Article

Cardiopharyngeal deconstruction and ancestral tunicate sessility

Alfonso Ferrández-Roldán[1,2], Marc Fabregà-Torrus[1,2,4], Gaspar Sánchez-Serna[1,2,4], Enya Duran-Bello[1,2], Martí Joaquín-Lluís[1,2], Paula Bujosa[1,2], Marcos Plana-Carmona[1,2], Jordi Garcia-Fernàndez[1,3], Ricard Albalat[1,2] & Cristian Cañestro[1,2 ✉]

A central question in chordate evolution is the origin of sessility in adult ascidians, and whether the appendicularian complete free-living style represents a primitive or derived condition among tunicates[1]. According to the 'a new heart for a new head' hypothesis, the evolution of the cardiopharyngeal gene regulatory network appears as a pivotal aspect to understand the evolution of the lifestyles of chordates[2–4]. Here we show that appendicularians experienced massive ancestral losses of cardiopharyngeal genes and subfunctions, leading to the 'deconstruction' of two ancestral modules of the tunicate cardiopharyngeal gene regulatory network that in ascidians are related to early and late multipotency involved in lineage cell-fate determination towards first and second heart fields and siphon muscles. Our work shows that the deconstruction of the cardiopharyngeal gene regulatory network involved the regressive loss of siphon muscle, supporting an evolutionary scenario in which ancestral tunicates had a sessile ascidian-like adult lifestyle. In agreement with this scenario, our findings also suggest that this deconstruction contributed to the acceleration of cardiogenesis and the redesign of the heart into an open-wide laminar structure in appendicularians as evolutionary adaptations during their transition to a complete pelagic free-living style upon the innovation of the food-filtering house[5].

The discovery that the branching of cephalochordates is basal within chordates, and tunicates therefore are the sister group of vertebrates, provided a novel view of adult ancestral chordates as free-living organisms in contrast to the sessile ascidian-like lifestyle alternating motile larva and sessile adults traditionally proposed by Garstang[6–8]. This novel view brought renewed interest in appendicularians, whose complete free-living style could parsimoniously represent the ancestral tunicate condition considering their most accepted position as the sister group of the remaining tunicates[9–12] (although see ref. [13]). In contrast to ascidians, the development of the heart in appendicularians remains poorly understood, and therefore whether differences between the cardiopharyngeal gene regulatory network (GRN) of appendicularians and ascidians reflect adaptations to their different lifestyles remains unknown.

Cardiac developmental atlas and ontogeny

The open-wide laminar heart of appendicularians is considered the simplest chordate heart, consisting just of two layers, the myocardium and the pericardium, the former of which pumps against the stomach[14] (Fig. 1). Here we provide a developmental atlas of the appendicularian heart and show that cardiogenesis in *Oikopleura dioica* is fast, spanning only 3.5 h from the early hatchling stage (5 h post-fertilization (hpf)) at which no morphological evidence of the cardiac primordium could

yet be distinguished, until the late hatchling stage (8.5 hpf) when the heart began beating (video in Supplementary File 1).

Analysis of muscular *Actin 1* (*ActnM1*)[15] expression (in lieu of *Mesp*, which is the preferred precardiac marker in ascidians[16], but is absent in appendicularians (Fig. 1b)) integrated with data from 4D microscopy nuclear tracing[17] (Fig. 1c, Extended Data Fig. 1) identified B8.9 blastomere at the incipient tailbud stage as the first cardiac progenitor cell (CPC). Our analysis revealed that cardiac cells shared lineage with the first three anterior tail muscle cells, following the same ontogenetic origin as in ascidians[17] and therefore providing evidence that ascidian and appendicularian hearts are homologous.

Loss of cardiopharyngeal GRN early module

In vertebrates, the cardiopharyngeal field is the developmental domain that gives rise to the heart and branchiomeric muscles from a common pool of early cardiopharyngeal multipotent progenitors. After a binary-stepwise process of fate choices, early cardiopharyngeal progenitors give rise to the first and second heart fields and to branchiomeric muscles in the head and neck[2]. In ascidians, pharyngeal muscles (that is, siphon and longitudinal muscles) are considered homologous to vertebrate branchiomeric muscles, and their cardiopharyngeal GRN is highly conserved with vertebrates using a homologous binary-stepwise model[4,17–20]. Consistent with this model, the precardiac master regulator
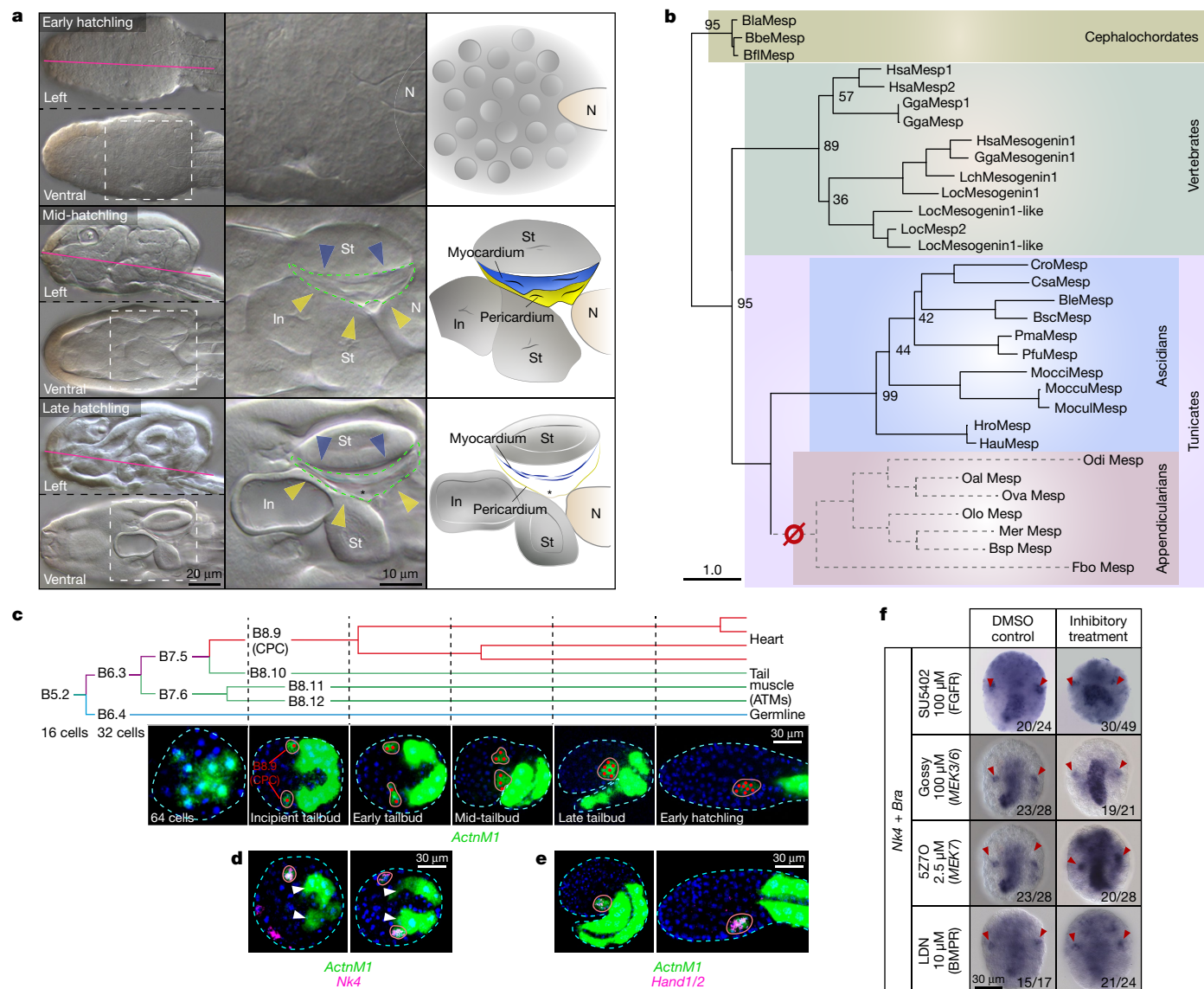
Fig. 1 | O. dioica cardiogenesis. a, Developmental cardiac atlas. In early hatchlings, no cardiac morphology is discernible near the notochord (N). In mid-hatchlings, the heart primordium is made of two layers—the myocardium (blue arrowheads) and the pericardium (yellow arrowheads)—located in the space in-between the stomachs (St), intestine (In) and notochord. In late hatchlings, the internal cardiac space (asterisk) has expanded, and the heart that has asymmetrically relocated towards the left begins beating. Schematic representations are displayed (right). **b**, ML phylogenetic tree shows the loss of *Mesp* in appendicularians (red) and its presence in the rest of chordates (cephalochordates in light green, vertebrates in dark green and ascidians in blue; species abbreviations can be found in Supplementary File 2). The scale bar indicates amino acid substitutions. Bootstrap values are shown. **c**, Integration of the cell lineage fate map reconstructed from 4D nuclear tracing (modified from ref. [16]) with confocal optical sections showing expression of *ActnM1* (green) and nuclear staining (blue) reveals B8.9 as the first CPC of the cardiac lineage (red lines) at the incipient tailbud stage (2 h 20 min post-fertilization), and its split from the anterior tail muscle cell lineage (B10-12; green lines) at the 64-cell stage, and the germline (B6.4; blue line) at the 32-cell stage (Extended Data Fig. 1). The red lines encircle CPCs (red dots). The dashed lines delineate embryos. **d**, Fluorescent whole-mount in situ hybridization (FWMISH) showed transient co-expression of *ActnM1* and *Nk4* in CPCs from incipient tailbud to mid-tailbud stages, but never in anterior tail muscles (white arrowheads). **e**, After *Nk4* downregulation, *ActnM1* and *Hand1/2* co-expressed in the CPCs at late tailbud and early hatchling stages. **f**, WMISH with the cardiac marker *Nk4* and the notochord marker *Brachyury* in DMSO control and treated embryos with inhibitors of the FGF–MEK and BMP signalling pathways (inhibited targets are in parenthesis; the number counts of the phenotypes out of the analysed embryos are also shown) show no effect of these signalling pathways on CPC determination nor the activation of the cardiogenic kernel. BMPR, BMP receptor; FGFR, FGF receptor.

Mesp is expressed in multipotent pre-CPCs in both vertebrates and ascidians[21,22], followed by FGF–MAPK signalling mediated by ETS1/2 phosphorylation, and finally the activation of the cardiac kernel (that is, *Gata4/5/6*, *FoxF*, *Nk4* and *Hand1/2*) and BMP signalling[18,20,23–25].

