3. Differential expression analysis

Differential expression analysis was performed by pairwise comparison of samples of consecutive stages. The analysis identified 11,996 differentially expressed genes (DEGs) during *M. galloprovincialis* development (Supplementary File 2, Supplementary Table 3). More than 1,500 DEGs were already detected at the morula stage (4 hpf), indicating an early initiation of zygotic gene expression (Fig. 1D). The highest number of DEGs (≅ 5,700) was reached at the blastula stage (8 hpf) and was followed by a progressive decrease as development proceeded. We next annotated the DEGs based on their functional domains (Supplementary Figure 2). DEGs were screened for transcription factor domains (Supplementary Figure 2A), for domains known to be involved in shell formation (Supplementary Figure 2B), and for signaling receptors and membrane transporters (Supplementary Figure 2C). In total, 711 transcription factors with 65 different domains were annotated, with the most recurrent (> 20 genes) being zinc finger, homeodomain, helix-loop-helix, and fork head domains (Supplementary Figure 2A, Supplementary Table 4). With respect to shell formation, 779 genes were found to contain domains associated with shell matrix proteins and shell calcification (Supplementary Figure 2B) (Ramesh et al.,

2019). The most recurrent domains included calcium binding proteins (EF-hand and calponin), von Willebrand factors (WKA), extracellular matrix proteins (fibronectin), carbohydrate binding proteins (lectin and concanavalin), chitin binding proteins (CB), ion channels, and transporters (Supplementary Figure 2B). Regarding receptors, 225 DEGs were annotated as G protein-coupled receptors of the rhodopsin and secretin families, 92 as ephrin-A receptors, 23 belonged to the nuclear receptor superfamily, with other DEGs including scavenger, low density (LD) lipoprotein, and TGFβII receptors (Supplementary Figure 2C). As many as 125 DEGs were predicted to be membrane transporter-associated domains (Supplementary Figure 2C). Altogether, these results confirmed the relevance of the identified DEGs with respect to the major regulatory (*i.e.*, transcription factors) and ontogenetic (*e.g.*, shell matrix proteins) programs controlling early development of bivalve mollusks.

2.4. Transcriptome validation by qPCR and HCR *in situ* hybridization

The differential expression analyses were corroborated by qPCR carried out on RNA samples from both biological replicates. Four target genes were selected based on their transcriptomic profiles and transcript abundance. *Foxb2* (MGAL_10B093191) and *Wnt8a* (MGAL_10B085403) were selected because they are characterized by similar trends of their expression in the course of development (both peaking at the SF gastrula), but with significantly different mean expression values (high and low, respectively, for *Foxb2* and *Wnt8a*) (Fig. 1E, Supplementary Figure 3). *SYTL4* (MGAL_10B044489) and *SLC46A3* (MGAL_10B020966) were selected because they have highly dynamic expression profiles (Supplementary Figure 3). For all four genes, the qPCR experiments validated the expression profiles obtained from the developmental transcriptome (Fig. 1E, Supplementary Figure 3). To allow a spatiotemporal mapping of gene expression, we adapted the hybridization chain reaction (HCR) protocol to *M. galloprovincialis* embryos and larvae (Supplementary Figure 4) (Choi et al., 2014; Piovani et al., 2023). The specificity of the HCR protocol was initially validated by comparison with colorimetric *in situ* hybridization of genes with known developmental expression patterns: *Hox1, Tyrosinase,* and *Tektin* (Supplementary Figure 4A) (Miglioli et al., 2021b; Piovani et al., 2023). Thereafter, we assessed the efficiency of the HCR protocol using *Foxb2* and *Wnt8a* (Supplementary Figure 4B). While the expression of *Foxb2* (higher baseline expression) was conspicuous at all surveyed developmental stages, that of *Wnt8a* (lower baseline expression) was less distinctive (Supplementary Figure 4B). A notable exception was the

expression of *Wnt8a* in the SF gastrula (16 hpf), which is the developmental stage with the highest expression levels of both genes (Supplementary Figure 4B).

The HCR protocol proved extremely well-suited for co-expression analyses of developmental genes and thus allowed the parallel mapping of the expression patterns of *Foxb2* and *Wnt8a*: while *Foxb2* was detectable in the stomodeal and trochal area, *Wnt8a* was expressed in a discrete group of cells that will give rise to one of the posterior neural ganglions of the larva (Fig. 1F,G, Supplementary Figure 4B). These observations were consistent with the expression patterns of these two genes in embryos and larvae of other spiralians (Lartilot et al., 2002; Marlow et al., 2014; Tomer et al., 2010). To go even further, we coupled the HCR protocol with immunocytochemistry, using immunoreactivity against serotonin (5-HT) as a marker for the serotonergic system in the developing nervous system of mussel larvae (Miglioli et al., 2021b; Voronezhskaya et al., 2008). In addition to the identification of serotonergic neurons, co-labeling with serotonin also allowed a standardized orientation of the larvae, which significantly facilitated the interpretation of *in situ* hybridization signals in trochophore and D-veliger larvae of *M. galloprovincialis* (Fig. 1F, G). Taken together, our protocol coupling HCR *in situ* hybridization with immunocytochemistry allows high-resolution mapping of spatiotemporal gene expression dynamics in embryos and larvae of *M. galloprovincialis*.

Transcriptomic and anatomical characterization of *Mytilus galloprovincialis* early developmentTaking advantage of the developmental transcriptome and the HCR protocol, we next carried out a characterization of *M. galloprovincialis* early development at both the transcriptomic and morphological level. Hierarchical clustering of all transcriptome libraries was performed to identify statistically distinct developmental periods (adjusted-unbiased p-values, au ≥ 0.99) (Fig. 2A). The analysis divided the libraries in two main developmental clusters in line with developmental progression: one with embryonic libraries from egg to SF gastrula and one regrouping all the larval libraries from trochophore 1 to D-veliger 4 (Fig. 2A). Within the embryonic cluster, the egg and morula libraries were distinct from the blastula, gastrula, and SF gastrula libraries, indicating that embryonic development can be subdivided, on a transcriptomic level, into an early and a late embryo period (Fig. 2A). Within the larval cluster, the trochophore libraries formed two sub-clusters, one established by trochophore 1 and 2 and one by trochophore 3 and 4, suggesting a split into an early and a mature trochophore period (Fig. 2A). Likewise, the early veliger and D-veliger libraries were separated into two sub-clusters, establishing the

early veliger and D-veliger as two additional periods defining the early development of *M. galloprovincialis* (Fig. 2A).

To assess potential morphological differences in support of the developmental clusters identified in the transcriptome dataset, HCR *in* situ hybridization was performed on representative stages of each developmental cluster starting at the SF gastrula (16 hpf), the embryonic stage before the start of larval development (Fig. 2A,B, Supplementary Figure 5). *Tektin* (*Tek*), *Myosin heavy chain* (*Mhc*), and *Tyrosinase* (*Tyr*) were selected as tissue markers for, respectively, the ciliated epithelium, the muscular system, and the shell field, as previously described for developing bivalve mollusks (Dyachuk and Odintsova, 2009; Miglioli et al., 2019; Piovani et al., 2023; Salamanca-Díaz et al., 2022; Yang et al., 2017). In SF gastrulae (16 hpf), *Tek* was expressed in the anterior and posterior tufts, with an inconspicuous signal detectable throughout the embryo. In trochophore 1 larvae (20 hpf), *Tek* expression was still localized in the apical tuft and anterior hemisphere of the larva, with the signal expanding in trochophore 3 larvae (28 hpf), anteriorly, to the apical tuft and prototroch ciliary band and, posteriorly, to two islets of ciliated epithelium close to the stomodeum and the future anus. In early veliger 1 larvae (36 hpf), *Tek* was detectable in the prototroch on the ventral side of the larva and, in D-veliger 1 larvae (44 hpf), *Tek* expression marked the velum as well as the oral and anal ciliary tufts, indicating that the prototroch was completely internalized at this stage of larval development (Fig. 2B, C, Supplementary Figure 1).

In SF gastrulae (16 hpf), *Mhc* was expressed in two groups of muscle cells and this expression subsequently expanded to additional groups of muscle cells in trochophore 1 larvae (20 hpf). By trochophore 3 larva (28 hpf), *Mhc* expression indicated a significant expansion of muscle tissue, with three pairs of symmetric larval retractor muscles developing along the prototroch and towards the posterior ciliated epithelium. The muscles were connected dorsally by longitudinal and transversal commissures. In early veliger 1 larvae (36 hpf), the *Mhc* signal became even more complex, as the larval retractor muscles reached the ciliated epithelium, the larval protractor muscle was formed and elongated towards the prototroch, and the anterior adductor muscle was clearly identifiable. In D-veliger 1 larvae (44 hpf), the expression of *Mhc* revealed that the anterior protractor muscles were associated with the ciliated epithelium and that the anterior retractors were connected by a muscle ring at the dorsal side of the velum (Fig. 2B,C, Supplementary Figure 1). This indicates that the larva has gained the ability to retract the body inside the shell (Dyachuk and Odintsova, 2009).

In SF gastrulae (16 hpf), *Tyr* expression was detectable in two islets of cells in close proximity to the invaginated shell field and marked the evagination and expansion of the shell field in trochophore 1 larvae (20 hpf). By the trochophore 3 larva (28 hpf), *Tyr* expression was concentrated in islets of cells at the D-border of the larval shell, which formed a straight hinge indicating that the evagination of the shell field was completed. In early veliger 1 larvae (36 hpf), *Tyr* expression suggested that the larval shell was covering most of the larval body and, in D-veliger 1 larvae (44 hpf), expression of *Tyr* at the periphery of the larva indicated that the shell was enclosing the complete larval body (Fig. 2B,C, Supplementary Figure 1). Taken together, these observations support the developmental staging obtained by hierarchical clustering, as each developmental period was associated with marked changes in larval anatomy: early trochophores were characterized by an invaginating shell field and the differentiation of discrete islets of muscle cells; mature trochophores by a completely evaginated shell field, a differentiated prototroch, and developing retractor muscles; early veligers by the shell covering most of the larval body, a ventralized prototroch, and anterior adductor and protractor muscles; D-veligers by a shell that completely encloses the larval body and muscles that can retract the ciliated epithelium within the larval body (Fig. 2C).

2.5. Gene clusters and biological processes defining early development of *Mytilus galloprovincialis*

Gene soft clustering was performed on the whole transcriptome to identify potential gene sets and biological processes enriched at specific developmental stages (Fig. 3, Supplementary Figure 6). The analysis identified nine gene clusters that, with the exception of early veliger 1 and 2 stages, closely matched the developmental stage clusters defined above (Fig. 3, Supplementary Table 5). Early embryonic development was subdivided into a maternal and a zygotic gene cluster. For the maternal cluster (*i.e.*, the egg stage), enriched gene ontology terms were associated with protein phosphorylation, microtubule-based movement, and signal transduction (Fig. 3, Supplementary Figure 6). The zygotic cluster was enriched in genes involved in RNA splicing and DNA replication (Fig. 3, Supplementary Figure 6). Late embryonic development, *i.e.*, the blastula and gastrula gene clusters, were both enriched in DNA recombination genes (Fig. 3, Supplementary Figure 6). The gene cluster corresponding to the early trochophore was characterized by genes associated with intracellular protein transport, RNA processing, and regulation of transcription, while the mature trochophore gene cluster was associated with genes involved in signal transduction and apoptosis (Fig. 3, Supplementary Figure 6).

The gene clusters defining the D-veliger, the mature D-veliger, and the late D-veliger were generally enriched in gene ontology terms associated with the G-protein coupled receptor signalling pathway and transmembrane transport, in addition to translation (for the mature D-veliger) and carbohydrate metabolism (for the late D-veliger) (Fig. 3, Supplementary Figure 6). In summary, this analysis corroborates the notion that the early ontogeny of *M. galloprovincialis* includes six main developmental periods (early embryo, late embryo, early trochophore, mature trochophore, early veliger, and D-veliger), each characterized by a distinctive set of anatomical features and transcriptomic signatures.

2.6. Comparison with the developmental transcriptome of the Pacific oyster *Crassostrea gigas*

To assess whether similar transcriptomic dynamics occur during the early development of other bivalve mollusks, gene soft clustering was performed on the developmental transcriptome of the Pacific oyster *C. gigas* (Liu et al., 2021) and compared to that of *M. galloprovincialis* (Fig. 4). Nine of the ten soft clusters identified in the *C. gigas* transcriptome showed similar transcriptomic dynamics to those observed in *M. galloprovincialis,* covering the main developmental periods of gene expression, from the egg to the late D-veliger stage (Fig. 4A, Supplementary Table 6). Like our analysis in *M. galloprovincialis*, that of the *C. gigas* developmental transcriptome failed to identify a gene cluster for the early veliger libraries. Comparison of the orthologous gene clusters in *M. galloprovincialis* and *C. gigas* showed that the D-veliger (40%), late D-veliger (32%), mature D-veliger (31%), zygotic (31%), early trochophore (28%), and mature trochophore (27%) clusters shared the highest percentage of orthologous genes, with the maternal (13%), blastula (11%), and gastrula (9%) clusters being characterized by much lower percentages of orthologous genes (Fig. 4B, Supplementary Table 7). The orthologous gene sets included several transcription factors (such as fork-head, homeobox, and nuclear receptors) and signaling components (of the Wnt, TGF, and EGF pathways, for example), suggesting that the embryonic and early larval development of the two bivalves might rely on conserved gene regulatory programs (Supplementary Table 7). However, only 17% of the *M. galloprovincialis* genes had an ortholog in *C. gigas* (*i.e.*, 10,909 of about 65,000 genes). This might be explained by the overall number of genes annotated in the genomes of the two species, which is significantly higher in *M. galloprovincialis* than in *C. gigas* (about 60,000 genes versus about 30,000 genes) (Gerdol et al., 2020; Peñaloza et al., 2021), but nonetheless implies that the comparative analysis by gene orthology could only be applied to a subset of the genes characterizing the early development of *M.*

galloprovincialis. Taken together, these results suggest that the temporal dynamics of gene expression during early development of *M. galloprovincialis* and *C. gigas* are relatively similar. However, the low percentages of orthologous genes between most of the corresponding developmental clusters of the two bivalve species indicates that there might be a certain degree of divergence in the genetic programs controlling early development of *M. galloprovincialis* and *C. gigas*. To shed additional light on conserved aspects of early *M. galloprovincialis* and *C. gigas* development, we identified representative marker genes for the early trochophore, mature trochophore, and D-veliger periods. Genes were selected based on membership to the same developmental clusters in the two bivalve species and up-regulation in the respective mussel libraries. The criteria were matched by 48, 11, and 6 genes for, respectively, the early trochophore, the mature trochophore, and the D-veliger (Supplementary Table 7). We then selected two genes for each developmental period for HCR *in situ* hybridization (Fig. 4C): *Dmbt1* and *Dimm* for the early trochophore, *Cbp2* and *NR3A* for the mature trochophore, and *ColVI-like* and *EF-hand* for the D-veliger (Table 1). *Dmbt1* and *Dimm* were both expressed in the shell field of *M. galloprovincialis* at the early trochophore stage, *Cbp2* and *NR3A*, respectively, along the hinge axis and throughout the larval shell of the mature trochophore, and *ColVI-like* and *EF-hand*, respectively, along the D-border of the shell and in the ciliated epithelium of the velum at the D-veliger stage (Fig. 5). The expression of five of the six genes in tissues secreting the developing shell strongly indicates that the genetic program regulating larval shell formation might be, to some level, conserved between the two bivalve species.

	***Mg*ID**	***Cg*ID**	***Mg* annotation**
Early trochophore	MGAL_10B027895	CGI_10026985	*Deleted in malignant brain tumors 1 (Dmbt1)*
	MGAL_10B020733	CGI_10020332	*Dimmed (Dimm)*
Late trochophore	MGAL_10B067867	CGI_10014550	*Chitin binding peritrophin-A (Cbp2)*
	MGAL_10B094150	CGI_10024100	*Nuclear receptor 3A (NR3A)*
D-veliger	MGAL_10B028954	CGI_10015228	*EF-hand*
	MGAL_10B070218	CGI_10024466	*Collagen VI-like (ColVI-like)*

Table 1. List of representative orthologous genes in the mussel *Mytilus galloprovincialis* (*Mg*ID) and the Pacific oyster *Crassostrea gigas* (*Cg*ID) selected for *in situ* hybridization.

3. Discussion

3.1. Fine-grained transcriptomic approaches combined with thorough co-labelling techniques provide valuable insights into bivalve mollusk development

The results presented in this study provide the first thorough characterization of the early development of *M. galloprovincialis* at both a transcriptomic and a morphological level. The sampling schedule at 4-hour intervals allowed us to closely map the gene expression dynamics underpinning developmental progression. Analysis of the functional domains of DEGs highlighted the power of our transcriptome to delineate early bivalve development. Transcription factors known to act in the regulation of early metazoan development, such as homeobox and fork head genes, and cohorts of genes involved in early shell formation were thus highly represented among these DEGs (Guo et al., 2023; Lartilot et al., 2002; Morino et al., 2013; Paps et al., 2015; Ramesh et al., 2019; Setiamarga et al., 2013; Zhao et al., 2018; Zheng et al., 2019). In addition, the analysis revealed a strong enrichment of receptors and transporters, which is consistent with previous analyses in bivalve mollusks, indicating that these mediators of intercellular signaling might act in the regulation of larval morphogenesis, neurogenesis, and shell biogenesis (Guo et al., 2023; Liu et al., 2020; Marin et al., 2013; Miglioli, 2022; Miglioli et al., 2021a; Song et al., 2016).

To correlate the transcriptional and anatomical changes marking early *M. galloprovincialis* development, we set up a high-resolution, whole-mount protocol for multicolor *in situ* hybridization compatible with immunostaining approaches. In combination with the developmental transcriptome, this protocol allowed the parallel mapping of tissue marker genes previously described for other bivalve species in *M. galloprovincialis* embryos and larvae: *Tektin* (*Tek*) for the ciliated epithelium, *Myosin heavy chain* (*Mhc*) for the muscular system, and *Tyrosinase* (*Tyr*) for the tissue forming the shell (Dyachuk and Odintsova, 2009; Miglioli et al., 2019; Piovani et al., 2023; Salamanca-Díaz et al., 2022; Yang et al., 2017; Zhang et al., 2006). We were hence able to sketch the morphogenesis of *M. galloprovincialis* embryos and larvae, demonstrating that developmental progression at the transcriptomic level is closely matched by anatomical changes, at least at the level of ciliated epithelia, musculature, and shell growth. Although not novel *per se*, the combination of transcriptomic data and fluorescent whole-mount staining protocols has so far found only very limited use in surveys of mollusk development (Krasity et al., 2015;

Lopez-Anido et al., 2023; Piovani et al., 2023). The datasets and protocols presented here thus represent significant advances that will facilitate future studies on bivalve mollusk development and support the establishment of *M. galloprovincialis* as a model system in developmental biology.

3.2. Transcriptomic signatures and larval morphology define distinct periods of bivalve mollusk development

Hierarchical clustering of libraries and gene soft clustering revealed that the early development of *M. galloprovincialis* is divided into six different periods, each characterized by distinctive gene expression dynamics (Futschik and Carlisle, 2005; Leclère et al., 2019; Suzuki and Shimodaira, 2006). To assess if these developmental gene expression dynamics are conserved in other mollusks, we performed gene clustering analyses on the developmental transcriptome of the Pacific oyster *C. gigas.* Comparisons of the resulting *C. gigas* gene clusters with those obtained in *M. galloprovincialis* revealed similar temporal dynamics of gene expression that were, however, not strictly based on expression of orthologous gene sets. This is in line with previous findings indicating that temporal shifts in the use of common gene sets are common in bilaterian larvae and that this process might have facilitated the evolutionary diversification of larval forms and life cycles (Martín-Zamora et al., 2023b). Furthermore, trochophore development in both species was characterized by an early and a mature transcriptomic signature, which we were able to associate, in *M. galloprovincialis,* with distinctive morphological features and thus specific larval morphotypes. The fact that we were not able to associate the early veliger libraries of both species to specific gene clusters might suggest that this developmental period represents a transition phase from the transcriptomic programs of the trochophore to that of the D-veliger. Taken together, our work identified similar temporal gene expression dynamics in the two bivalve mollusk species that, albeit not driven by strictly orthologous gene sets, corroborated the existence of six conserved developmental periods during bivalve mollusk ontogeny, each characterized by distinctive transcriptomic and anatomical characters: early embryo, late embryo, early trochophore, mature trochophore, early veliger, and D-veliger. These results suggest that the staging of bivalve larval development should be carried out with great care to avoid generalized stage assignments that could lead to erroneous comparisons.

3.3. Marker gene expression reveals the progression of bivalve mollusk development

In this work, early development of *M. galloprovincialis* was morphologically characterized using expression of the tissue-specific marker genes *Tektin* (*Tek*), *Myosin heavy chain* (*Mhc*), and *Tyrosinase* (*Tyr*) respectively labeling the ciliated epithelium, musculature, and shell-forming tissue. *Tek* was selected as a marker of ciliated cells based on the single-cell atlas of the Pacific oyster *C. gigas* (Piovani et al., 2023), and *Mhc* has previously been used to characterize the muscle system in *M. coruscus* (Dyachuk and Odintsova, 2009) and is expressed in muscles of the quagga mussel *Dreissena rostriformis* (Salamanca-Díaz et al., 2022), while *Tyr* expression is known to be restricted to the outer edge of growing mantle tissue in both oysters and mussels (Miglioli et al., 2019; Yang et al., 2017). Co-labelling of developing *M. galloprovincialis* with these three marker genes allowed us to define key morphological events in each larval developmental period and to highlight the main anatomical differences between them. These results suggest that gene expression surveys using *Tek*, *Mhc*, and *Tyr* are an effective tool for assessing developmental progression of bivalve mollusks, one that could be generalized to homogenize ontogenetic staging systems within this animal clade.

Detailed comparisons between the developmental transcriptomes of *M. galloprovincialis* and *C. gigas* identified sets of highly conserved genes defining the bivalve early trochophore, mature trochophore, and D-veliger periods. Of these, we found that two early trochophore genes, *Dmbt1* and *Dimm*, were expressed in the developing shell field. Both genes have previously been described as regulators of terminal differentiation, *Dmbt1* in columnar cells and *Dimm* in neurosecretory cells (Liu et al., 2016; Takito and Al-Awqati, 2004). Expression of the two mature trochophore genes *Cbp2* (a chitin binding peritrophin) and *NR3A* (a nuclear receptor) was also associated with the developing shell, so was that of the D-veliger gene *ColVI-like*, which was detectable along the edge of the shell-forming tissue. Only the D-veliger gene *EF-hand* was not expressed in developing shell tissues but was found in the ciliated epithelium instead.

The expression dynamics of the conserved genes identified in our comparative analysis closely follow the stereotypical phases of shell formation in bivalve mollusks: terminal differentiation of shell field cells in the early trochophore, organic matrix growth in the mature trochophore, and calcification in the D-veliger (Kniprath, 1980, 1981). Cbp2 proteins, for example, are required for the assembly of the organic layer of the larval shell in the mature trochophore, while ColVI-like proteins are necessary for the successive formation of the nacre in the D-veliger (Marin et al., 2013; Zhao et al., 2018; Zheng et al., 2017) . These results strongly suggest that the genes identified by our comparative analysis are part of the molecular toolbox for biomineralization in bivalve mollusks

(Davison and Neiman, 2021; Marin et al., 2013; McDougall and Degnan, 2018) and indicate that the gene networks mediating shell formation are amongst the most conserved genetic components of early bivalve mollusk development. In addition, the implication of some of these genes (*Dimm* and *NR3A*) in neurogenesis in ecdysozoans and deuterostomes (Coumailleau et al., 2015; Liu et al., 2016) lends support to the notion that genes involved in nervous system development might have been co-opted for shell formation during the early evolutionary diversification of mollusks (McDougall and Degnan, 2018; Wollesen et al., 2017).

3.4. The Mediterranean mussel, *Mytilus galloprovincialis*, a novel model system for developmental studies of mollusks

The results reported in this study represent the first thorough transcriptomic analysis of the development of the Mediterranean mussel *M. galloprovincialis*. This work further provides a collection of protocols, for qPCR, immunohistochemistry, and HCR *in situ* hybridization that are essential for the validation and exploitation of transcriptomic data. We found that the combination of fluorescent immunohistochemistry and HCR *in situ* hybridization is particularly powerful for assessing the co-localization of developmental factors in developing mussel embryos and larvae. Potential applications of this transcriptome, and associated experimental protocols, include developmental studies, cross-species comparisons, and ecotoxicological surveys. Taking advantage of the developmental transcriptome and the novel protocols, we are the first to document the expression of a suite of genes illustrating the developmental progression of *M. galloprovincialis*, hence defining a set of markers that can be used in other bivalve mollusk systems to characterize early development. We are confident that this represents the first step for *M. galloprovincialis* to become a reference for developmental studies of bivalve mollusks. The next steps need to include the establishment of protocols for transgenesis to allow functional analyses, which are currently limited to pharmacological approaches (Davison and Neiman, 2021). Tackling this challenge should be a priority for the scientific community to further improve the experimental amenability of this promising animal model.