Our genomic survey of seven appendicularian species and 11 ascidians revealed the absence of *Mesp*, *Ets1/2b*, *Gata4/5/6*, *Mek1/2* and *Hand-r* homologues in all analysed appendicularians, whereas they were present in all ascidians (Fig. 1b, Extended Data Figs. 2–8). Phylogenetic analyses suggested that the absence of these genes was probably due to ancestral gene losses that occurred at the base of the appendicularian lineage after its split from ascidians. Following the surprising absence of a homologue of *Mesp*, considering its precardiac master role in ascidians and vertebrates[21,22,26], we tested for the possibility of 'function shuffling' among *Mesp*-related basic helix–loop–helix (bHLH) genes. However, we did not observe any expression domain of *Math* nor *Neurogenin* (the closest bHLH genes according to blast analysis)
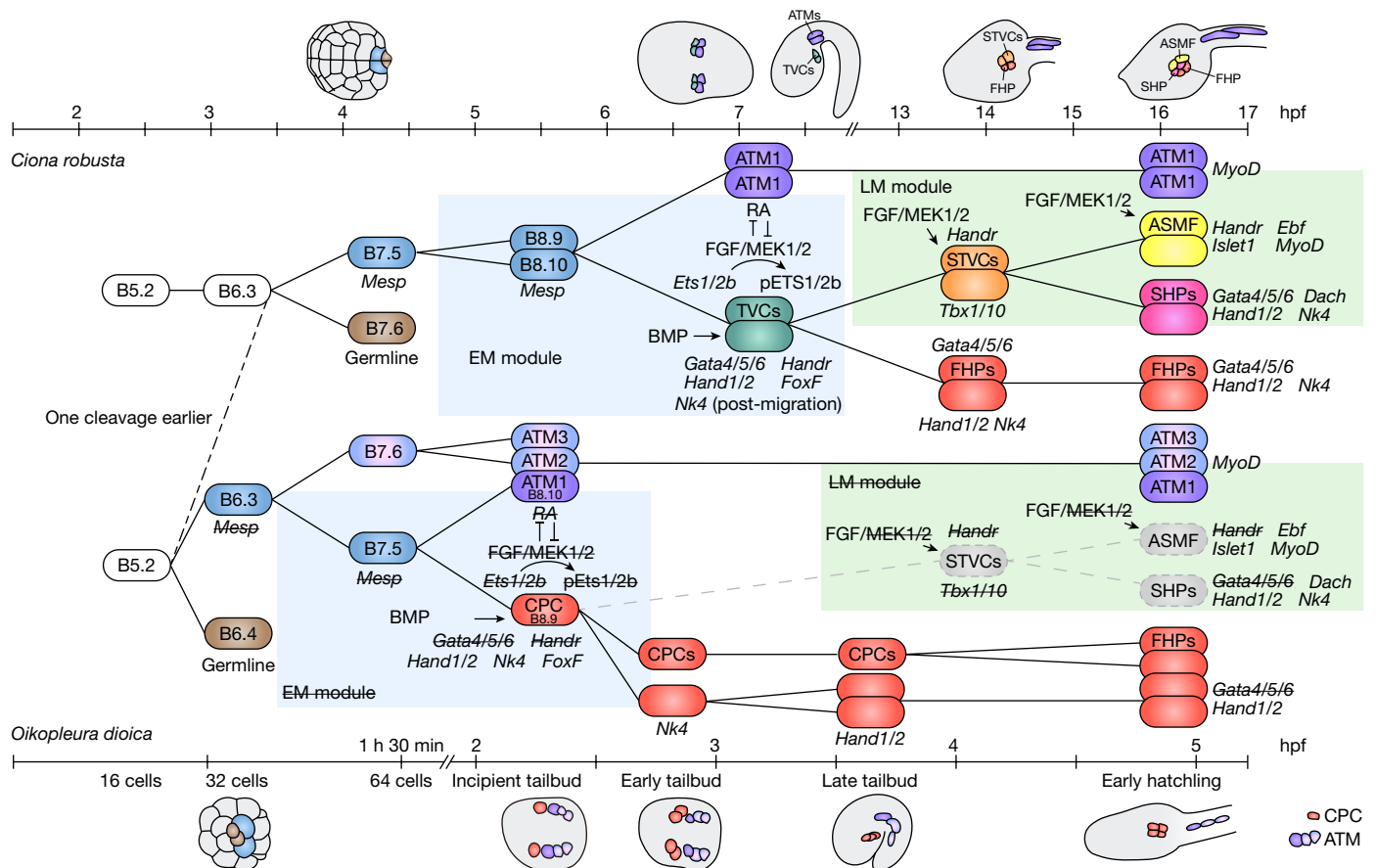
Fig. 2 | Comparison of the cardiopharyngeal cell lineage and GRN in ascidians and appendicularians. The determination of the cardiopharyngeal lineage (blue) and its split from the germline (brown) occurs one cleavage earlier in *O. dioica* than in the ascidian *C. robusta*. In contrast to *O. dioica*, in ascidians, the precardiac *Mesp*-positive cell B7.5 divides before the split of the anterior tail muscle lineage (ATM; purple) and the cardiopharyngeal lineage (blue). In *O. dioica*, the daughter cell of B7.5 rapidly activates the expression of the cardiac kernel (that is, *Nk4* and *Hand1/2*) and becomes a CPC (red), whereas in ascidians, their counterpart TVCs (green) maintain a multipotent state, which will give rise to the first heart precursors (FHPs; red), second heart precursors (SHPs; pink) and atrial siphon muscle field (ASMF; yellow), the last

two through intermediate secondary multipotent cells (STVCs; orange). The lack of *Dach* expression in the heart of *O. dioica* suggests the absence of a homologue of the ascidian second heart field. The numerous losses of cardiopharyngeal genes (strikethrough) and subfunctions (grey) highlight the deconstruction of the 'early' and 'late' ancestral multipotent GRN modules related with the early precardiac multipotency (EM module) of the TVCs in ascidians, and the late multipotency (LM module) of the STVCs and their derivatives such as pharyngeal muscles and second heart field, respectively. The developmental timelines depict the acceleration of cardiogenesis in *O. dioca* compared with ascidians, with the differentiation of the first CPC (red) as soon as 2.5 hpf in *O. dioca*.

that were compatible with precardiac progenitors. We also tested for the possibility of function shuffling among appendicularian-specific duplications of paralogues closely related to the lost genes (that is, two *Ets1/2a* and four *Gata1/2/3*) as well as *MEK7* and *MEK3/6*, but again no tissue-specific expression domains compatible with CPCs were observed. We found homologues of *Nk4*, *Hand1/2* and *FoxF*. While *Nk4*, first, and then *Hand1/2* were sequentially expressed in the CPCs at incipient tailbud and mid-tailbud stages, respectively (Fig. 1d, e), no expression of *FoxF* was observed in the CPCs. These results showed that the first CPC (B8.9) resulting from the split of the anterior tail muscle lineage (B8.10) did not maintain a multipotent state (as their counterpart trunk ventral cells (TVCs) do in ascidians), but rapidly activated the expression of the cardiac kernel (that is *Nk4* and *Hand1/2*). In contrast to ascidians and vertebrates, this activation had become independent of *Mesp*, *Ets1/2*-mediated FGF signalling, *FoxF* and *Gata4/5/6* (ref. [27]). The co-elimination of *Mesp*, *Ets1/2b*, *MEK1/2*, *Gata4/5/6* and the loss of the cardiac subfunction of *FoxF* highlight the deconstruction of what can be considered an ancestral 'early multipotent' module that in ascidians and vertebrates is related to the early maintenance of the multipotent state of the precardiac progenitors[3], and consequently result in an accelerated cardiogenesis in *O. dioica* (Fig. 2).

The deconstruction of this early multipotent module suggested that the conserved cardiogenic roles of FGF–MAPK and BMP signalling in ascidians and vertebrates could have also been altered in appendicularians. To test this hypothesis, we performed inhibitory treatments against the FGF receptor, the surviving paralogues *MEK3/6* and *MEK7*, and BMP receptors, at different concentrations and time windows, and treated embryos were analysed by whole-mount in situ hybridization (Extended Data Fig. 9, Supplementary File 3). Results revealed that the formation of the CPCs and the onset of *Nk4* expression were not altered in treated embryos in which FGF–MAPK or BMP signalling pathways had been inhibited (Fig. 1f). These results suggested that the determination and differentiation of the CPCs and the onset of the cardiogenetic kernel had become independent of these two signalling pathways during the evolution of appendicularians.

Loss of cardiopharyngeal GRN late module

In ascidians, TVCs undertake a series of asymmetric cell divisions and regulatory transient secondary multipotent states (that is, STVCs) that not only give rise to first and second heart precursors but also to the atrial siphon muscle founder cells[4]. We surveyed the appendicularian homologue genes encoding cardiopharyngeal transcription factors
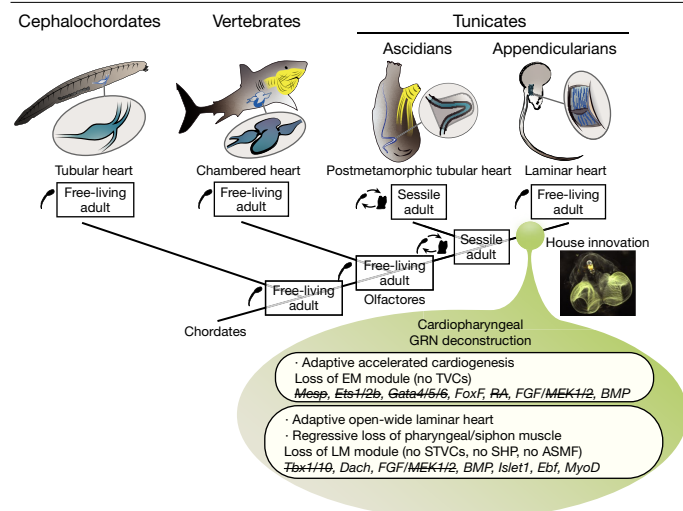
Article

Fig. 3 | Evolutionary scenario of the deconstruction of the cardiopharyngeal GRN and acquisition of an adult free-living style in appendicularians. The evolutionary innovation of an accelerated cardiogenesis, an open-wide laminar heart and the loss of the pharyngeal/siphon muscle in appendicularians can be connected to the deconstruction of their cardiopharyngeal GRN. These three innovations accompanied and plausibly facilitated the transition from an ascidian-like ancestral condition alternating motile larva and sessile adults, to a complete free-living lifestyle of appendicularians upon their innovation of the house. The strikethrough indicates loss of the cardiopharyngeal gene.

such as *Hand-r*, *Tbx1/10*, *Islet1* and *Ebf* (also known as *Coe*) that in ascidians become activated in a FGF–MAPK-dependent manner to determine the trajectory towards the atrial siphon muscles, including the activation of *MyoD* (also known as *Mrf*)[4,19,28] (Extended Data Fig. 10). In addition to the aforementioned absence of *Hand-r*, no expression of *Tbx1/10* was present in any appendicularian, suggesting again an ancestral gene loss in the appendicularian lineage. We found single homologues for *Islet1*, *Ebf* and *MyoD*, the two former genes were expressed in the nervous system, and the later gene in the oikoplastic epithelium, but no expression was found for any of these homologues in the trunk that could suggest the presence of a homologous tissue to the atrial or any other pharyngeal muscle. To test for the presence of a presumptive second heart field in *O. dioica*, we analysed the expression of the homologue of *Dach*, which in ascidians is activated by *Tbx1/10* in the absence of FGF–MAPK signalling, and is necessary to determine the identity of the second heart precursors[19]. Whole-mount in situ hybridization revealed that, while the single homologue of *Dach* was expressed in the nervous system, the endostyle and the trunk epidermis, no expression was detected in the heart, suggesting the absence of a second heart field homologue in *O. dioica* (Extended Data Fig. 10). In addition to *Dach*, our genome survey in *O. dioica* revealed the absence of 12 out of 25 genes that a recent single-cell transcriptomic analysis had revealed to be specific for the first or second heart field precursors in ascidians[19] (Supplementary File 4). In summary, our results highlight that during the deconstruction of the cardiopharyngeal GRN, the loss of *Tbx1/10* and the loss of cardiopharyngeal subfunctions for *Dach*, *Islet1*, *Ebf* and *MyoD* might represent the loss of an ancestral 'late multipotent' module that is required for regulating secondary multipotency and differentiation of the second heart field and atrial muscle in ascidians, structures that appear to have been lost during the evolution of appendicularians[15] (Fig. 2).

Discussion

Homology and one cleavage earlier trend
Understanding the evolution of the cardiopharyngeal GRN in appendicularians and ascidians is key to infer the ancestral lifestyle of tunicates.

Our work provides strong evidence supporting the homology between the hearts of ascidians and appendicularians despite their remarkable differences of morphology, developmental pace and physiology. A key difference is that the specification of the cardiopharyngeal cell lineage and its split from the tail muscle cell lineage occurs one cleavage earlier in appendicularians than in ascidians (Fig. 2). This 'one cleavage earlier' trend can be considered a general aspect of the evolution of the development of appendicularians, consistent with the observation already made by Delsman[29] that gastrulation in *O. dioica* occurred one cleavage earlier than in ascidians, and later corroborated by modern studies[16,30]. This trend has probably contributed to developmental acceleration, morphological simplification and reduction in cell number in this group of tunicates.

Deconstruction and evolutionary impact
One of the most striking findings of our work is the numerous losses of genes and subfunctions that highlight a process of deconstruction of the cardiopharyngeal GRN in appendicularians. The term 'deconstruction', originally coined in philosophy and later applied in literature, architecture, fashion and cookery, or even developmental biology[31], is not synonymous with destruction or homogenously distributed erosion, but instead it refers to the process of dismantling or breaking apart elements that traditionally are combined, and whose analysis facilitates the recognition of structural modules. Our EvoDevo work here, in agreement with the modular model for the control of heart cell identity proposed by Wang et al.[19], unveils the deconstruction of 'evolvable modules' of the cardiopharyngeal GRN, accompanied by developmental system drift and GRN rewiring during the evolution of the heart and pharyngeal muscle in appendicularians (Fig. 2). The loss of the 'early' and 'late' multipotent ancestral modules correlates with the loss of the multipotent states that in ascidians are maintained in the TVC and STVCs, respectively. The losses of these two ancestral modules can be connected to three evolutionary innovations that accompanied, and plausibly facilitated, the evolution from an ancestral sessile ascidian-like adult lifestyle to the pelagic fully free lifestyle of appendicularians: (1) an accelerated cardiogenesis, (2) the formation of an open-wide laminar heart, and (3) the loss of the siphon muscle (Fig. 3).

First, the accelerated cardiogenesis driven by the 'one-cleavage earlier' CPC specification and by the deconstruction of the GRN is probably the result of a primary adaptation to the faster development in appendicularians than in other tunicates. Moreover, accelerated cardiogenesis also enabled the heart to adaptively begin beating as soon as 8.5 hpf in *O. dioica*—in contrast to a few days post-metamorphosis in ascidians—driving haemolymph circulation to be ready when juveniles inflate the first house (10 hpf) and begin pelagic filter feeding.