4. Materials and Methods

4.1. Spawning, fertilization, and rearing of embryos and larvae

Sexually mature *M. galloprovincialis* adults were harvested from the natural population in the bay of Villefranche-sur-Mer, France (43.682°N, 7.319°E) during the spawning seasons

of 2021, 2022, and 2023 (January through April). Animals were acclimated to and maintained in laboratory conditions, as previously described (Miglioli et al., 2021a). To induce gamete release, 30 adult mussels were exposed to a heat shock. Briefly, epibionts, dirt, and algae were scrubbed off the shells under running filtered seawater. Byssus and threads were carefully cut off. Clean mussels were placed on ice for 10 minutes and then immersed, in individual 200 mL containers and on a rocking shaker, in filtered seawater kept at 28°C until gamete release. Spawning mussels were immediately moved to filtered seawater at 16°C. Prior to fertilization, gametes were collected, washed, and maintained in Millipore-filtered seawater (using a 0.2 µm Millipore filter). Fertilization was carried out as previously described with a 1:10 ratio of eggs to sperm (Miglioli et al., 2021a). After 30 minutes, the fertilization success rate (*i.e.*, the ratio of the number of fertilized eggs over the total number eggs x 100) was established by microscopic observation. Parental pairs with a high fertilization success rate (> 90%) were selected for the following experiments. Two experimental replicates were set up, each consisting of a homogeneous pool of fertilized eggs from two independent parental pairs. Embryos from each replicate were distributed into 200 mL flasks for suspended cultures and brought to a density of 100 embryos per mL with Millipore-filtered seawater. A total of 14 flasks per replicate were prepared, corresponding to the pre-defined sampling strategy: in addition to sampling before fertilization, sample collections were carried out every 4 hours from 4 hpf to 52 hpf, with an additional sampling timepoint at 72 hpf. Embryo cultures were maintained at 16°C throughout the experiment (ASTM E724-21, 2021; ISO 17244:2015, 2020; Miglioli et al., 2021b). Samples for immunohistochemistry and *in situ* hybridization were produced following the same culture protocol. After filtration, larvae were fixed in 4% paraformaldehyde in 1x PBS overnight at 4°C, washed three times for 15 minutes in 1x PBS with 0.01% Tween 20, and then stored in 100% methanol at -20°C (Miglioli et al., 2021b).

4.2. Sampling, total RNA extraction, and high-throughput sequencing

RNA extraction and purification were carried out with the RNAqueous-Micro Kit (Invitrogen) and the RNeasy MinElute Cleanup Kit (Qiagen) following published protocols (Leclère et al., 2019). Unfertilized eggs were directly collected from spawning females, and embryos and larvae were sampled with a 60 µm mesh filter and concentrated in RNAse-free 1.5 mL tubes. The tubes were then placed on ice and gently centrifuged multiple times to remove the seawater from the samples. The pellet was immediately resuspended in lysis buffer, homogenized by energetic pipetting, and flash-frozen in liquid nitrogen. For

each sample, 10 µL of the 1.5 mL filtrate were placed in a separate tube and fixed overnight in 4% paraformaldehyde in 1x PBS at 4°C (Miglioli et al., 2021b). The fixed larvae were stained with Hoechst (Invitrogen) and imaged with a SP8 confocal microscope (Leica Microsystems) to validate proper developmental progression (Miglioli et al., 2021b). Lysed samples were stored at -80 °C until RNA extraction and purification (Leclère et al., 2019). RNA concentration and quality was assessed with a Nanodrop (Thermo Scientific) and a 2,100 Bioanalyzer (Agilent). Samples were shipped on dry ice to BGI (Hong Kong) for high-throughput sequencing. At least thirty million clean paired-end reads of 200 bp each were produced using the BGISEQ-500 Platform. Read cleaning and trimming was performed by BGI (Hong Kong). After sequencing, the raw reads were filtered. Data filtering included removal of adapter sequences, contaminations, and low-quality reads from the raw data. The entire read was deleted, if more than 28% matched the adapter sequence, if more than 40% had a quality value lower than 20 or if there were more than 3% of unidentifiable nucleotides in the read. This step was completed by the SOAP nuke software developed by BGI. Libraries were not subjected to further trimming (Dobin et al., 2013).

4.3. Genome-guided transcriptome assembly and differential expression analysis

Sequencing quality of the libraries was assessed with FastQC (v. 0.11.7) (Andrews, 2010). The libraries were mapped on the indexed *M. galloprovincialis* reference genome with STAR (v. 2.7.10a) using default parameters (Dobin et al., 2013; Gerdol et al., 2020). Read counts were obtained using StringTie (v2.2.1) in counting mode with the genome annotation file (Pertea et al., 2015, 2016). The resulting gene count matrix was exported using prepDE.py and processed in Rstudio (v. 4.2.2) with the DESeq2 (v. 1.24.0) package using default parameters and a design file indicating replicate number and developmental time for each library (Love et al., 2014). The dataset was then pre-filtered by keeping all genes with at least 10 reads in at least two replicated libraries to reduce the potential batch effect derived from the variability in gene presence-absence characterizing the species (Gerdol et al., 2020). The resulting matrix was normalized using the regularized logarithms of counts (DESeq2::rlog) and used to perform principal component analyses and heatmaps (Kolde, 2019). Stage-by-stage differential expression analysis was performed with DESeq2 using the pairwise comparison function, with the reference stage for comparisons relevelled for each contrast (Love et al., 2014). Pairwise comparisons were performed on samples of consecutive stages (0 hpf versus 4 hpf, 4 hpf versus 8 hpf, etc.).

DEGs were filtered by adjusted p-value (padj.cutoff < 0.05) and log2 fold change (lfc.cutoff > 0.58).

4.4. Developmental staging and gene clustering

To discriminate stages of *M. galloprovincialis* development from a transcriptomic perspective, bootstrapped hierarchical clustering with pvclust (v. 2.2-0) default parameters was performed on regularized logarithms (rlog) of the DESeq2 count matrix using the mean expression values of replicated libraries (Leclère et al., 2019; Love et al., 2014; Suzuki and Shimodaira, 2006). Soft clustering with the Mfuzz (v. 2.56.0) package was performed on mean expression values of replicated libraries normalized by transcripts per million to identify representative gene clusters of different developmental phases and potential transcriptomic waves (Futschik and Carlisle, 2005). The number of soft clusters was determined using the elbow method to the minimum centroid distance (Futschik and Carlisle, 2005). Genes were assigned to the cluster with the highest membership probability.

4.5. Functional annotation, gene ontologies, and enrichment analyses

The predicted protein sequences of the DEGs were annotated using BLASTp (v. 2.11.0) against the UniProt/SwissProt database, conserved domains were identified with hmmerscan from HMMER 3.3 against the Pfam database, and transcription factor domains were screened with AnimalTFDB (v2.0) (Finn et al., 2015, 2017; Jones et al., 2014; Mahram and Herbordt, 2015; Mistry et al., 2021; Sayers et al., 2021; Shen et al., 2023). Ontology enrichment analysis was performed for each gene soft cluster. The protein sequence corresponding to genes of interest were extracted from the *M. galloprovincialis* proteome and annotated with Gene Ontology (GO) identifiers using Interproscan (v. 95.0) (Finn et al., 2017; Gerdol et al., 2020; Jones et al., 2014). Enrichment analysis of Biological Processes (BP) was conducted with topGO R (v3.17). Significance of the enrichment was computed with Fisher's exact test (Alexa and Rahnenführer, 2009).

4.6. Crassostrea gigas gene clustering and comparison with *Mytilus galloprovincialis*

The transcripts per million matrix of the *C. gigas* developmental transcriptome from the egg to the D-veliger stage was downloaded from MolluscDB and gene soft clustering was

performed as previously described (Futschik and Carlisle, 2005; Liu et al., 2021). Orthologous genes were defined with DIAMOND (v0.9.22) using the rbhXpress script (Buchfink et al., 2014; El Hilali and Copley, 2023). Representative genes for different periods of early development conserved between *C. gigas* and *M. galloprovincialis* were selected based on membership to corresponding gene clusters and differential expression dynamics in the respective *M. galloprovincialis* transcriptome libraries.

4.7. Transcriptome validation by real time quantitative polymerase chain reaction

Complementary first strand DNAs (cDNAs) were synthesized from total RNA extracts of both biological replicates at all sample timepoints using the SuperScript VILO Kit (Invitrogen). The synthesis of cDNA was performed using random primers following the manufacturer's instructions: 10 minutes at 25°C, 60 minutes at 42°C, and 5 minutes at 85°C. To validate the differential gene expression analysis and the cluster analysis, four target genes were selected: *Foxb2* (MGAL_10B093191), *Wnt8a* (MGAL_10B085403), *SYTL4* (MGAL_10B044489), and *SLC46A3* (MGAL_10B020966). *EF-α1* was used for data normalization as one of the best-performing reference gene controls (Balbi et al., 2016). Primer pairs were designed within the open reading frame of each gene using Primer3 (Koressaar and Remm, 2007). Sequences, efficiency, and amplicon sizes are available in the supplementary material (Supplementary Table 8). The qPCR experiments were carried out for each set of primers, in experimental triplicates. The amplification was followed in real time using LightCycler 480 SYBR Green 1 master mix (Roche) in 45 cycles (initial 5 minutes at 95ºC, followed by 45 cycles of 20 seconds at 95ºC, 20 seconds at 60ºC, and 30 seconds at 72ºC) in a LightCycler 480 (Roche). Quantification was carried out in parallel for each of the selected genes and the reference gene. Quantification of expression at a given stage was determined in relation to the expression of the reference gene using the comparative CT method (Schmittgen and Livak, 2008). To compare qPCR and transcriptome results, expression of the target genes from both datasets were normalized with respect to their maximum value. Statistical differences in qPCR results were assessed using a Student's t-test (de Winter, 2013).

4.8. *In situ* hybridization using hybridization chain reaction and immunohistochemistry

For hybridization chain reaction (HCR)-based *in situ* hybridization experiments, hybridization probe sets were designed using a user-friendly Python interface

(https://github.com/rwnull/insitu_probe_generator) and synthesized by OligoPools (Twist Bioscience). Amplifiers (B2: 546 nm, B1: 647 nm, and B3: 514 nm) were purchased from Molecular Instruments (https://www.molecularinstruments.com/). Probe sets with relative amplifier pairings are summarized in Supplementary Table 9. Larval samples stored in methanol were re-hydrated with four washes of 5 minutes followed by a 10-minute incubation in 1x PBS with 0.01% Triton X-100 at room temperature. Permeabilization was carried out by a 30-minute incubation in the detergent solution (1% SDS, 0.005% Tween 20, 50 mM HCl pH7.5, 1 mM EDTA pH8.0, 0.15 M NaCl) at room temperature. After two washes of 1 minute and two washes of 5 minutes in 1x PBS with 0.01% Triton X-100 at room temperature, the specimens were placed in pre-hybridization for 30 minutes at 37°C with gentle shaking in hybridization buffer (HB) (30% formamide, 5x SSC, 9 mM citric acid pH6.0, 0.01% Tween 20, 2 mg heparin, 1x Denhardt's solution, 10% dextran sulphate). After addition of the probes (at a final concentration of 10 nM), hybridization was carried out overnight at 37°C. Following hybridization, 1.5 volumes of washing solution (30% formamide, 5x SSC, 9 mM citric acid, pH6.0, 0.01% Tween 20, 2 mg Heparin) with respect to the volume of the hybridization solution were added to the samples, and four washes of 15 minutes were subsequently performed with washing solution at 37°C. Specimens were then washed, at room temperature, twice for 5 minutes in 5x SSC with 0.01% Tween 20, once for 5 minutes in 1x SSC with 0.01% Tween 20, and 30 minutes in amplification buffer (5x SSC, 0.01% Tween 20, 10% dextran sulphate), before amplifiers were added (at a final concentration of 60 nM) to start the amplification chain reaction. Amplifiers were prepared as follows: the required volume of hairpins from each amplifier were placed in a 0.6 ml tube and incubated for 90 seconds at 95°C. After the heat shock, amplifiers were placed in the dark at room temperature for 30 minutes. The amplifiers were then resuspended in amplification buffer and distributed to the samples. Amplification was carried out at room temperature in the dark overnight. Excess hairpins were removed in two washes of 5 minutes and two washes of 30 minutes in 1x SSC with 0.01% Tween 20, followed by two washes of 5 minutes in 1x PBS with 0.01% Tween 20.

Co-labeling by immunocytochemistry with anti-serotonin (5-HT) rabbit antibody (1:10,000; #20080, Immunostar) and goat anti-rabbit Alexa Fluor 488 conjugated secondary antibody (1:500; #4412S, Cell Signaling Technology) was performed immediately after HCR amplification using a previously described protocol (Miglioli et al., 2021b). The same protocol was used to carry out labelling with anti-acetylated α-tubulin mouse antibody (1:10,000; #T6793, Sigma-Aldrich) and goat anti-mouse Alexa Fluor 488 conjugated

secondary antibody (1:500; #A-11001, Thermo Fisher Scientific). Specimens were then prepared for imaging by incubation in 1 µg/L Hoechst (Invitrogen) in 1x PBS for 30 minutes at room temperature. Following three washes of 5 minutes in 1x PBS with 0.01% Tween 20, specimens were mounted with the antifading agent CitiFluor AF1 (Agar Scientific) and imaged on a SP8 confocal microscope (Leica Microsystems), as previously described (Miglioli et al., 2021a). Images were analyzed with ImageJ (Schindelin et al., 2012), and fluorescent signals were corrected as follows: autofluorescence from tissue and shell was subtracted with “math”, the signal was homogenized with the “despeckle” or “smooth” functions, and z-stacks were merged by either “maximum intensity” or “sum slices”. Fluorescent HCR *in situ* hybridization signals were validated with colorimetric *in situ* hybridization experiments using genes with known developmental expression as positive controls and with no-probe HCR *in situ* hybridization assays as negative controls (Supplementary Table 10). Colorimetric *in situ* hybridization was performed as previously described for *M. galloprovincialis* embryos and larvae (Miglioli et al., 2019).

5. Acknowledgments

The authors are indebted to Jenifer Croce, Guy Lhomond, and Lucas Leclère for technical assistance and fruitful discussions. The authors would like to thank Sébastien Schaub of the Plateforme d'Imagerie par Microscopie (PIM) as well as Laurent Gilletta and Axel Duchene of the Service Aquariologie (SA) of the Institut de la Mer de Villefranche (IMEV), France, that is supported by EMBRC-France and funded by the Agence Nationale de la Recherche (ANR) (ANR-10-INBS-02). We thank the members of the Ascidian BioCell and EvoInSiDe teams at the Laboratoire de Biologie du Développement de Villefranche-sur-Mer, France, as well as the members of the team of Environmental Physiology at the Università Degli Studi Di Genova, Italy, for their support of this work.

6. Competing Interests

The authors declare that they have no competing financial interests or personal relationships that could have influenced the work reported in this paper.

7. Funding

This work was funded by grants from the CNRS DBM2021 program and the ANR (ANR-21-CE34-0006-02), both of which are French granting agencies, to Michael Schubert and Rémi Dumollard.

8. Data Availability

All libraries analyzed in this article have been deposited in the NCBI SRA database (Accession: PRJNA996031; ID: 996031).

9. References

Alexa, A. and Rahnenführer, J. (2009). Gene set enrichment analysis with topGO. *Bioconductor Improv.* **27**, 1–26.

Andrews, S. (2010). FASTQC. A quality control tool for high throughput sequence data.

ASTM E724-21 (2021). Standard guide for conducting static short-term chronic toxicity tests starting with embryos of four species of saltwater bivalve molluscs. *American Society for Testing and Materials (ASTM) International*.

Balbi, T., Franzellitti, S., Fabbri, R., Montagna, M., Fabbri, E. and Canesi, L. (2016). Impact of bisphenol A (BPA) on early embryo development in the marine mussel *Mytilus galloprovincialis*: Effects on gene transcription. *Environ. Pollut.* **218**, 996–1004.

Brzozowska, R., Sui, Z. and Kang, K. H. (2012). Testing the usability of sea mussel (*Mytilus sp.*) for the improvement of seawater quality–An experimental study. *Ecol. Eng.* **39**, 133–137.

Buchfink, B., Xie, C. and Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60.

Carrier, T. J., Reitzel, A. M. and Heyland, A. (2018). *Evolutionary ecology of marine invertebrate larvae*. Oxford University Press, Oxford, United Kingdom.

Choi, H. M. T., Beck, V. A. and Pierce, N. A. (2014). Next-generation *in situ* hybridization chain reaction: Higher gain, lower cost, greater durability. *ACS Nano* **8**, 4284–4294.

Coumailleau, P., Pellegrini, E., Adrio, F., Diotel, N., Cano-Nicolau, J., Nasri, A., Vaillant, C. and Kah, O. (2015). Aromatase, estrogen receptors and brain development in fish and amphibians. *Biochim. Biophys. Acta* **1849**, 152–162.

Daguin, C. and Borsa, P. (2000). Genetic relationships of *Mytilus galloprovincialis* Lmk. populations worldwide: Evidence from nuclear-DNA markers. In *Evolutionary Biology of the Bivalvia Vol. 177* (ed. E. M. Harper, J. D. Taylor and J. A. Crame), pp.389–397. London, UK: Geological Society, Special Publication.

Davison, A. and Neiman, M. (2021). Mobilizing molluscan models and genomes in biology. *Philos. Trans. R. Soc. B* **376**, 20200163.

De Robertis, E. M. and Tejeda-Muñoz, N. (2022). Evo-Devo of Urbilateria and its larval forms. *Dev. Biol.* **487**, 10–20.

de Winter, J. C. F. (2013). Using the student's t-test with extremely small sample sizes. *Pract. Assess. Res. Eval.* **18**, 10.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.

Dyachuk, V. and Odintsova, N. (2009). Development of the larval muscle system in the mussel *Mytilus trossulus* (Mollusca, Bivalvia). *Dev. Growth Differ.* **51**, 69–79.

El Hilali, S. and Copley R. R. (2023). macrosyntR: Drawing automatically ordered Oxford Grids from standard genomic files in R. *bioRxiv* 2023.01.26.525673.

FAO (2022). *The state of world fisheries and aquaculture 2022. Towards blue transformation.*

Rome, Italy.

Finn, R. D., Clements, J., Arndt, W., Miller, B. L., Wheeler, T. J., Schreiber, F., Bateman, A. and Eddy, S. R. (2015). HMMER web server: 2015 Update. *Nucleic Acids Res.* **43**, W30–W38.

Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., Chang, H. Y., Dosztanyi, Z., El-Gebali, S., Fraser, M., et al. (2017). InterPro in 2017–Beyond protein family and domain annotations. *Nucleic Acids Res.* **45**, D190–D199.

Futschik, M. E. and Carlisle, B. (2005). Noise-robust soft clustering of gene expression time-course data. *J. Bioinform. Comput. Biol.* **3**, 965–988.

Gerdol, M., Moreira, R., Cruz, F., Gómez-Garrido, J., Vlasova, A., Rosani, U., Venier, P., Naranjo-Ortiz, M. A., Murgarella, M., Greco, S., et al. (2020). Massive gene presence-absence variation shapes an open pan-genome in the Mediterranean mussel. *Genome Biol.* **21**, 275.

Goulding, M. Q. and Lambert, J. D. (2016). Mollusc models I. The snail *Ilyanassa*. *Curr. Opin. Genet. Dev.* **39**, 168–174.

Guo, X., Li, X., Zhao, F., Liu, D., Yang, Z., Li, M., Li, Y., Wei, H., Wang, H., Qin, Z., et al. (2023). Full-length transcriptome analysis provides insights into larval shell formation in *Mulinia lateralis*. *Front. Mar. Sci.* **9**, 1111241.

Henry, J. Q. (2014). Spiralian model systems. *Int. J. Dev. Biol.* **58**, 389–401.

ISO 17244:2015 (2020). Water quality–Determination of the toxicity of water samples on the embryo-larval development of Japanese oyster (*Crassostrea gigas*) and mussel (*Mytilus edulis* or *Mytilus galloprovincialis*). *International Standardization Organization (ISO)*.

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240.

Kniprath, E. (1980). Larval development of the shell and the shell gland in *Mytilus* (Bivalvia). *Wilhelm Roux Arch. Dev. Biol.* **188**, 201–204.

Kniprath, E. (1981). Ontogeny of the molluscan shell field: A review. *Zool. Scr.* **10**, 61–79.

Kolde, R. (2019). pheatmap: Pretty heatmaps. *R package version* 1.0.12.

Koressaar, T. and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**, 1289–1291.

Krasity, B. C., Troll, J. V., Lehnert, E. M., Hackett, K. T., Dillard, J. P., Apicella, M. A., Goldman, W. E., Weiss, J. P. and McFall-Ngaia, M. J. (2015). Structural and functional features of a developmentally regulated lipopolysaccharide-binding protein. *mBio* **6**, e01193-15.

Kurita, Y., Deguchi, R. and Wada, H. (2009). Early development and cleavage pattern of the Japanese purple mussel, *Septifer virgatus*. *Zoolog. Sci.* **26**, 814–820.

Lartilot, N., Le Gouar, M. and Adoutte, A. (2002). Expression pattern of *fork head* and *goosecoid* homologues in the mollusc *Patella vulgata* supports the ancestry of the anterior mesendoderm across Bilateria. *Dev. Genes Evol.* **212**, 551–561.

Leclère, L., Horin, C., Chevalier, S., Lapébie, P., Dru, P., Peron, S., Jager, M., Condamine, T., Pottin, K., Romano, S., et al. (2019). The genome of the jellyfish *Clytia hemisphaerica* and the evolution of the cnidarian life-cycle. *Nat. Ecol. Evol.* **3**, 801–810.

Liu, Y., Luo, J. and Nässel, D. R. (2016). The *Drosophila* transcription factor dimmed affects neuronal growth and differentiation in multiple ways depending on neuron type and developmental stage. *Front. Mol. Neurosci.* **9**, 97.

Liu, Z., Zhou, Z., Zhang, Y., Wang, L., Song, X., Wang, W., Zheng, Y., Zong, Y., Lv, Z. and Song, L. (2020). Ocean acidification inhibits initial shell formation of oyster larvae by suppressing the biosynthesis of serotonin and dopamine. *Sci. Total Environ.* **735**, 139469.

Liu, F., Li, Y., Yu, H., Zhang, L., Hu, J., Bao, Z. and Wang, S. (2021). MolluscDB: An integrated functional and evolutionary genomics database for the hyper-diverse animal phylum Mollusca. *Nucleic Acids Res.* **49**, D988–D997.

Lopez-Anido, R. N., Batzel, G. O., Ramirez, G., Goodheart, J. A., Wang, Y., Neal, S., and Lyons, D. C. (2023). Spatial-temporal expression analysis of lineage-restricted shell matrix proteins reveals shell field regionalization and distinct cell populations in the slipper snail *Crepidula atrasolea*. *bioRxiv* 2023.03.18.532128.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.

Lyons, D. C. and Henry, J. Q. (2022). Slipper snail tales: How *Crepidula fornicata* and *Crepidula atrasolea* became model molluscs. *Curr. Top. Dev. Biol.* **147**, 375–399.

Mahram, A. and Herbordt, M. C. (2015). NCBI BLASTP on high-performance reconfigurable computing systems. *ACM Trans. Reconfigurable Technol. Syst.* **7**, 33.

Marin, F., Marie, B., Hamada, S. B., Silva, P., Le Roy, N., Guichard, N., Wolf, S., Montagnani, C., Joubert, C., Piquemal, D., et al. (2013). 'Shellome': Proteins involved in mollusc shell biomineralization–Diversity, functions. In *Recent Advances in Pearl Research* (ed. S. Watabe, K. Maeyama, H. Nagasawa), pp.149–166. Tokyo, Japan: Terra Scientific Publishing Company.

Marlow, H., Tosches, M. A., Tomer, R., Steinmetz, P. R., Lauri, A., Larsson, T. and Arendt, D. (2014). Larval body patterning and apical organs are conserved in animal evolution. *BMC Biol.* **12**, 7.

Martín-Durán, J. M. and Marlétaz, F. (2020). Unravelling spiral cleavage. *Development* **147**, dev181081.

Martín-Zamora, F. M., Davies, B. E., Donnellan, R. D., Guynes, K. and Martín-Durán, J. M. (2023a). Functional genomics in Spiralia. *Brief. Funct. Genomics* **22**, 487–497.

Martín-Zamora, F. M., Liang, Y., Guynes, K., Carrillo-Baltodano, A. M., Davies, B. E., Donnellan, R. D., Tan, Y., Moggioli, G., Seudre, O., Tran, M., et al. (2023b). Annelid functional genomics reveal the origins of bilaterian life cycles. *Nature* **615**, 105–110.