Second, the low number of cardiac cells (that is, the myocardium is made of only six cells[14]) together with the apparent loss of the second heart homologue in appendicularians is compatible with the transformation of an ascidian-like tubular heart into an open-wide laminar heart that beats against the stomach. Considering that haemolymph circulation in appendicularians is not only powered by the heart but also by tail movements[32], the adaptive innovation of a laminar cardiac structure plausibly offered a more efficient system to pump haemolymph waves propelled by the tail movements through an open-wide structure than through the less accessible space of a tubular ascidian-like heart.

Third, the loss of the 'body wall' and pharyngeal/siphon muscles in the trunk of *O. dioica*[15] can be considered the result of regressive evolution during the transition from a sessile ascidian-like to the pelagic style of appendicularians, in which their functions in sessile ascidians (that is, siphon opening/closing and water squirting as a response to large debris, predators, low tide, or the ejection of faeces or gametes) became useless upon the innovation of the house in appendicularians.

Future perspective
This work exemplifies how the study of gene loss and the application of the concept of deconstruction in evolutionary biology facilitates

the recognition of modules and rewiring of the GRN to better understand the evolution of species. Our study, for instance, supports an evolutionary scenario in which the deconstruction of the cardiopharyngeal GRN was linked to regressive loss of features that characterize the ascidian-like sessile lifestyle such as the siphon muscles, and to the evolution of the accelerated cardiogenesis and the transformation to a laminar heart that could have been adaptively selected during the transition of appendicularians to a pelagic complete free-living active style connected to the innovation of the house (Fig. 3). Our evidence, supporting the view that the last common tunicate ancestor had a biphasic lifestyle alternating motile larva and sessile adults[3,16], is compatible with the commonly accepted assumption that appendicularian branching is basal among tunicates, but it is also compatible with the possibility that appendicularians are phylogenetically related to some groups of ascidians (that is, Aplousobranchia[13]). Thus, our work provides a useful framework for future comparative studies of the cardiopharyngeal GRN among different tunicates, as well as for future efforts to clarify the potential neotenic origin of appendicularians and their phylogenetic relationship with other tunicates.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-04041-w.

1. Satoh, N. in *Chordate Origins and Evolution* (ed. Satoh, N.) 17–30 (Academic Press, 2016).
2. Diogo, R. et al. A new heart for a new head in vertebrate cardiopharyngeal evolution. *Nature* **520**, 466–473 (2015).
3. Razy-Krajka, F. & Stolfi, A. Regulation and evolution of muscle development in tunicates. *Evodevo* **10**, 1–34 (2019).
4. Stolfi, A. et al. Early chordate origins of the vertebrate second heart field. *Science* **565**, 565–569 (2010).
5. Mikhaleva, Y., Skinnes, R., Sumic, S., Thompson, E. M. & Chourrout, D. Development of the house secreting epithelium, a major innovation of tunicate larvaceans, involves multiple homeodomain transcription factors. *Dev. Biol.* **443**, 117–126 (2018).
6. Garstang, W. The morphology of the Tunicata, and its bearings on the phylogeny of the Chrodata. *Quar. J. Micr. Sci.* **72**, 51–186 (1928).
7. Bourlat, S. J. et al. Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature* **444**, 85–88 (2006).
8. Delsuc, F., Brinkmann, H., Chourrout, D. & Philippe, H. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* **439**, 965–968 (2006).
9. Swalla, B. J., Cameron, C. B., Corley, L. S. & Garey, J. R. Urochordates are monophyletic within the deuterostomes. *Syst. Biol.* **49**, 52–64 (2000).
10. Delsuc, F. et al. A phylogenomic framework and timescale for comparative studies of tunicates. *BMC Biol.* **16**, 39 (2018).
11. Kocot, K. M., Tassia, M. G., Halanych, K. M. & Swalla, B. J. Phylogenomics offers resolution of major tunicate relationships. *Mol. Phylogenet. Evol.* **121**, 166–173 (2018).
12. Braun, K., Leubner, F. & Stach, T. Phylogenetic analysis of phenotypic characters of Tunicata supports basal Appendicularia and monophyletic Ascidiacea. *Cladistics* **36**, 259–300 (2020).
13. Stach, T. Ontogeny of the appendicularian *Oikopleura dioica* (Tunicata, Chordata) reveals characters similar to ascidian larvae with sessile adults. *Zoomorphology* **126**, 203–214 (2007).
14. Nishida, H., Ohno, N., Caicci, F. & Manni, L. 3D reconstruction of structures of hatched larva and young juvenile of the larvacean *Oikopleura dioica* using SBF-SEM. *Sci. Rep.* **11**, 1–14 (2021).
15. Almazán, A., Ferrández-Roldán, A., Albalat, R. & Cañestro, C. Developmental atlas of appendicularian *Oikopleura dioica* actins provides new insights into the evolution of the notochord and the cardio-paraxial muscle in chordates. *Dev. Biol.* **448**, 260–270 (2019).
16. Stach, T., Winter, J., Bouquet, J.-M. M., Chourrout, D. & Schnabel, R. Embryology of a planktonic tunicate reveals traces of sessility. *Proc. Natl Acad. Sci. USA* **105**, 7229–7234 (2008).
17. Davidson, B. *Ciona intestinalis* as a model for cardiac development. *Semin. Cell Dev. Biol.* **18**, 16–26 (2007).
18. Christiaen, L., Stolfi, A. & Levine, M. BMP signaling coordinates gene expression and cell migration during precardiac mesoderm development. *Dev. Biol.* **340**, 179–187 (2010).
19. Wang, W. et al. A single-cell transcriptional roadmap for cardiopharyngeal fate diversification. *Nat. Cell Biol.* **21**, 674–686 (2019).
20. Racioppi, C., Wiechecki, K. A. & Christiaen, L. Combinatorial chromatin dynamics foster accurate cardiopharyngeal fate choices. *eLife* **8**, 1–33 (2019).
21. Lescroart, F. et al. Early lineage restriction in temporally distinct populations of Mesp1 progenitors during mammalian heart development. *Nat. Cell Biol.* **16**, 829–840 (2014).
22. Satou, Y., Imai, K. S. & Satoh, N. The ascidian Mesp gene specifies heart precursor cells. *Development* **131**, 2533–2541 (2004).
23. Davidson, B., Shi, W., Beh, J., Christiaen, L. & Levine, M. FGF signaling delineates the cardiac progenitor field in the simple chordate, *Ciona intestinalis*. *Genes Dev.* **20**, 2728–2738 (2006).
24. Bernadskaya, Y. Y., Brahmbhatt, S., Gline, S. E., Wang, W. & Christiaen, L. Discoidin-domain receptor coordinates cell-matrix adhesion and collective polarity in migratory cardiopharyngeal progenitors. *Nat. Commun.* **10**, 57 (2019).
25. Shi, Y., Katsev, S., Cai, C. & Evans, S. BMP signaling is required for heart formation in vertebrates. *Dev. Biol.* **224**, 226–237 (2000).
26. Davidson, B., Shi, W. & Levine, M. Uncoupling heart cell specification and migration in the simple chordate *Ciona intestinalis*. *Development* **132**, 4811–4818 (2005).
27. Schachterle, W., Rojas, A., Xu, S.-M. & Black, B. L. ETS-dependent regulation of a distal Gata4 cardiac enhancer. *Dev. Biol.* **361**, 439–449 (2012).
28. Wang, W., Razy-Krajka, F., Siu, E., Ketcham, A. & Christiaen, L. NK4 antagonizes Tbx1/10 to promote cardiac versus pharyngeal muscle fate in the ascidian second heart field. *PLoS Biol.* **11**, e1001725 (2013).
29. Delsman, H. C. Contributions on the ontogeny of *Oikopleura dioica*. *Verch. Rijksinst. Onderz. Zee* **3**, 1–24 (1910).
30. Fujii, S., Nishio, T. & Nishida, H. Cleavage pattern, gastrulation, and neurulation in the appendicularian, *Oikopleura dioica*. *Dev. Genes Evol.* **218**, 69–79 (2008).
31. Hogan, B. Deconstructing the genesis of animal form. *Development* **131**, 2515–2520 (2004).
32. Fenaux, R. in *The Biology of Pelagic Tunicates* (ed. Bone, Q.) 25–34 (Oxford Univ. Press, 1998).

Article

Methods

Biological material

O. dioica specimens were obtained from the Mediterranean coast of Barcelona (Catalonia, Spain). Culturing of *O. dioica* and embryo collections have been performed as previously described[33].

Genome database searches and phylogenetic analysis

Protein sequences from the tunicate *Ciona robusta* and the vertebrate *Homo sapiens* were used as queries in BLASTp and tBLASTn searches in genome databases of selected species: https://blast.ncbi.nlm.nih.gov/Blast.cgi for *Branchiostoma floridae*, *Branchiostoma belcheri*, *Branchiostoma lanceolatum*, *Gallus gallus*, *Lepisosteus oculatus* and *Latimeria chalumnae*; http://www.aniseed.cnrs.fr/ for the ascidian species[34] *Ciona savignyi*, *Phallusia fumigata*, *Phallusia mammillata*, *Halocynthia roretzi*, *Halocynthia aurantium*, *Botryllus schlosseri*, *Botryllus leachii*, *Molgula occulta*, *Molgula oculata* and *Molgula occidentalis*; and http://oikoarrays.biology.uiowa.edu/Oiko/ for *O. dioica*, and for six other appendicularian species with public genomes available in GeneBank (*Oikopleura albicans* SCLG01000000, *Oikopleura vanhoeffeni* SCLH01000000, *Oikopleura longicauda* SCLD01000000, *Mesochordaeus erythrocephalus* SCLF01000000, *Bathochordaeus stygius SCLEO1OOOOOO* and *Fritillaria borealis* SDII01000000)[35]. The orthology between potential cardiac genes was initially assessed by blast reciprocal best hit (BRBH) and subsequently by phylogenetic analysis based on ML inferences calculated with PhyML v3.0 and an automatic substitution model[36] using protein alignment generated by MUSCLE and reviewed manually with the package AliView v1.17.1 (ref. [37]). Species abbreviations and gene accession numbers are provided in Supplementary File 2.

Cardiac lineage tracing using 4D microscopy

Cardiac lineage tracing was performed using Supplementary Video 2 from Stach et al.[16]. We followed cell divisions starting from B5.2 blastomere at the 16-cell stage until the late tailbud stage when the CPC divides from the first anterior tail muscle cell. Blastomere nomenclature follows that of Conklin for ascidians (vegetal blastomeres in capital letters, animal blastomeres in small letters, and blastomeres from the right underlined)[38].

Cloning and expression analysis

O. dioica genes were PCR amplified from cDNA obtained as previously described[39]. Then, they were cloned using the Topo TA Cloning Kit (K4530-20, Invitrogen) to synthesize antisense digoxigenin (DIG) and fluorescein (FITC) riboprobes for whole-mount in situ hybridization (WMISH)[39–41] and double fluorescent whole-mount in situ hybridization (FWMISH). Gene probes, forward primers, reverse primers, template, length and RNA-pol/digestion enzymes are as follow: *OdActnM1* cross-hybridizing: 5′GTCCCCGCCATG TACGTCTG3′, 5′GCATCGGAATCGCTCGTTACCA3′, gDNA exon 2 partial, 389 bp, T3/NotI; *OdActnM1* specific: 5′GATCGTCCACCGAAA GTGC3′, 5′GTCAGCAACTGTTTGAATATATTG3′, cDNA 3′ UTR, 351 bp, T7/PstI; *OdBrachyury*: 5′GGTTCGCACTGGATGAAACAGCC3′, 5′TATCCGTTCTGACACCAGTCGTTC3′, gDNA exon 3, 630 bp, T3/NotI; *OdDach*: 5′GAGATGGATCCTGCGCAGC3′, 5′GTAAGTTTAA GAATATCCGAAAATCC3′, cDNA full length, 770 bp, T7/SpeI; *OdEbf* (*COE*): 5′GAGATCATGTGTTCCCGATGTTG3′, 5′GTTGAGTGAAA GAAAACCTTGCT3′, cDNA exon 5 to exon 8, 578 bp, T3/NotI; *OdERK*: 5′GAAGGAGCCTACGGCATAG3′, 5′GCTAGAATACATCCGACAGAC3′, cDNA exon 1 to exon 5, 563 bp, T3/NotI; *OdEts1/2a1*: 5′GACGGCATT GATGGATTACAGCTAT3′, 5′GTTACTCTTCAGTCTCTGGCTC3′, cDNA exon 3/4 to exon 7, 480 bp, T3/NotI; *OdEts1/2a2*: 5′GACTCTCTTTGC CATCAATCC3′, 5′GCCTTTTTTCGTCCAGCTAATG3′, cDNA exon 2 to exon 5, 728 bp, PCR4-TOPO,T3/NotI; *OdFoxF*: 5′GGCTGGAAGAATTC CGTCCG3′, 5′GAGCTGATTCGCATGGGCAGG3′, cDNA exon 3 to exon 7, 639 bp, T7/SpeI; *OdGata1/2/3b*: 5′GCCTCTTCTGATTCGCCATTC3′, 5′GGAATGACTGTTTGGTGTTGG3′, cDNA exon 1 to exon 3, 791 bp, T3/XbaI; *OdGata1/2/3d*: 5′GGGCAGAATATGAAAATGTATTTT3′, 5′GCTGACC GTCCGCTAGTC3′, cDNA full length, 1,086 bp, T7/BamHI; *OdHand1/2*: 5′GATGGAGTTGAATTTGTATTCCGATC3′, 5′GATTCTTTTCTAATCAG ATGGGCA3′, cDNA full length, 462 bp, T7/SpeI; *OdIslet*: 5′GGTCTCCGG TGATGAATTCCT3′, 5′GCTTTGTCGTAGGTTTAGGCCA3′, cDNA exon 4 to exon 8, 766 bp, T7/PstI; *OdMath6*: 5′GCCGAATTCCACACAACGAAG3′, 5′CAACGTTAGCAGGTATAGAATAG3′, cDNA exon 1/2 to 3′ UTR, 602 bp, T7/SpeI; *OdMEK3/6*: 5′..AGTGACGAGCAACGACCCTC3′, 5′..ATGC ATCTTTGCAGCTCTTCGT3′, cDNA exon 1 to exon 7, 1,034bp, T3/NotI; *OdMEK7*: 5′..CCCCGTTCCTGTTCTAGGGTC3′, 5′..GGCCCGTTACAAAG ACGCTG3′, cDNA exon 1 to exon 5, 842 bp, T7/SpeI; *OdMyoD*: 5′GTTAC AAAATGACTATGACGGAAAC3′, 5′GGCTTCCAAGTTTCTTGACCAG3′, cDNA full length, 813 bp, T7/PstI; *OdNeurogenin*: 5′..GAGCACTTCTCC AAAAACAGAG3′, 5′.. GTTTTTACATTGTCGGAAATTCTG3′, cDNA full length, 833 bp, T3/NotI; *OdNk4*: 5′GACCGAAAAATTACAACTATGAGC3′, 5′GCTGTAGCGCCGAGCTCAC3′, cDNA exon 1 to exon 3, 649 bp, T3/NotI.