McDougall, C. and Degnan, B. M. (2018). The evolution of mollusc shells. *Wiley Interdiscip. Rev. Dev. Biol.* **7**, e313.

Miglioli, A. (2022). Pathways of endocrine disruption in the larval development of the Mediterranean mussel *Mytilus galloprovincialis. Ph.D. thesis of the Università degli studi di*

Genova, Italy, and Sorbonne Université, France.

Miglioli, A., Dumollard, R., Balbi, T., Besnardeau, L. and Canesi, L. (2019). Characterization of the main steps in first shell formation in *Mytilus galloprovincialis*: Possible role of tyrosinase. *Proc. R. Soc. B* **286**, 20192043.

Miglioli, A., Balbi, T., Montagna, M., Dumollard, R. and Canesi, L. (2021a). Tetrabromobisphenol A acts a neurodevelopmental disruptor in early larval stages of *Mytilus galloprovincialis*. *Sci. Total Environ.* **793**, 148596.

Miglioli, A., Balbi, T., Besnardeau, L., Dumollard, R. and Canesi, L. (2021b). Bisphenol A interferes with first shell formation and development of the serotoninergic system in early larval stages of *Mytilus galloprovincialis*. *Sci. Total Environ.* **758**, 144003.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., et al. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419.

Morino, Y., Okada, K., Niikura, M., Honda, M., Satoh, N., and Wada, H. (2013). Genome-wide survey of genes encoding transcription factors in the Japanese pearl oyster, *Pinctada fucata*: I. homeobox genes. *Zoolog. Sci.* **30**, 851–857.

Müller, B. and Grossniklaus, U. (2010). Model organisms–A historical perspective. *J. Proteomics* **73**, 2054–2063.

Nielsen, C. (2005). Trochophora larvae: Cell-lineages, ciliary bands and body regions. 2. Other groups and general discussion. *J. Exp. Zool. B* **304**, 401–447.

Nielsen, C. (2018). Origin of the trochophora larva. *R. Soc. Open Sci.* **5**, 180042.

Nigon, V. M. and Félix, M. A. (2017). History of research on *C. elegans* and other free-living nematodes as model organisms. In *WormBook: The Online Review of* C. elegans *Biology* (ed. The *C. elegans* Research Community), pp.1–84. http://www.wormbook.org.

Paps, J., Xu, F., Zhang, G. and Holland, P. W. H. (2015). Reinforcing the egg-timer: Recruitment of novel Lophotrochozoa homeobox genes to early and late development in the Pacific oyster. *Genome Biol. Evol.* **7**, 677–688.

Peñaloza, C., Gutierrez, A. P., Eöry, L., Wang, S., Guo, X., Archibald, A. L., Bean, T. P. and Houston, R. D. (2021). A chromosome-level genome assembly for the Pacific oyster *Crassostrea gigas*. *Gigascience* **10**, giab020.

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T. and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295.

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667.

Piovani, L., Leite, D. J., Yañez Guerra**, L. A., Simpson, F., Musser, J. M., Salvador-Martínez, I., Marlétaz, F., Jékely, G. and Telford, M. J.** (2023). Single-cell atlases of two lophotrochozoan larvae highlight their complex evolutionary histories. *Sci. Adv.* **9**, eadg6034.

Ramesh, K., Yarra, T., Clark, M. S., John, U. and Melzner, F. (2019). Expression of calcification-

related ion transporters during blue mussel larval development. *Ecol. Evol.* **9**, 7157–7172.

Salamanca-Díaz, D. A., Schulreich, S. M., Cole, A. G. and Wanninger, A. (2022). Single-cell RNA sequencing atlas from a bivalve larva enhances classical cell lineage studies. *Front. Ecol. Evol.* **9**, 783984.

Sayers, E. W., Beck, J., Bolton, E. E., Bourexis, D., Brister, J. R., Canese, K., Comeau, D. C., Funk, K., Kim, S., Klimke, W., et al. (2021). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **49**, D10–D17.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682.

Schmittgen, T. D. and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* **3**, 1101–1108.

Setiamarga, D. H. E., Shimizu, K., Kuroda, J., Inamura, K., Sato, K., Isowa, Y., Ishikawa, M., Maeda, R., Nakano, T., Yamakawa, T., et al. (2013). An *in-silico* genomic survey to annotate genes coding for early development-relevant signaling molecules in the pearl oyster, *Pinctada fucata*. *Zoolog. Sci.* **30**, 877–888.

Shen, W. K., Chen, S. Y., Gan, Z. Q., Zhang, Y. Z., Yue, T., Chen, M. M., Xue, Y., Hu, H. and Guo, A. Y. (2023). AnimalTFDB 4.0: A comprehensive animal transcription factor database updated with variation and expression annotations. *Nucleic Acids Res.* **51**, D39–D45.

Song, H., Yu, Z. L., Sun, L. N., Gao, Y., Zhang, T. and Wang, H. Y. (2016). *De novo* transcriptome sequencing and analysis of *Rapana venosa* from six different developmental stages using Hi-seq 2500. *Comp. Biochem. Physiol. D* **17**, 48–57.

Stephenson, R. and Metcalfe, N. H. (2013). *Drosophila melanogaster*: A fly through its history and current use. *J. R. Coll. Physicians Edinb.* **43**, 70–75.

Suzuki, R. and Shimodaira, H. (2006). Pvclust: An R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**, 1540–1542.

Takito, J. and Al-Awqati, Q. (2004). Conversion of ES cells to columnar epithelia by hensin and to squamous epithelia by laminin. *J. Cell Biol.* **166**, 1093–1102.

Tomer, R., Denes, A. S., Tessmar-Raible, K. and Arendt, D. (2010). Profiling by image registration reveals common origin of annelid mushroom bodies and vertebrate pallium. *Cell* **142**, 800–809.

Voronezhskaya, E. E., Nezlin, L. P., Odintsova, N. A., Plummer, J. T. and Croll, R. P. (2008). Neuronal development in larval mussel *Mytilus trossulus* (Mollusca: Bivalvia). *Zoomorphology* **127**, 97–110.

Wada, H., Phuangphong, S., Hashimoto, N. and Nagai, K. (2020). Establishment of the novel bivalve body plan through modification of early developmental events in mollusks. *Evol. Dev.* **22**, 463–470.

Wang, J., Zhang, L., Lian, S., Qin, Z., Zhu, X., Dai, X., Huang, Z., Ke, C., Zhou, Z., Wei, J., et al. (2020). Evolutionary transcriptomics of metazoan biphasic life cycle supports a single intercalation origin of metazoan larvae. *Nat. Ecol. Evol.* **4**, 725–736.

Wollesen, T., Scherholz, M., Rodríguez Monje, S. V., Redl, E., Todt, C. and Wanninger, A. (2017). Brain regionalization genes are co-opted into shell field patterning in Mollusca. *Sci. Rep.* **7**, 5486.

Wu, L., Ferger, K. E. and Lambert, J. D. (2019). Gene expression does not support the developmental hourglass model in three animals with Spiralian development. *Mol. Biol. Evol.* **36**, 1373–1383.

Xu, F., Domazet-Lošo, T., Fan, D., Dunwell, T. L., Li, L., Fang, X. and Zhang, G. (2016). High expression of new genes in trochophore enlightening the ontogeny and evolution of trochozoans. *Sci. Rep.* **6**, 34664.

Yang, B., Pu, F., Li, L., You, W., Ke, C. and Feng, D. (2017). Functional analysis of a tyrosinase gene involved in early larval shell biogenesis in *Crassostrea angulata* and its response to ocean acidification. *Comp. Biochem. Physiol. B* **206**, 8–15.

Yang, Z., Zhang, L., Hu, J., Wang, J., Bao, Z. and Wang, S. (2020). The evo-devo of molluscs: Insights from a genomic perspective. *Evol. Dev.* **22**, 409–424.

Zhang, C., Xie, L., Huang, J., Chen, L. and Zhang, R. (2006). A novel putative tyrosinase involved in periostracum formation from the pearl oyster (*Pinctada fucata*). *Biochem. Biophys. Res. Commun.* **342**, 632–639.

Zhao, R., Takeuchi, T., Luo, Y. J., Ishikawa, A., Kobayashi, T., Koyanagi, R., Villar-Briones, A., Yamada, L., Sawada, H., Iwanaga, S., et al. (2018). Dual gene repertoires for larval and adult shells reveal molecules essential for molluscan shell formation. *Mol. Biol. Evol.* **35**, 2751–2761.

Zheng, Z., Du, X., Xiong, X., Jiao, Y., Deng, Y., Wang, Q. and Huang, R. (2017). *PmRunt* regulated by *Pm-miR-183* participates in nacre formation possibly through promoting the expression of *collagen VI-like* and *Nacrein* in pearl oyster *Pinctada martensii*. *PLoS One* **12**, e0178561.

Zheng, Z., Hao, R., Xiong, X., Jiao, Y., Deng, Y. and Du, X. (2019). Developmental characteristics of pearl oyster *Pinctada fucata martensii*: Insight into key molecular events related to shell formation, settlement and metamorphosis. *BMC Genomics* **20**, 122.

10. Figure Captions

Figure 1: Novel resources and techniques for the study of gene expression dynamics during early development of *Mytilus galloprovincialis*. A) Diagrams of the typical morphotype of each stage of the developmental transcriptome. Lateral views. B) Dot plot of the number of genes expressed in each library before DESeq2 normalization. The black dotted line represents the mean trend throughout development. C) Principal component analysis (PCA) of DESeq2 normalized libraries showing principal component (PC) axes 1 and 2. D) Histogram showing the number of up- (red) and down- (blue) regulated differentially expressed genes (DEGs) at each stage. Color code is according to log fold change values. E) Validation of transcriptomic analyses by qPCR with candidate genes *Fox2b* and *Wnt8a*. Left panel shows expression in the transcriptome in counts per million (CPM), right panels show expression obtained by qPCR; asterisks indicate statistically significant differential expression. F, G) Maximum projection superimpositions of *Foxb2* (yellow) and *Wnt8a* (red) fluorescent *in situ* hybridization with serotonin (5-HT) immunohistochemistry (green) and Hoechst (grey) in (F) trochophore 2 (24 hpf) and (G) D-veliger 2 (48 hpf) larvae. Lateral views. Scale bars: 20 µm. Abbreviations: hpf: hours post fertilization; A: anterior; P: posterior; D: dorsal; V: ventral; a: anus; ao/cg: apical organ/cerebral ganglion; at: apical tuft; h: hinge; m: mouth; pt: prototroch; SF, sf: shell field; st: stomodeum/presumptive mouth; ve: velum.

Figure 2: Transcriptomic and anatomical characterization of *Mytilus galloprovincialis* early development. A) Hierarchical clustering of average expression values of DEseq2 normalized libraries. Approximately unbiased (au) p-values are shown in red on each branch. Names for each cluster, corresponding to six developmental periods, were assigned according to developmental timepoint and overall morphology. B) Anatomical characterization of representative stages of the developmental periods identified by hierarchical clustering using hybridization chain reaction (HCR) *in situ* hybridization. Tissue marker genes used: *Tektin* (*Tek*) for ciliated epithelium, *Myosin heavy chain* (*Mhc*) for muscular system, and *Tyrosinase* (*Tyr*) for the developing larval shell. Lateral views. Scale bars: 20 µm. Additional developmental stages are shown in Supplementary Figure 5. C) Schematic representation of the main morphological differences between the developmental periods. Lateral views. Color code for developmental expression: red, *Tektin* (*Tek*); yellow, *Myosin heavy chain* (*Mhc*); green, *Tyrosinase* (*Tyr*). Abbreviations: a: anus; ad: anterior adductor muscle; at: apical tuft; h: hinge; lp: larval protractor muscle; lr^{1-4}: larval retractor muscles 1 through 4; m: mouth; mc:

muscle cells; pt: prototroch; SF, sf: shell field; st: stomodeum/presumptive mouth; ve: velum. Green and red arrowheads respectively highlight the dynamics of shell field and ciliated epithelium development.

Figure 3: Identification of gene sets characterizing early development of *Mytilus galloprovincialis*. Heatmap of soft clusters in scaled transcripts per million (z-scores) arranged by time of appearance during development. Row annotations identify the gene clusters, column annotation reports the hierarchical clusters. The top three gene ontology (GO) terms by adjusted p-value are shown as word clouds next to the respective gene cluster. Text color according to p-value (red: low; blue: high) and font size depending on number of significant genes per term. Abbreviations: SF: shell field; Troch.: Trochophore.