For FWMISH, fixed embryos were rehydrated in PBT (PBS with 0.2% Tween-20), treated with 50 mM DTT in PBT (for 10 min at room temperature), washed in 0.1 M triethanolamine in PBT (2 × 5 min at room temperature), treated with two successive dilutions of acetic anhydride (0.25% and 0.5%) in 0.1 M triethanolamine (10 min at room temperature), and washed in PBT (2 × 5 min at room temperature). Prehybridization was carried out in a mixture of 50% formamide, 5× SSC, 0.1 mg/ml heparin, 0.15% Tween-20, 5 mM EDTA, 0.5 mg/ml yeast RNA and 1× Denhardt's reagent for 2 h at 63 °C. Then, hybridization was carried out in the same solution but adding the two probes at 0.5–1 ng/μl each, overnight at 63 °C. Next day, embryos were washed in successive dilution of SSC (2 × 10 min in 2× SSC/0.2% Tween-20 at 65 °C; 2 × 10 min in 0.2× SSC/0.2% Tween-20 at 65 °C; 1 × 5 min in 0.1× SSC/0.2% Tween-20 at room temperature) and 2 × 5 min in MABT (0.1 M maleic acid, 0.15 M NaCl and 0.1% Tween-20, pH 7.5). Blocking was performed by washing the embryos in a mixture of MABT, 2.5 mg/ml BSA and 5% sheep serum. Finally, anti-FITC-POD antibody (1:1,000 in blocking solution) (11426346910, Roche) was added to the samples for overnight incubation at 4 °C. The day after, samples were washed in MABT (eight times for 15 min at room temperature) and then in TNT (0.1 M Tris-HCl pH 7, 0.15 M NaCl and 0.3% Triton X-100) for 10 min. For the staining, embryos were incubated in TSA-tetramethylrhodamine (NEL742001KT, Perkin Elmer) for 10 min. Then, they were washed in TNT for 10 min, in PBT for 10 min, in 2% H_2O_2/PBT for 45 min, in PBT (two times for 5 min) and in MABT (two times for 5 min). Then, a second blocking step was performed followed by the addition of an anti-DIG-POD antibody (1:1,000 in blocking solution; 11207733910, Roche) that was incubated overnight at 4 °C. The morning after, as the previous day, the samples were washed and the coloration reaction was added, that in this case included TSA-FITC system-green (NEL741E001KT, Perkin Elmer) for 1 h 30 min. After coloration, embryos were washed in TNT (2 × 5 min at room temperature) and PBT (2 × 5 min at room temperature). Mounting was made in 80% glycerol/PBS with Hoechst-33342 1 μM (62249, Invitrogen) as previously described[15]. A confocal microscopy LSM880 (Zeiss) was used for imaging of samples and FIJI[42] was used to compose the confocal series and adjust the brightness and contrast.

Pharmacological treatments

For FGF receptor inhibition, animals were treated with 50 μM and 100 μM of SU5402 (SML0443, Merck) and AZD4547 (9403, BioVision) from the two-cell stage (30 min post-fertilization (mpf)) and from 32-cell stage (70 mpf), respectively, to hatchling stage (4 hpf) in darkness. For *MEK3/6* inhibition, animals were treated with gossypetin (1176, Extrasynthese) 100 μM from the 2-cell to the 32-cell stage in darkness. For *MEK7* inhibition, animals were treated with 1 μM and 25 μM of 5Z-7-oxozeaenol (O9890, Merck) from the 2-cell stage to the 32-cell

stage, respectively. For BMP inhibition, animals were treated with 10 μM of LDN (SML1119, Merck) and dorsomorphin (P5499, Merck) from the 2-cell stage and from the 32-cell stage until hatchling stage. To perform these treatments, eggs were pooled in 4 ml of SSW and fertilized with 200 μl of sperm dilution (the sperm of three males in 5 ml of SSW). At the desired time, embryos were transferred to a 3-mm Petri dish plate with 4 ml of treatment solution at 19 °C. Control embryos were incubated in DMSO 0.2% or 0.3% (v/v) depending on the concentration of the treatment. The effects of the treatments were scored by in situ hybridization. For tailbud embryos, we used cross-hybridizing *ActnM1*, *Nk4* and *Brachyury* probes[15,43], whereas for hatchling embryos, we used the specific *ActnM1* probe[15].

Statistics and reproducibility

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment. Descriptions of morphological features in live animals or expression domains obtained by WMISH or FWMISH were performed in at least five specimens (usually from 10 to 20) in each analysed developmental stage. WMISH and FWMISH were performed at least twice for each probed gene. Inhibitory treatments were performed at least twice for each condition, and numbers of phenotype counts out of the total number of analysed embryos are indicated in Fig. 1f, Extended Data Fig. 9 and Supplementary File 3. The sex condition of embryos does not influence experimental design.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Accession numbers and URLs of databases from publicly available sources are provided in the Methods, Supplementary Information and Supplementary File 2.

33. Martí-Solans, J. et al. *Oikopleura dioica* culturing made easy: a low-cost facility for an emerging animal model in EvoDevo. *Genesis* **53**, 183–193 (2015).
34. Brozovic, M. et al. ANISEED 2017: extending the integrated ascidian database to the exploration and evolutionary comparison of genome-scale datasets. *Nucleic Acids Res.* **46**, D718–D725 (2018).
35. Naville, M. et al. Massive changes of genome size driven by expansions of non-autonomous transposable elements. *Curr. Biol.* **29**, 1161–1168 (2019).
36. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
37. Larsson, A. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **30**, 3276–3278 (2014).
38. Conklin, E. G. The organization and cell lineage of the ascidian egg. *J. Acad. Nat. Sci. Phila.* **13**, 1–119 (1905).
39. Martí-Solans, J. et al. Coelimination and survival in gene network evolution: dismantling the RA-signaling in a chordate. *Mol. Biol. Evol.* **33**, 2401–2416 (2016).
40. Bassham, S. & Postlethwait, J. Brachyury (T) expression in embryos of a larvacean urochordate, *Oikopleura dioica,* and the ancestral role of T. *Dev. Biol.* **220**, 322–332 (2000).
41. Cañestro, C. & Postlethwait, J. H. Development of a chordate anterior–posterior axis without classical retinoic acid signaling. *Dev. Biol.* **305**, 522–538 (2007).
42. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
43. Torres-Águila, N. P. et al. Diatom bloom-derived biotoxins cause aberrant development and gene expression in the appendicularian chordate *Oikopleura dioica*. *Commun. Biol.* **1**, 121 (2018).

Extended Data Fig. 1 | 4D-reconstruction of a virtual cardiac cell tracing, based on nuclear position from the 30-cell stage to tailbud stages of *O. dioica* embryos (modified from Stach 2008)[16]. B8.9 appears as the first CPC. Blastomere nomenclature follows that of Conklin for ascidians (vegetal blastomeres in capital letters, animal blastomeres in small letters, and blastomeres from the right underlined)[38], and their fate are indicated in different colors: muscle+heart (purple), posterior tail muscle cells (yellow), anterior tail muscle cells (ATM, green), heart (red), germ-line (blue). Circles and hexagons represent blastomeres derived from right and left sides of the embryo, respectively. Dashes encircle sister cells resulting from a cell division.
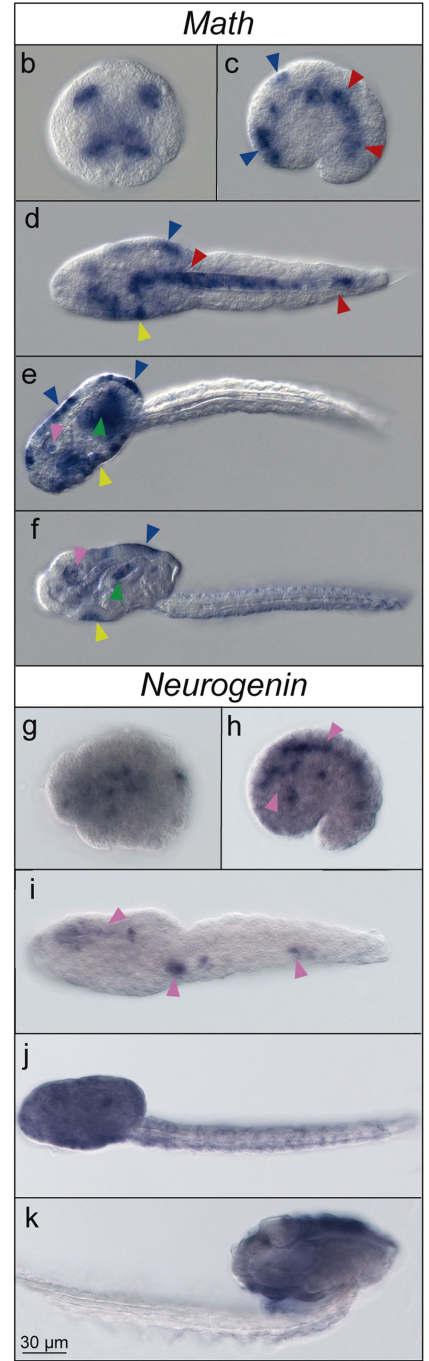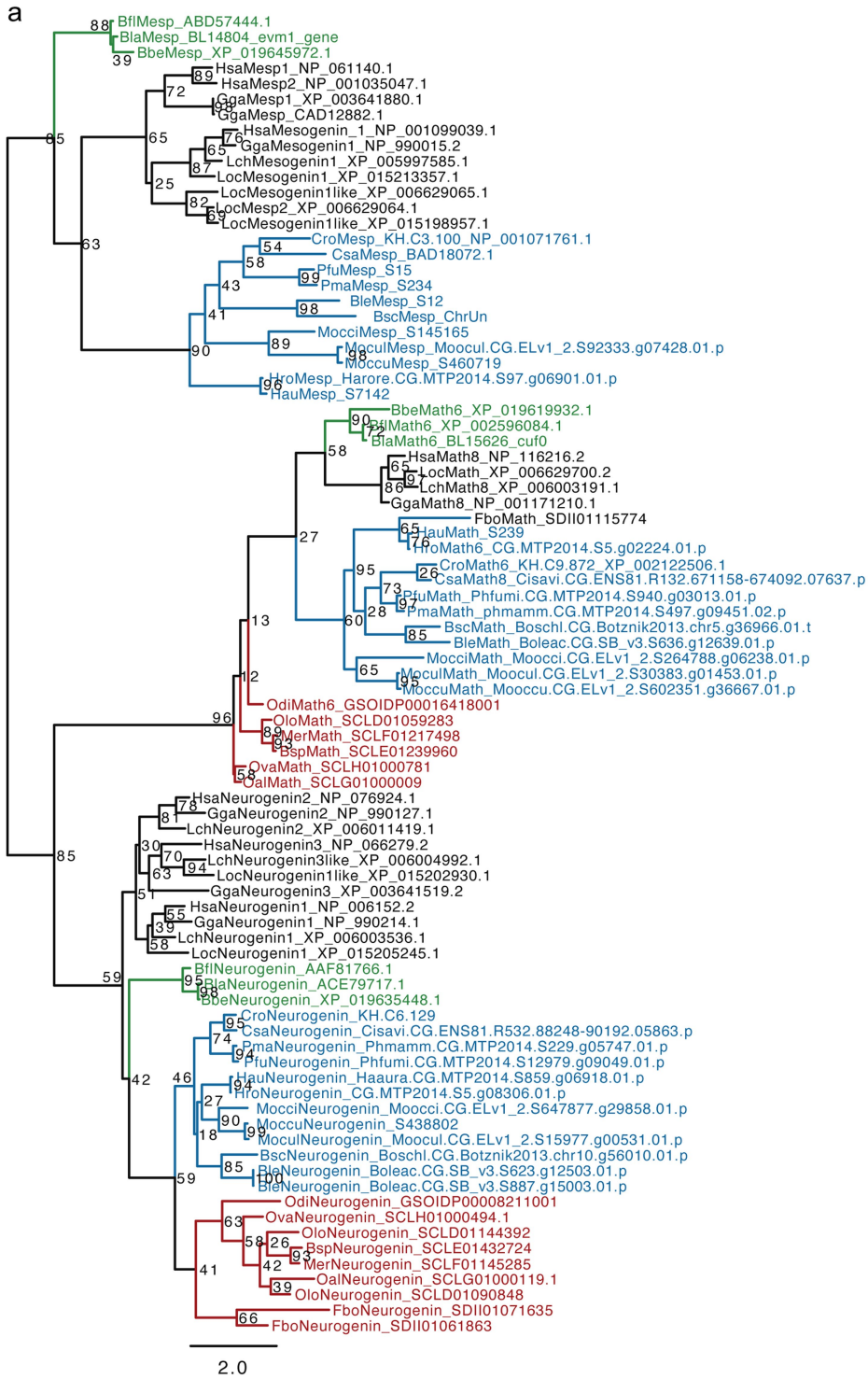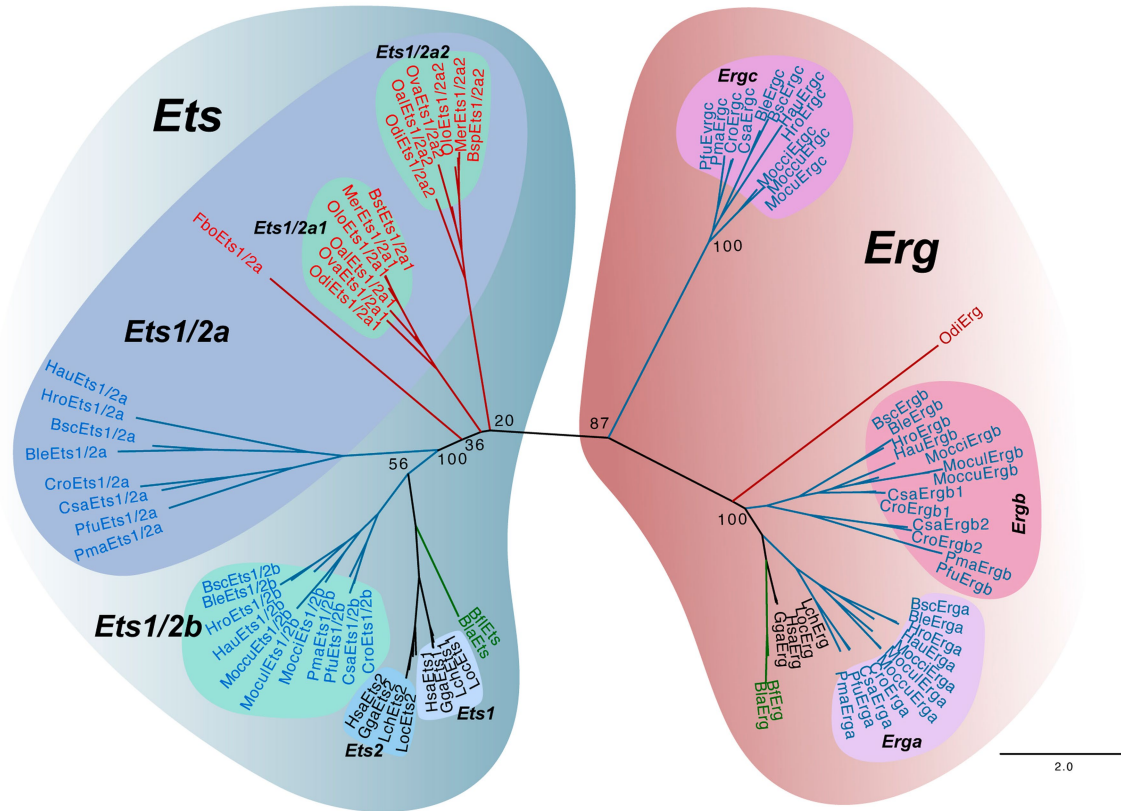
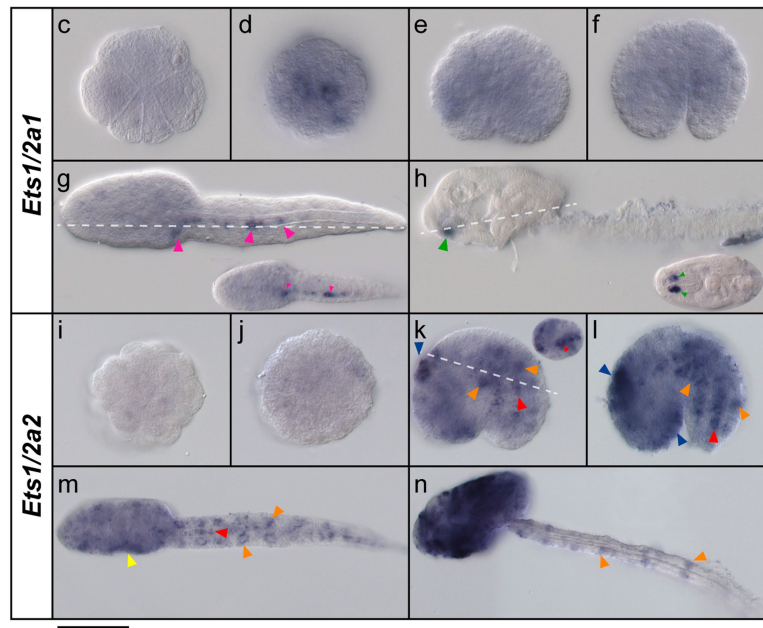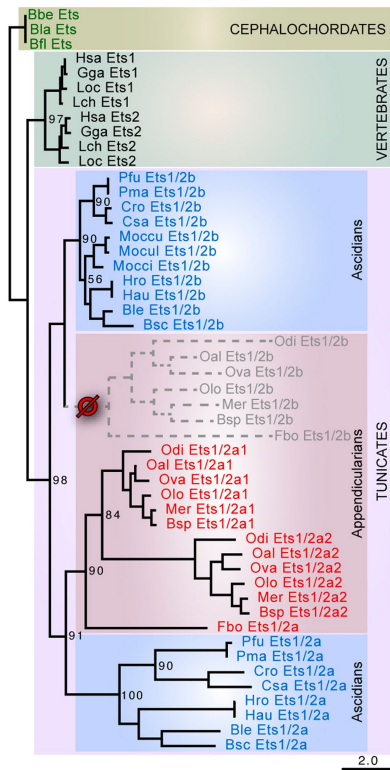**Extended Data Fig. 2** | See next page for caption.

Article

Extended Data Fig. 2 | Mesp ML phylogenetic tree and *Math* and *Neurogenin* expression. a, Unrooted phylogenetic tree, represented in a rectangular layout for the sake of clarity, showing the presence of bHLH homologs of *Neurogenin* and *Math* in appendicularians, but the absence of *Mesp*. The presence of *Mesp* in cephalochordates, vertebrates and all analyzed ascidians suggests an ancestral loss of *Mesp* at the base of the appendicularian lineage after its split from the lineage leading to ascidians. Bootstrap values are shown in the nodes. Scale bar indicates amino acid substitutions. Vertebrates (black): *Gallus gallus* (Gga), *Homo sapiens* (Hsa), *Latimeria chalumnae* (Lch), *Lepisosteus oculatus* (Loc); Ascidian tunicates (blue): *Botrylloides leachii* (Ble), *Botrylloides schlosseri* (Bsc), *Ciona robusta* (Cro), *Ciona savignyi* (Csa), *Halocynthia aurantium* (Hau), *Halocynthia roretzi* (Hro), *Molgula occidentalis* (Mocci), *Molgula occulta* (Moccu), *Molgula oculata* (Mocul), *Phallusia fumigata* (Pfu), *Phallusia mammillata* (Pma); Appendicularian tunicates (red): *Bathochordaeus sp.* (Bsp), *Fritillaria borealis* (Fbo), *Mesochordaeus erythrocephalus* (Mer), *Oikopleura albicans* (Oal), *Oikopleura dioica* (Odi), *Oikopleura longicauda* (Olo), *Oikopleura vanhoeffeni* (Ova); Cephalochordates (green): *Branchiostoma belcheri* (Bbe), *Branchiostoma floridae* (Bfl), *Branchiostoma lanceolatum* (Bla). **b**–**f**, Developmental expression pattern of *O. dioica Math* homolog. Whole mount in situ hybridization in different stages of *O. dioica* development showing expression in the notochord in tailbud and early-hatchling embryos (red arrowheads) (**c**, **d**), in epidermis (blue arrowheads) (**c**–**f**), in the rectum domain in hatchling stages (yellow arrowheads) (**d**–**f**), in later stages of neural system development (pink arrowheads) (**e**, **f**), and in later stages of digestive system development (green arrowheads) (**e**, **f**). **g**–**k**, Developmental expression pattern of *O. dioica Neurogenin* homolog. Whole mount in situ hybridization in different stages of *O. dioica* development shows that *Neurogenin* expression was restricted to nervous system in tailbud and early-hatchling stages (pink arrowheads) (**h**, **i**) but no expression was detected in any region compatible with cardiac function. Images from tailbud in advance correspond to left lateral views orientated anterior towards the left and dorsal towards the top.
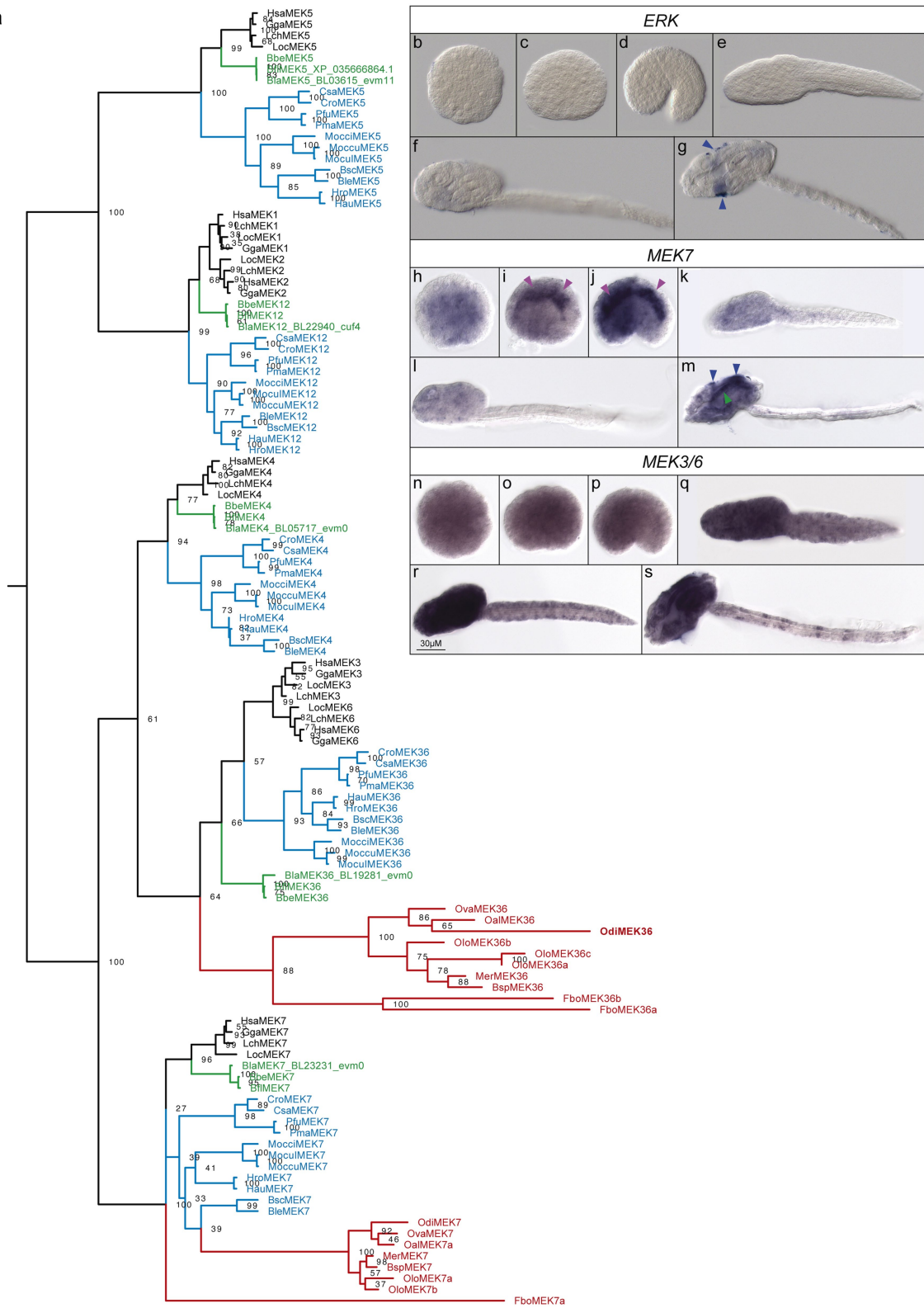
Extended Data Fig. 3 | See next page for caption.