Figure 4: Comparisons between the developmental transcriptomes of the Mediterranean mussel, *Mytilus galloprovincialis*, and the Pacific oyster, *Crassostrea gigas*. A) Heatmap of soft clusters in scaled transcripts per million (z-scores) identified in the developmental transcriptome of *C. gigas* and arranged by time of appearance during development. Row annotations identify the gene clusters. B) Histogram displaying the percentage of orthologous genes between *M. galloprovincialis* and *C. gigas* soft clusters. C) Expression profiles of conserved period marker genes in the *M. galloprovincialis* developmental transcriptome by developmental period (early trochophore: green; mature trochophore: orange; D-veliger: red). Abbreviations: SF: shell field; Troch.: Trochophore.

Figure 5: Expression patterns of conserved period marker genes in *Mytilus galloprovincialis*. Left panel: expression pattern of the early trochophore genes *Dmbt1* and *Dimm* at SF gastrula (16 hpf) and trochophore 1 (20 hpf). Central panel: expression pattern of the mature trochophore genes *Cbp2* and *NR3A* at trochophore 2 (24 hpf) and trochophore 3 (28 hpf). Right panel: expression pattern of the D-veliger genes *ColVI-like* and *EF-hand* at early veliger 2 (40 hpf) and D-veliger 1 (44 hpf). Lateral views, except the panel showing *Dimm* expression at SF gastrula (16 hpf), which is a dorsal view. Abbreviations: D.: Development; a: anus; at: apical tuft; h: hinge; m: mouth; pt: prototroch; SF, sf: shell field; st: stomodeum/presumptive mouth; ve: velum. Scale bars: 20 µm.