Article

Extended Data Fig. 3 | Ets ML phylogenetic tree and expression. a, Unrooted phylogenetic tree of the *Ets* and *Erg* protein families showed a high bootstrap value separating both protein families what corroborated the existence of two *Ets1/2* genes in appendicularians. Scale bar indicates amino acid substitutions. Vertebrates (black): *Gallus gallus* (Gga), *Homo sapiens* (Hsa), *Latimeria chalumnae* (Lch), *Lepisosteus oculatus* (Loc); Ascidian tunicates (blue): *Botrylloides leachii* (Ble), *Botrylloides schlosseri* (Bsc), *Ciona robusta* (Cro), *Ciona savignyi* (Csa), *Halocynthia aurantium* (Hau), *Halocynthia roretzi* (Hro), *Molgula occidentalis* (Mocci), *Molgula occulta* (Moccu), *Molgula oculata* (Mocul), *Phallusia fumigata* (Pfu), *Phallusia mammillata* (Pma); Appendicularian tunicates (red): *Bathochordaeus sp.* (Bsp), *Fritillaria borealis* (Fbo), *Mesochordaeus erythrocephalus* (Mer), *Oikopleura albicans* (Oal), *Oikopleura dioica* (Odi), *Oikopleura longicauda* (Olo), *Oikopleura vanhoeffeni* (Ova); Cephalochordates (green): *Branchiostoma belcheri* (Bbe), *Branchiostoma floridae* (Bfl), *Branchiostoma lanceolatum* (Bla). **b**, Phylogenetic analysis of chordate *Ets1/2*, using cephalochordate sequences as outgroup, suggested that the two *Ets1/2* genes of appendicularians were co-orthologs to the ascidian *Ets1/2a*. **c**–**h**, Whole mount in situ hybridization of *O. dioica* *Ets1/2a1* did not show any clear expression before hatchling stages (**c**–**f**). In early-hatchling stage *Ets1/2a1* revealed expression in the migratory endodermal strand cells (pink arrowheads) (**g**). In late-hatchling the expression signal was restricted to the buccal gland (green arrowheads) (**h**). **i**,**j**, *Ets1/2a2* did not show expression until tailbud stage. **k**, **l**, In tailbud embryos, expression signal was detected in tail muscle cells (orange arrowheads), the notochord (red arrowheads) and the epidermis of the trunk (blue arrowheads). **m**, In early-hatchling expression signal continued in the tail muscle and the notochord and increased in the anal domain (yellow arrowhead). **n**, In late-hatchling stage, the *Ets1/2a2* expression covered the entire oikoplastic epithelium, and continued in the muscle cells of the tail. Large images from tailbud in advance correspond to left lateral views oriented anterior towards the left and dorsal towards the top. Inset images are dorsal views of optical cross sections at the levels of dashed lines.

Extended Data Fig. 4 | See next page for caption.

Article

Extended Data Fig. 4 | FGF/MAPK ML phylogenetic tree and expression.
a, ML phylogenetic tree of the *MEK* subfamilies in chordates revealing the loss of the *MEK4*, *MEK5* and *MEK1/2* subfamilies in appendicularians, but the surviving of *MEK3/6* and *MEK7* subfamilies. Scale bar indicates amino acid substitutions. Bootstrap values are shown in the nodes. Vertebrates (black): *Gallus gallus* (Gga), *Homo sapiens* (Hsa), *Latimeria chalumnae* (Lch), *Lepisosteus oculatus* (Loc); Ascidian tunicates (blue): *Botrylloides leachii* (Ble), *Botrylloides schlosseri* (Bsc), *Ciona robusta* (Cro), *Ciona savignyi* (Csa), *Halocynthia aurantium* (Hau), *Halocynthia roretzi* (Hro), *Molgula occidentalis* (Mocci), *Molgula occulta* (Moccu), *Molgula oculata* (Mocul), *Phallusia fumigata* (Pfu), *Phallusia mammillata* (Pma); Appendicularian tunicates (red): *Bathochordaeus sp.* (Bsp), *Fritillaria borealis* (Fbo), *Mesochordaeus erythrocephalus* (Mer), *Oikopleura albicans* (Oal), *Oikopleura dioica* (Odi), *Oikopleura longicauda* (Olo), *Oikopleura vanhoeffeni* (Ova); Cephalo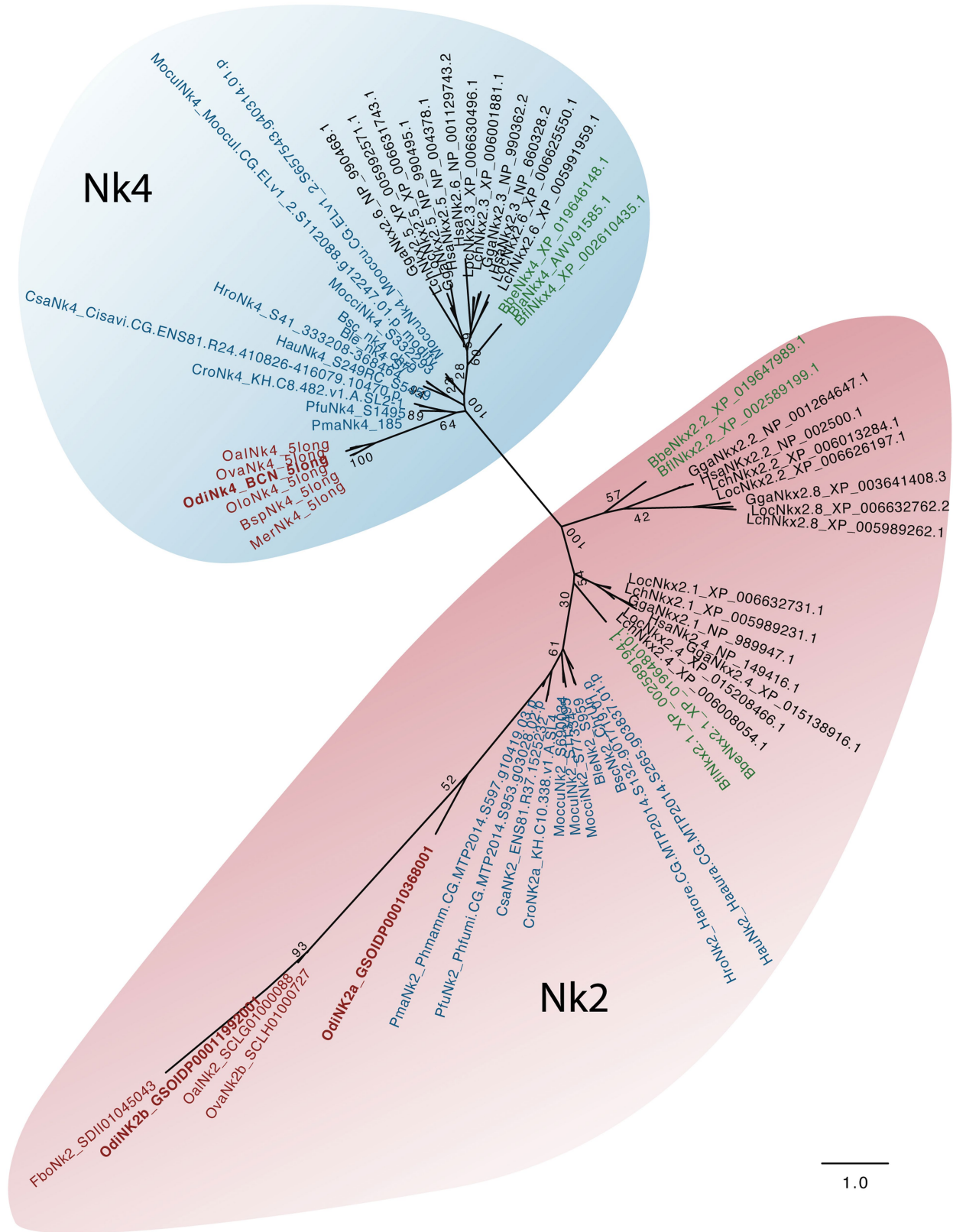chordates (green): *Branchiostoma belcheri* (Bbe), *Branchiostoma floridae* (Bfl), *Branchiostoma lanceolatum* (Bla). **b**–**g**, Whole mount in situ hybridization of *ERK* homolog in different stages of *O. dioica* development did not detect expression in any studied stage (**b**–**f**) until late-hatchling when expression was detected in an specific central domain in the oikoplastic epithelium (blue arrowheads) (**g**). **h**–**m**, Whole mount in situ hybridization of *MEK7* homolog in *O. dioica* revealed expression in the developing neural tissue in tailbud stages (pink arrowheads) (**i**,**j**), and in the esophagus (green arrowhead) and the oikoplastic epithelium (blue arrowheads) in the late-hatchling stage (**m**). **n**–**s**, Whole mount in situ hybridization of *MEK3/6* homolog in different stages of *O. dioica* development did not show any obvious tissue specific expression domain in the trunk, but the signal was generalized, with the exception of muscle cells in the tail at late-hatchling stages. Images from tailbud in advanced correspond to left lateral views orientated anterior towards the left and dorsal towards the top.
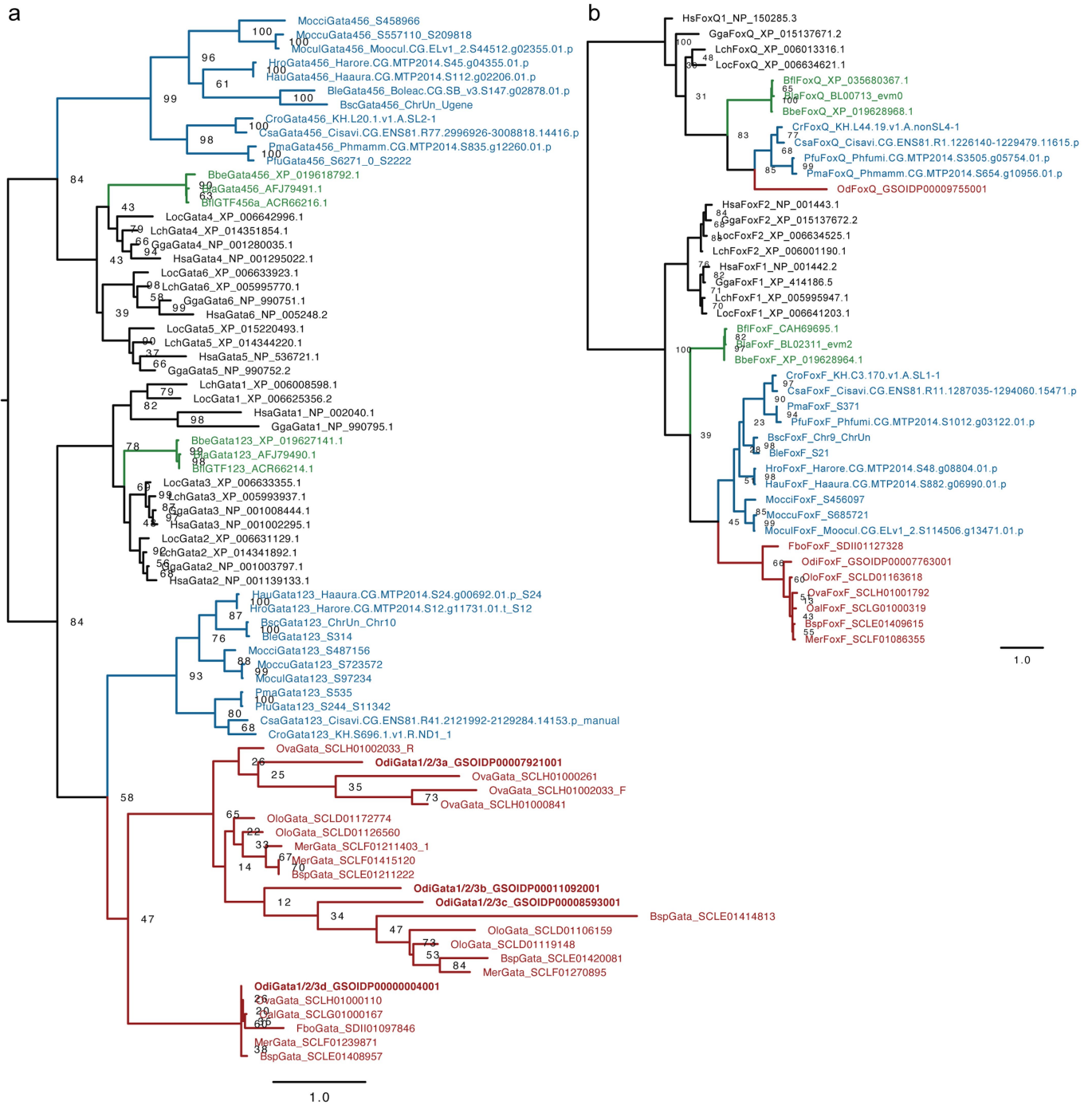
Extended Data Fig. 5 | Gata and FoxF ML phylogenetic trees in chordates.
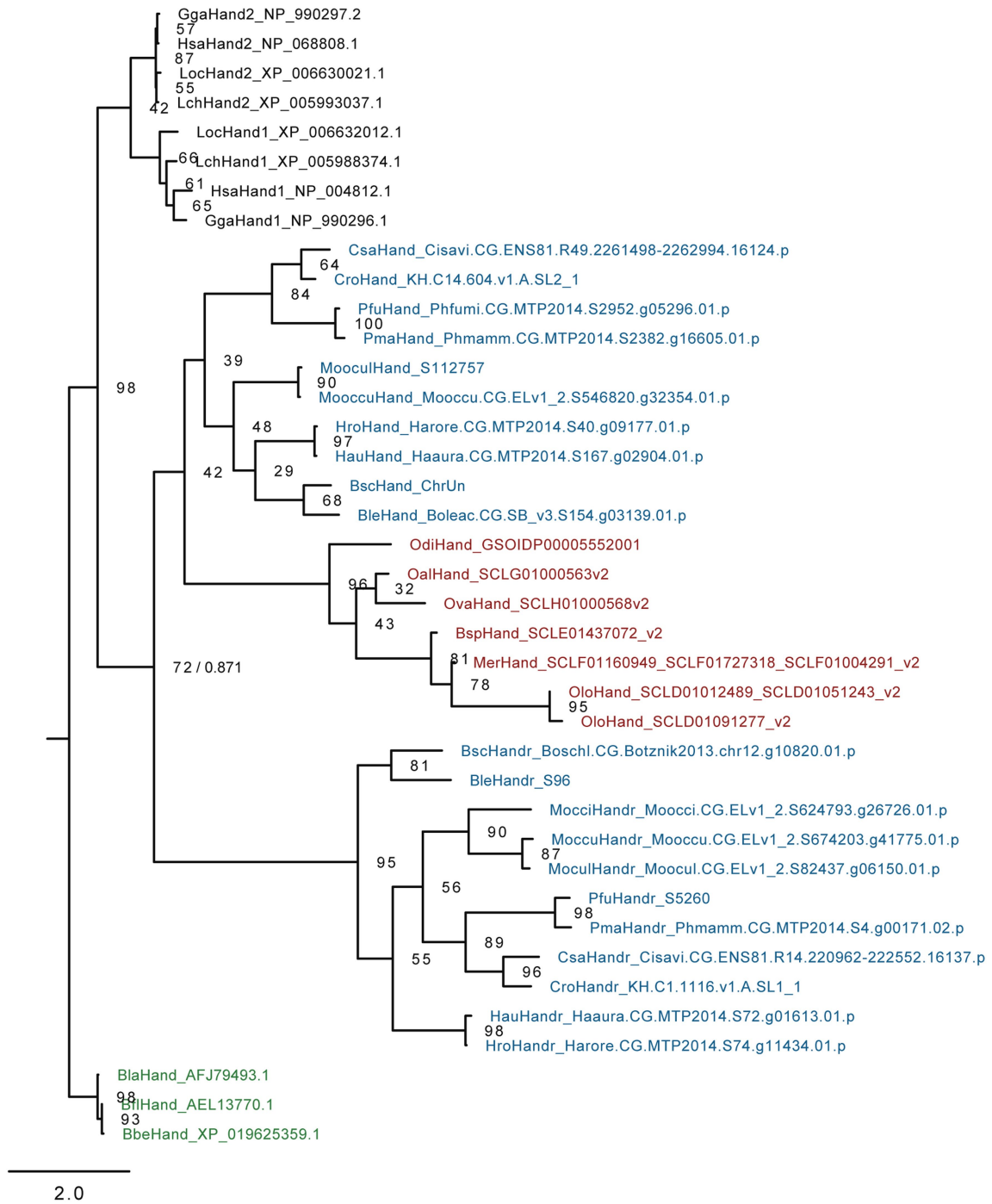**a**, *Gata* ML phylogenetic tree reveals the loss of the *Gata4/5/6* in appendicularians, but the surviving and lineage specific duplications of *Gata1/2/3* in appendicularians. **b**, *FoxF* ML phylogenetic tree reveals the presence of an ortholog of *FoxF* in appendicularians. The sister *FoxQ* subfamily was used as outgroup to root the tree. Scale bar indicates amino acid substitutions. Bootstrap values are shown in the nodes. Vertebrates (black): *Gallus gallus* (Gga), *Homo sapiens* (Hsa), *Latimeria chalumnae* (Lch), *Lepisosteus oculatus* (Loc); Ascidian tunicates (blue): *Botrylloides leachii* (Ble),

Botrylloides schlosseri (Bsc), *Ciona robusta* (Cro), *Ciona savignyi* (Csa), *Halocynthia aurantium* (Hau), *Halocynthia roretzi* (Hro), *Molgula occidentalis* (Mocci), *Molgula occulta* (Moccu), *Molgula oculata* (Mocul), *Phallusia fumigata* (Pfu), *Phallusia mammillata* (Pma); Appendicularian tunicates (red): *Bathochordaeus sp.* (Bsp), *Fritillaria borealis* (Fbo), *Mesochordaeus erythrocephalus* (Mer), *Oikopleura albicans* (Oal), *Oikopleura dioica* (Odi), *Oikopleura longicauda* (Olo), *Oikopleura vanhoeffeni* (Ova); Cephalochordates (Green): *Branchiostoma belcheri* (Bbe), *Branchiostoma floridae* (Bfl), *Branchiostoma lanceolatum* (Bla).
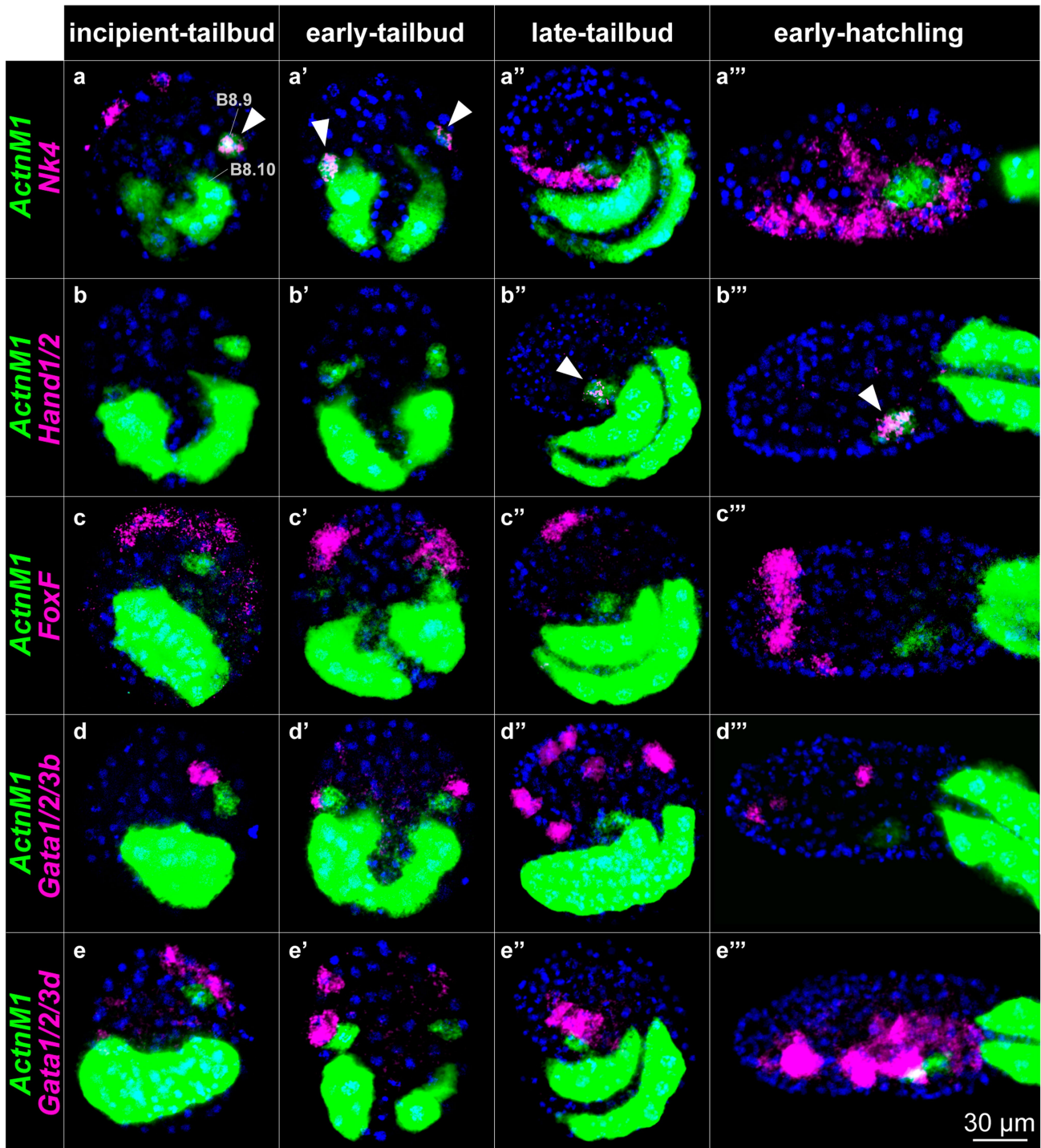
Extended Data Fig. 6 | NK ML phylogenetic tree in chordates reveals the presence of an ortholog of *Nk4* in appendicularians and two orthologs of the Nk2 subfamily. Scale bar indicates amino acid substitutions. Bootstrap values are shown in the nodes. Vertebrates (black): *Gallus gallus* (Gga), *Homo sapiens* (Hsa), *Latimeria chalumnae* (Lch), *Lepisosteus oculatus* (Loc); Ascidian tunicates (blue): *Botrylloides leachii* (Ble), *Botrylloides schlosseri* (Bsc), *Ciona robusta* (Cro), *Ciona savignyi* (Csa), *Halocynthia aurantium* (Hau), *Halocynthia roretzi* (Hro), *Molgula occidentalis* (Mocci), *Molgula occulta* (Moccu), *Molgula oculata* (Mocul), *Phallusia fumigata* (Pfu), *Phallusia mammillata* (Pma); Appendicularian tunicates (red): *Bathochordaeus sp.* (Bsp), *Fritillaria borealis* (Fbo), *Mesochordaeus erythrocephalus* (Mer), *Oikopleura albicans* (Oal), *Oikopleura dioica* (Odi), *Oikopleura longicauda* (Olo), *Oikopleura vanhoeffeni* (Ova); Cephalochordates (Green): *Branchiostoma belcheri* (Bbe), *Branchiostoma floridae* (Bfl), *Branchiostoma lanceolatum* (Bla).

Extended Data Fig. 7 | Hand ML phylogenetic tree suggests that member of this family in *O. dioica* is homologous to ascidian *Hand1/2*. Despite the tree suggests that the second paralog of ascidian (*Hand-r*) arose by a duplication at the base of the tunicate clade, and therefore subsequently lost in appendicularians. The low node support –bootstrap and approximate likelihood-ratio test (aLRT)– and the presence of shared long amino acid domain rich in K between the *Hand1/2* and *Hand-r* in ascidians, but absent in appendicularians, do not allow us to discard the possibility that *Hand-r* was originated by a duplication within the ascidian lineage, and its basal branching in the tunicate clade is due to a long branch attraction phenomenon. Scale bar indicates amino acid substitutions. Bootstrap values are shown in the nodes.