A

Orientation

Morula 4 hpf

Blastula 8 hpf

Gastrula 12 hpf

SF Gastrula 16 hpf

Trochophore 1 20 hpf

Trochophore 2 24 hpf

Trochophore 3 28 hpf

Trochophore 4 32 hpf

Early Veliger 1 36 hpf

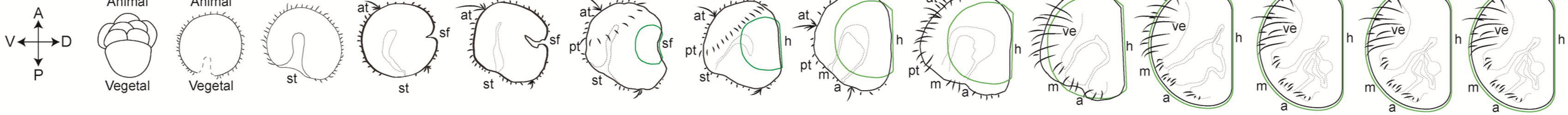
Early Veliger 2 40 hpf

D-Veliger 1 44 hpf

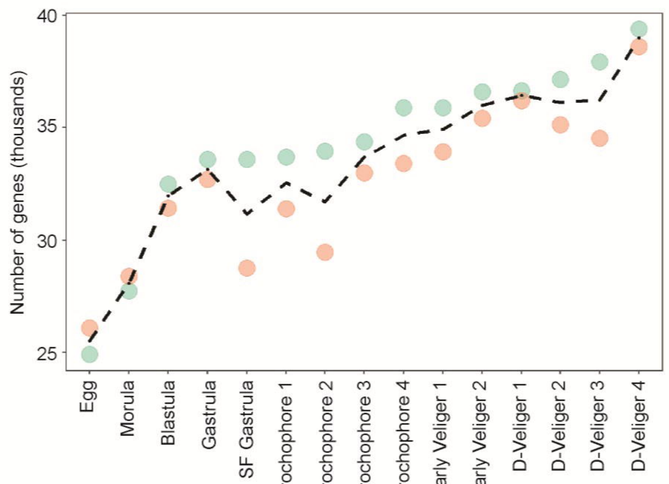
D-Veliger 2 48 hpf

D-Veliger 3 52 hpf

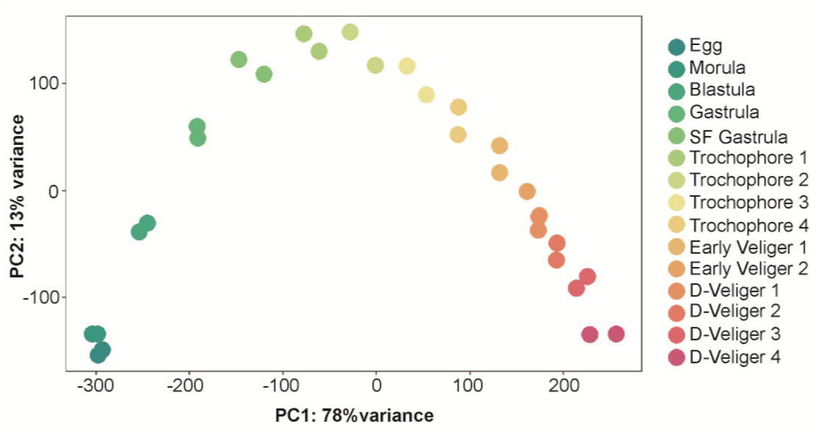
D-Veliger 4 72 hpf



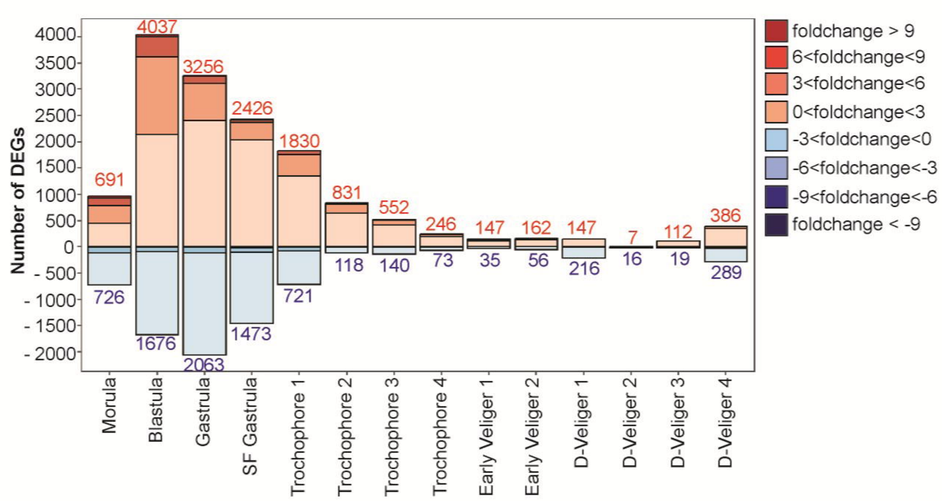
B



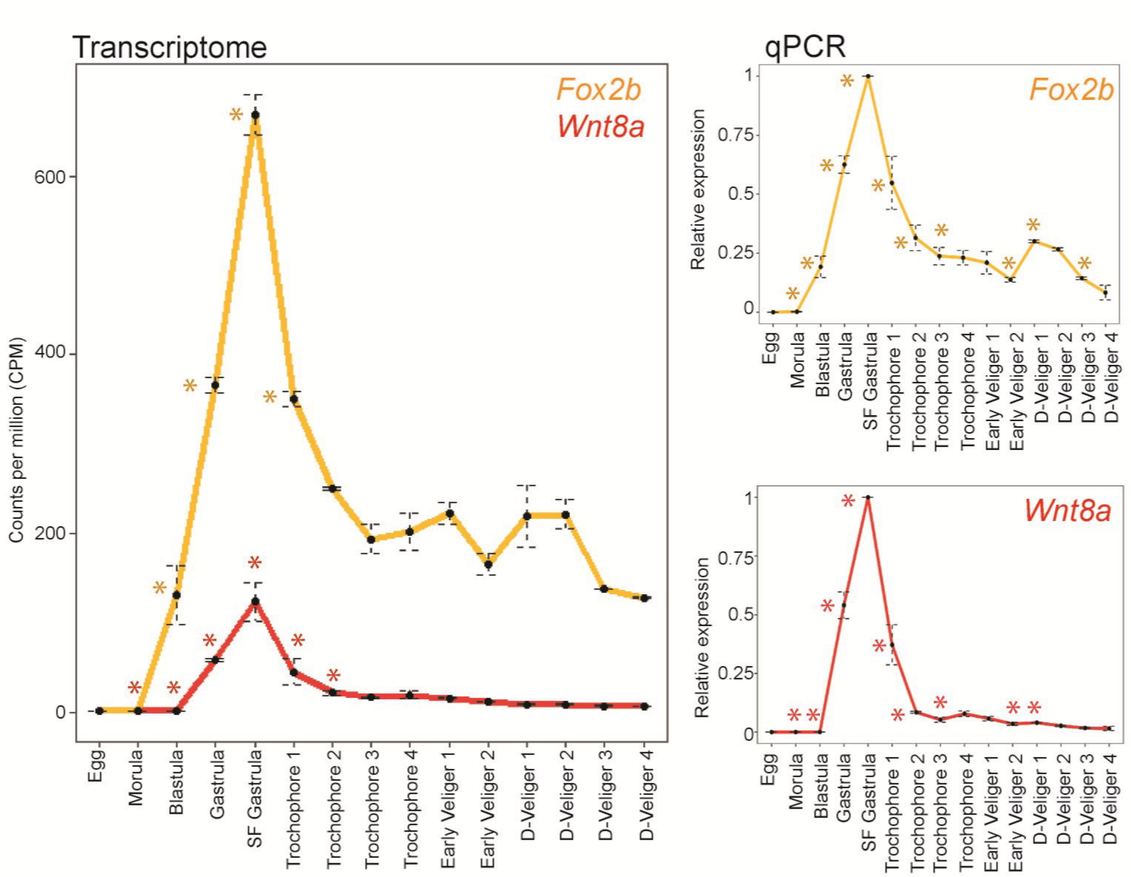
C



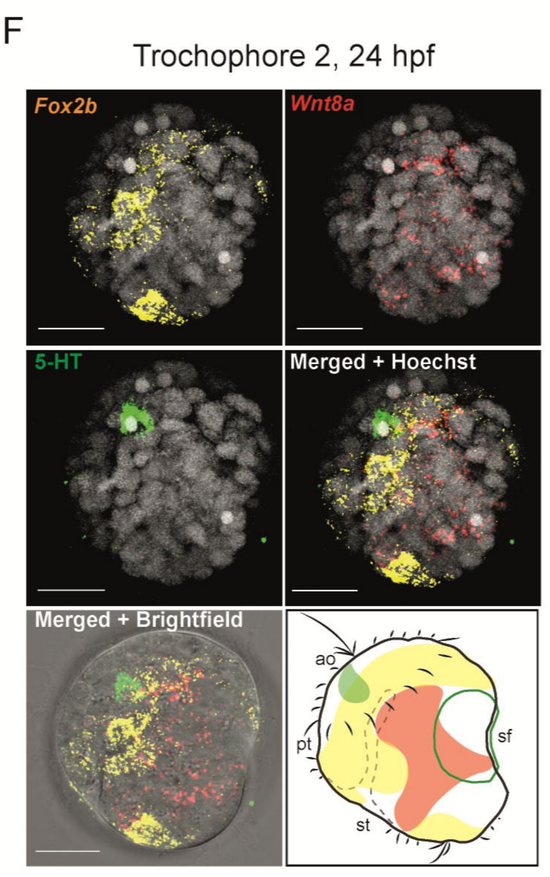
D



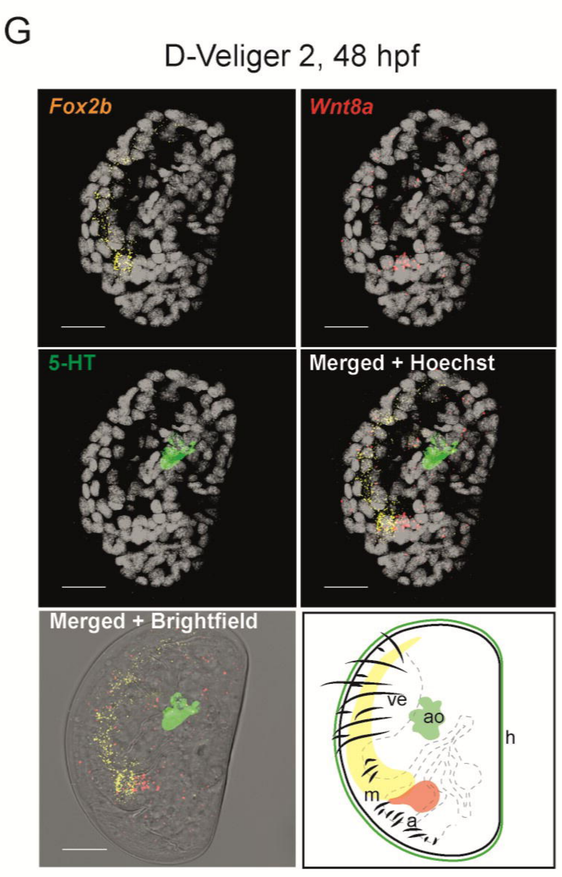
E



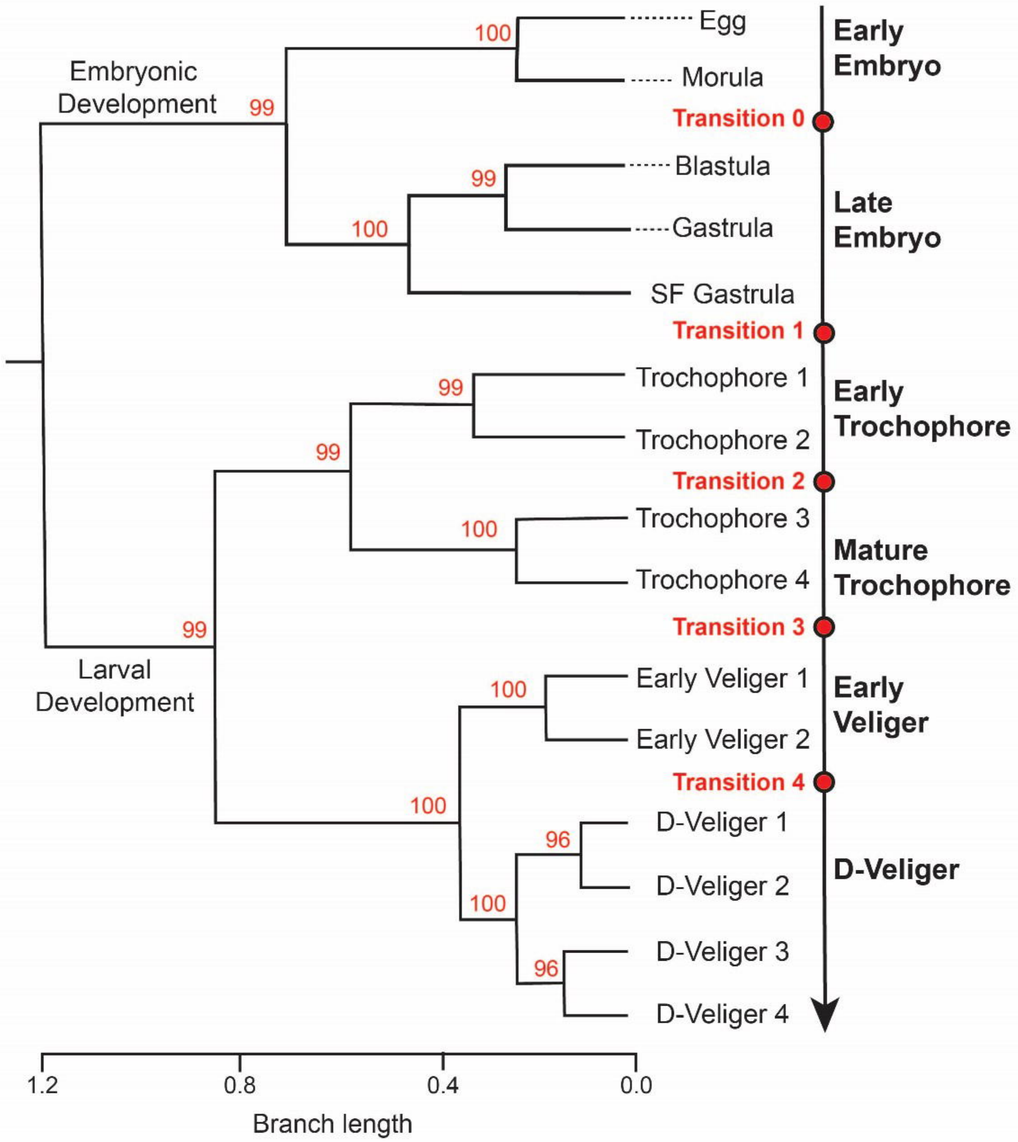
F



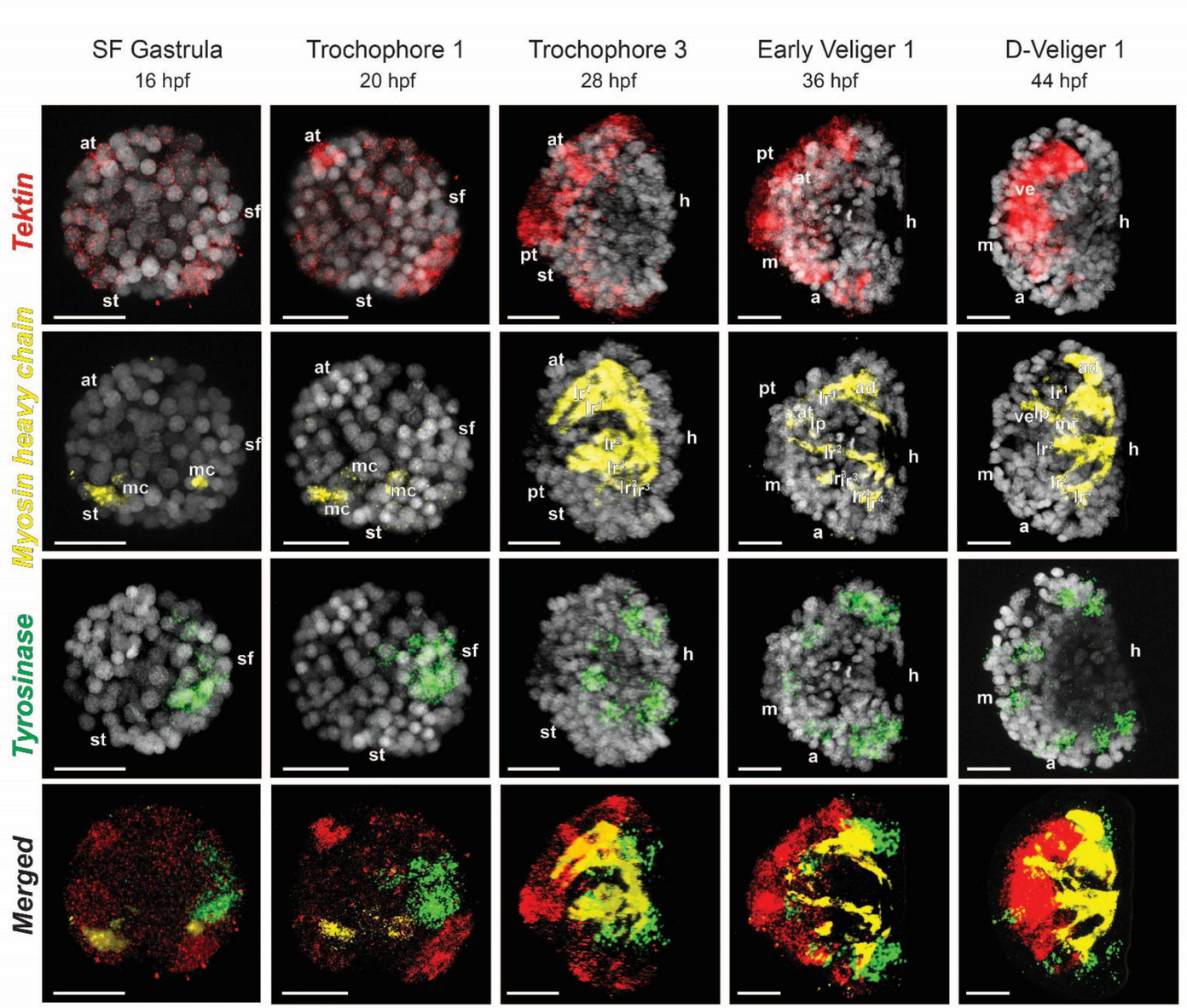
G



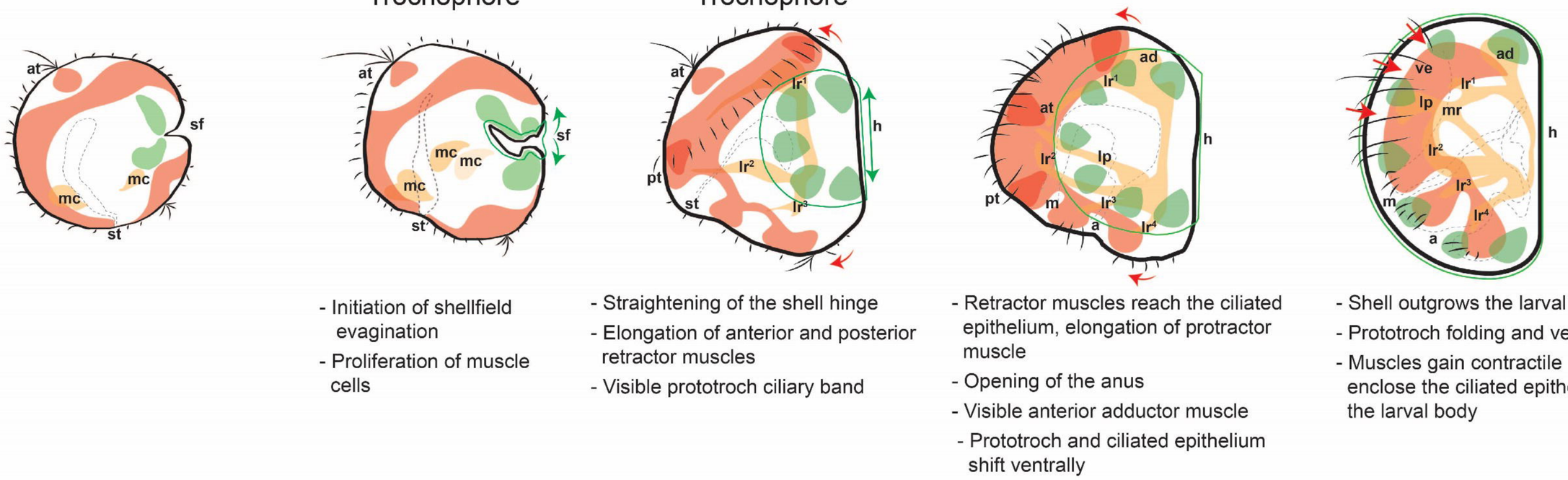
A


Egg
Morula
Blastula
Gastrula
SF Gastrula
Trochophore 1
Trochophore 2
Trochophore 3
Trochophore 4
Early Veliger 1
Early Veliger 2
D-Veliger 1
D-Veliger 2
D-Veliger 3
D-Veliger 4
Embryonic Development
Larval Development
Early Embryo
Late Embryo
Early Trochophore
Mature Trochophore
Early Veliger
D-Veliger
Transition 0
Transition 1
Transition 2
Transition 3
Transition 4
100
99
100
99
100
99
100
100
96
100
96
1.2
0.8
0.4
0.0
Branch length

B


SF Gastrula
16 hpf
Trochophore 1
20 hpf
Trochophore 3
28 hpf
Early Veliger 1
36 hpf
D-Veliger 1
44 hpf
Tektin
Myosin heavy chain
Tyrosinase
Merged
at
sf
st
mc
pt
h
ve
m
a
lr1
lr2
lr3
lr4
ad
lp
mr

C


SF Gastrula
Early Trochophore
Mature Trochophore
Early Veliger
D-Veliger
at
sf
st
mc
pt
h
ad
ve
lp
mr
m
a
lr1
lr2
lr3
lr4
- Initiation of shellfield evagination
- Proliferation of muscle cells
- Straightening of the shell hinge
- Elongation of anterior and posterior retractor muscles
- Visible prototroch ciliary band
- Retractor muscles reach the ciliated epithelium, elongation of protractor muscle
- Opening of the anus
- Visible anterior adductor muscle
- Prototroch and ciliated epithelium shift ventrally
- Shell outgrows the larval body
- Prototroch folding and velum formation
- Muscles gain contractile activity and enclose the ciliated epithelium within the larval body

Mytilus galloprovincialis

z-score
3 2 1 0 -1 -2 -3

Early Embryo | Late Embryo | Early Troch. | Mature Troch. | Early Veliger | D-Veliger

Hierarchical clusters

Maternal

Zygotic

Blastula

Gastrula

Early Trochophore

Mature Trochophore

D-Veliger

Mature D-Veliger

Late D-Veliger

Gene clusters

Egg
Morula
Blastula
Gastrula
SF Gastrula
Trochophore 1
Trochophore 2
Trochophore 3
Trochophore 4
Early Veliger 1
Early Veliger 2
D-Veliger 1
D-Veliger 2
D-Veliger 3
D-Veliger 4

microtubule-based movement
protein phosphorylation
small GTPase mediated signal transduction

DNA replication
mRNA splicing, via spliceosome
RNA splicing

deoxyribonucleotide biosynthetic process
DNA recombination
proteolysis

chromatin remodeling
DNA recombination
telomere maintenance

intracellular protein transport
rRNA processing
regulation of transcription, DNA-templated

regulation of apoptotic process
signal transduction

ion transmembrane transport
G protein-coupled receptor signaling pathway
transmembrane transport

transmembrane transport
translation
G protein-coupled receptor signaling pathway

G protein-coupled receptor signaling pathway
carbohydrate metabolic process
cell-matrix adhesion

A

Crassostrea gigas

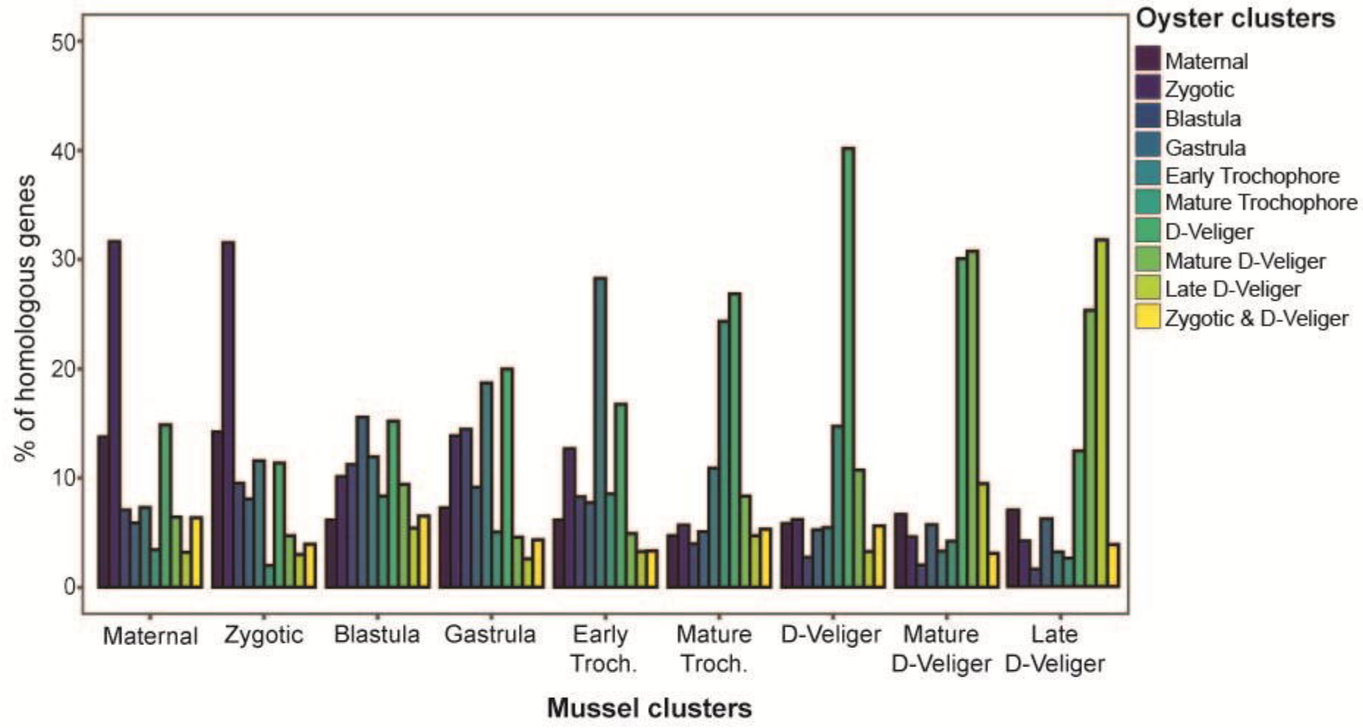
Maternal
Zygotic
Blastula
Gastrula
Early Trochophore
Mature Trochophore
D-Veliger
Mature D-Veliger
Late D-Veliger
Zygotic & D-Veliger

z-score: 4, 2, 0, -2, -4

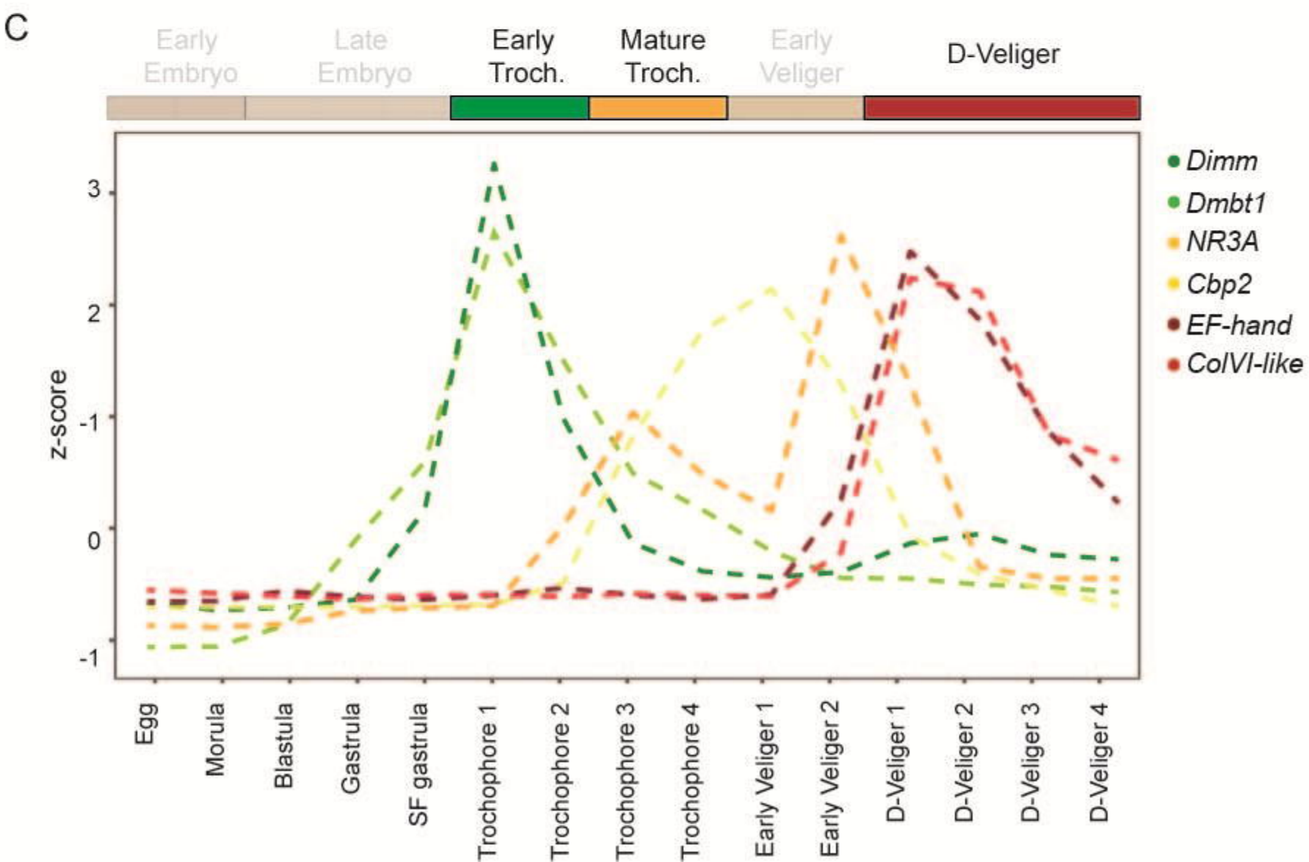
Gene clusters

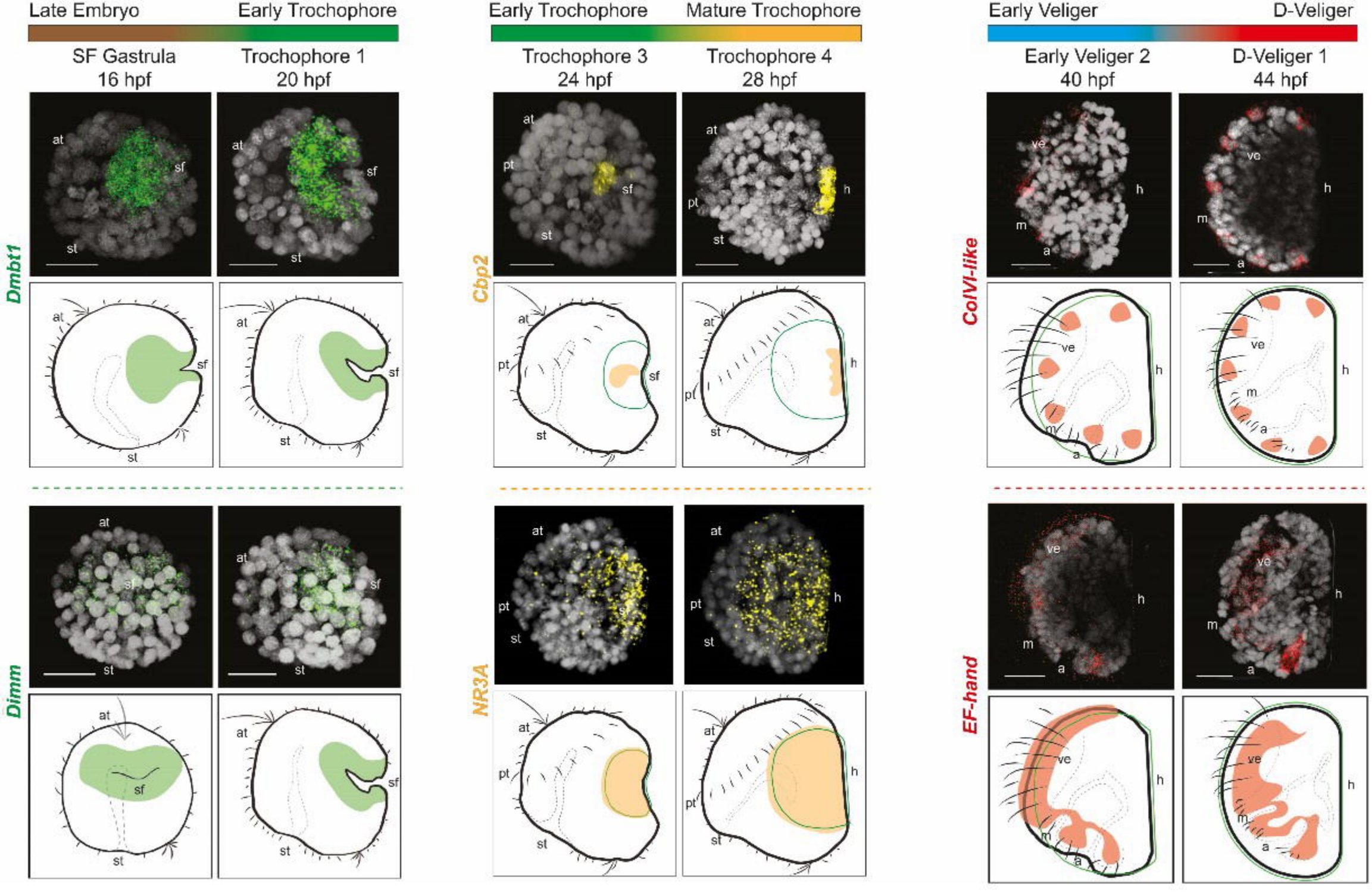
Egg
2 Cells
4 Cells
Early Morula
Morula
Blastula
Early Gastrula
Gastrula
Trochophore 1
Trochophore 2
Trochophore 3
Trochophore 4
Trochophore 5
Early Veliger 1
Early Veliger 2
D-Veliger 1
D-Veliger 2
D-Veliger 3
D-Veliger 4
D-Veliger 5
D-Veliger 6
D-Veliger 7

B



C




Late Embryo
Early Trochophore
Early Trochophore
Mature Trochophore
Early Veliger
D-Veliger
SF Gastrula
16 hpf
Trochophore 1
20 hpf
Trochophore 3
24 hpf
Trochophore 4
28 hpf
Early Veliger 2
40 hpf
D-Veliger 1
44 hpf
Dmbt1
Cbp2
ColVI-like
Dimm
NR3A
EF-hand
at
pt
sf
st
h
ve
m
a