Vertebrates (black): *Gallus gallus* (Gga), *Homo sapiens* (Hsa), *Latimeria chalumnae* (Lch), *Lepisosteus oculatus* (Loc); Ascidian tunicates (blue): *Botrylloides leachii* (Ble), *Botrylloides schlosseri* (Bsc), *Ciona robusta* (Cro), *Ciona savignyi* (Csa), *Halocynthia aurantium* (Hau), *Halocynthia roretzi* (Hro), *Molgula occidentalis* (Mocci), *Molgula occulta* (Moccu), *Molgula oculata* (Mocul), *Phallusia fumigata* (Pfu), *Phallusia mammillata* (Pma); Appendicularian tunicates (red): *Bathochordaeus sp.* (Bsp), *Fritillaria borealis* (Fbo), *Mesochordaeus erythrocephalus* (Mer), *Oikopleura albicans* (Oal), *Oikopleura dioica* (Odi), *Oikopleura longicauda* (Olo), *Oikopleura vanhoeffeni* (Ova); Cephalochordates (green): *Branchiostoma belcheri* (Bbe), *Branchiostoma floridae* (Bfl), *Branchiostoma lanceolatum* (Bla).

Extended Data Fig. 8 | Developmental coexpression patterns of *ActnM1* and potential cardiac transcription factors. Double fluorescent in situ hybridization of *ActnM1* with *Nk4*, *Hand1/2*, *FoxF*, *Gata1/2/3b* and *Gata1/2/3d*. *Nk4* expression signal was detected in ventral epidermis and the CPC (B8.9) from the incipient-tailbud stage (**a**) until the early-tailbud (**a'**). In later stages, we only detected expression in the epidermis, but not in the cardiac precursors (**a''**–**a'''**). *Hand1/2* was specifically expressed in the cardiac progenitors from late-tailbud to hatchling stages (**b''**–**b'''**). We did not detect expression of *FoxF*, *Gata1/2/3b* nor *Gata1/2/3d* in cardiac precursors, but they were expressed in different epidermal domains (**c**–**e'''**). The images correspond to the overlay of a stack of confocal sections with expression of the different genes. The small overlapping color in **e'''** is due to the overlay of the stack, and not to actual co-expression. White arrowheads indicate co-expression of *ActnM1* with the corresponding gene in cardiac progenitors. Incipient- and early-tailbud stages correspond to ventral views oriented anterior towards the top. Late-tailbud and early-hatchling stages correspond to lateral views oriented anterior towards the left and dorsal towards the top.

From 2-cell stage

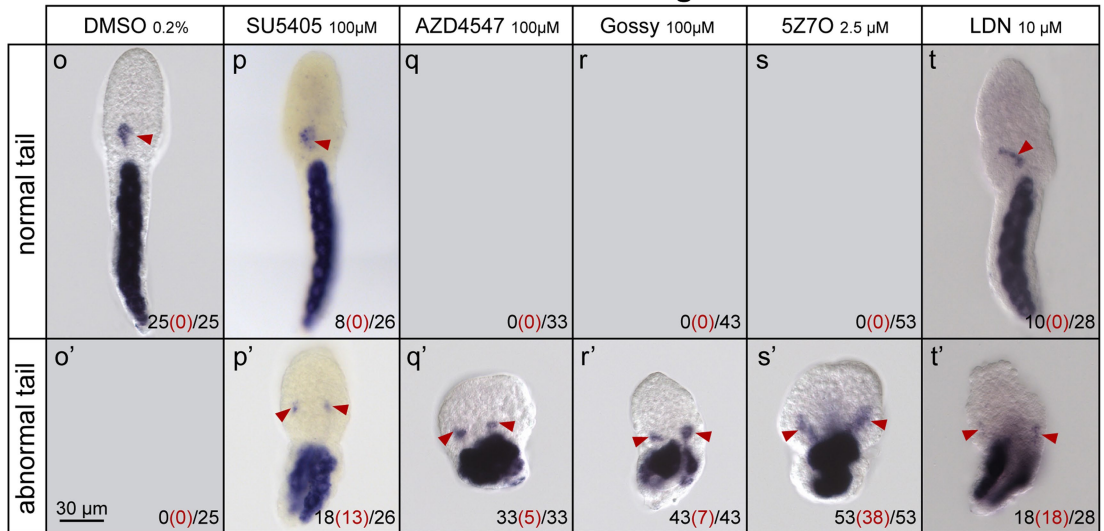| DMSO 0.2%	SU5402 50 μM	AZD4547 50 μM	Gossy 100 μM	5Z7O 1 μM	LDN 10 μM	DORSO 10 μM
a 20/20	b 8/51	c 28/48	d 30/53	e 29/52	f 21/26	g 20/22
30 μm	b' 13/51	c' 19/48	d' 16/53	e' 11/52	f' 5/26	g' 2/22
	b'' 30/51	c'' 1/48	d'' 7/53	e'' 12/52	0/26	0/22
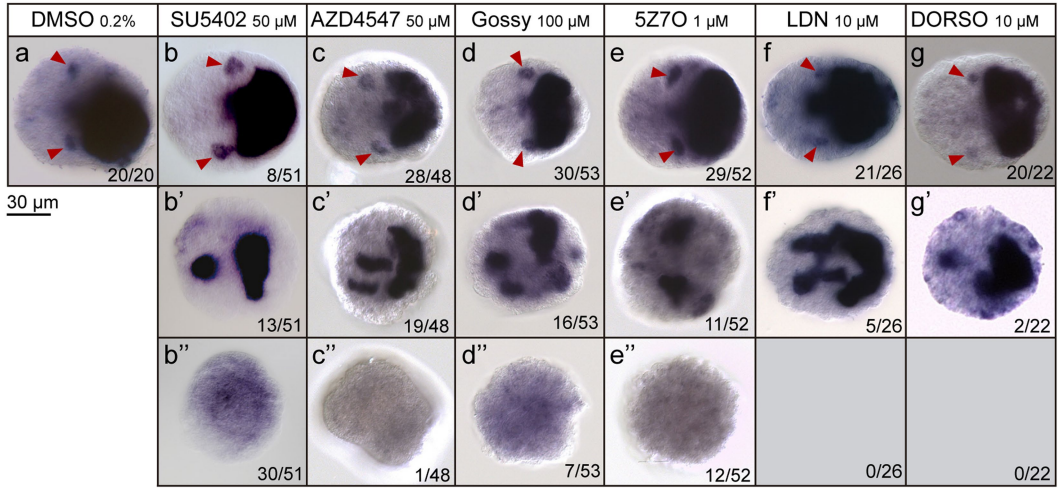
ActnM1

From 32-cell stage

DMSO 0.2%	SU5402 100 μM	AZD4547 100μM	Gossy 100 μM	5Z7O 2.5 μM	LDN 10 μM	DORSO 10 μM
h 20/24	i 30/49	j 13/24	k 19/21	l 20/28	m 21/24	n 16/19
30 μm	i' 19/49	j' 11/24	k' 2/21	l' 8/28	m' 3/24	n' 3/19
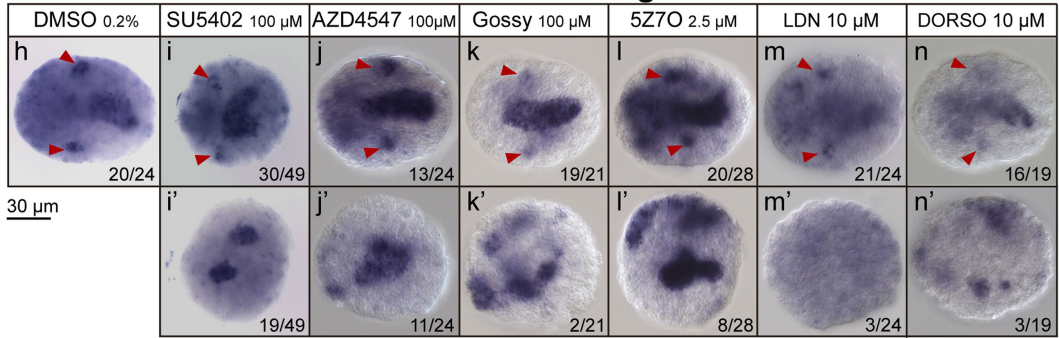
Nk4 + Brachyury

From 32-cell stage

|	DMSO 0.2%	SU5405 100μM	AZD4547 100μM	Gossy 100μM	5Z7O 2.5 μM	LDN 10 μM
normal tail	o 25(0)/25	p 8(0)/26	q 0(0)/33	r 0(0)/43	s 0(0)/53	t 10(0)/28
abnormal tail	o' 0(0)/25	p' 18(13)/26	q' 33(5)/33	r' 43(7)/43	s' 53(38)/53	t' 18(18)/28

(scale bar 30 μm)

ActnM1

Extended Data Fig. 9 | See next page for caption.

Article

Extended Data Fig. 9 | FGF, MEK and BMP inhibition during heart development of *O. dioica*. a–**g'**, Whole mount in situ hybridization of *ActnM1* in DMSO-control (**a**) and treated embryos with inhibitors of FGFR (SU5402 and AZD4547), MEK3/6 (Gossypetin), MEK7 (5Z-7-Oxozeaenol) and BMP inhibitors (LDN and Dorsomorphin) from 2-cell stage up to early-tailbud stage (**b**–**g'**). Embryos treated with FGFR and MEK inhibitors affected gastrulation and caused abnormal phenotypes in which mesodermal derivatives showed either abnormal domains (**b'**–**e'**) or complete absence (**b''**–**e''**). However, those treated embryos that reached fairly normal incipient morphologies (**b**–**g**), showed the presence of CPCs (red arrowheads). **h**–**n'**, Whole mount in situ hybridization of *NK4+Brachyury* in DMSO-control (**h**) and treated embryos with FGFR, MEK and BMP inhibitors (**i**–**n'**) from 32-cell stage to early-tailbud stage. A majority of the treated embryos showed the *Nk4* expression in the CPCs (**i**–**n**), even in some with obvious abnormalities in the notochord (**i**). Only in embryos with severe abnormal morphologies or arrested, we could not distinguish the CPCs from other *Nk4* expression domains (**i'**–**n'**). **o**–**t'**, Whole mount in situ hybridization of *ActnM1* in DMSO-control (**o**) and treated embryos with FGFR, MEK and BMP inhibitors from 32-cell stage to early-hatchling stage (**p**–**t'**). Most of the treated embryos showed abnormal tails (**p'**–**t'**), in which the elongation and rotation had been affected. Moreover, while the CPCs had converged near the midline into a single cardiac field, we observed that in many embryos with tail malformations, the CPCs had not converged and were still bilaterally separated at the right and left sides of the trunk (red numbers in brackets). These results suggests that FGF/MEK/MAPK and BMP signaling pathways may be involved in tail elongation/rotation and late cardiac organogenesis. Tailbud embryos images correspond to dorsal views with anterior to the left. Hatchling images represent dorsal views with anterior to the top.
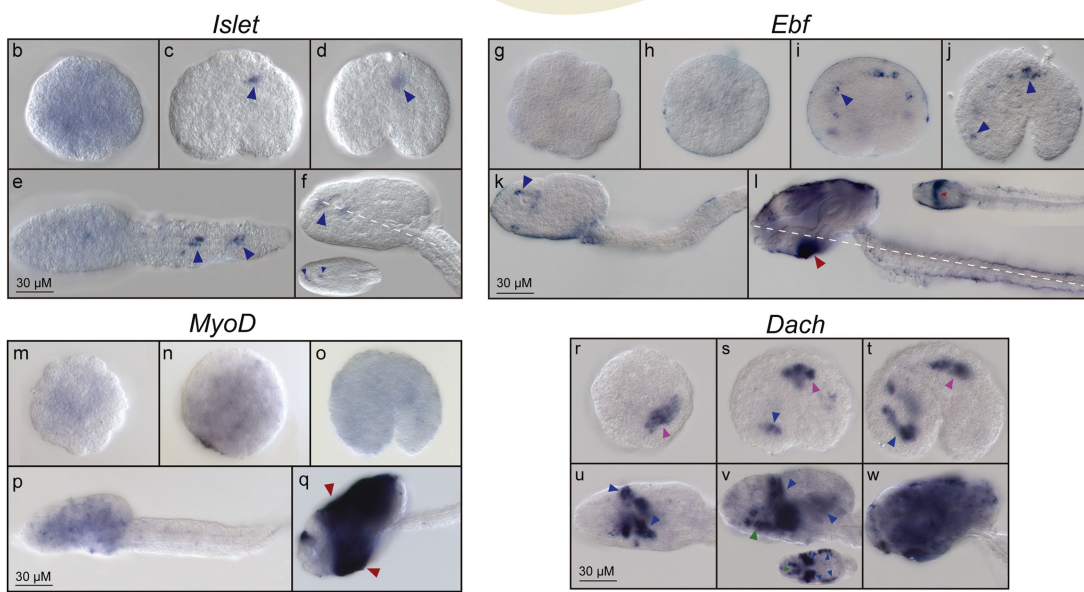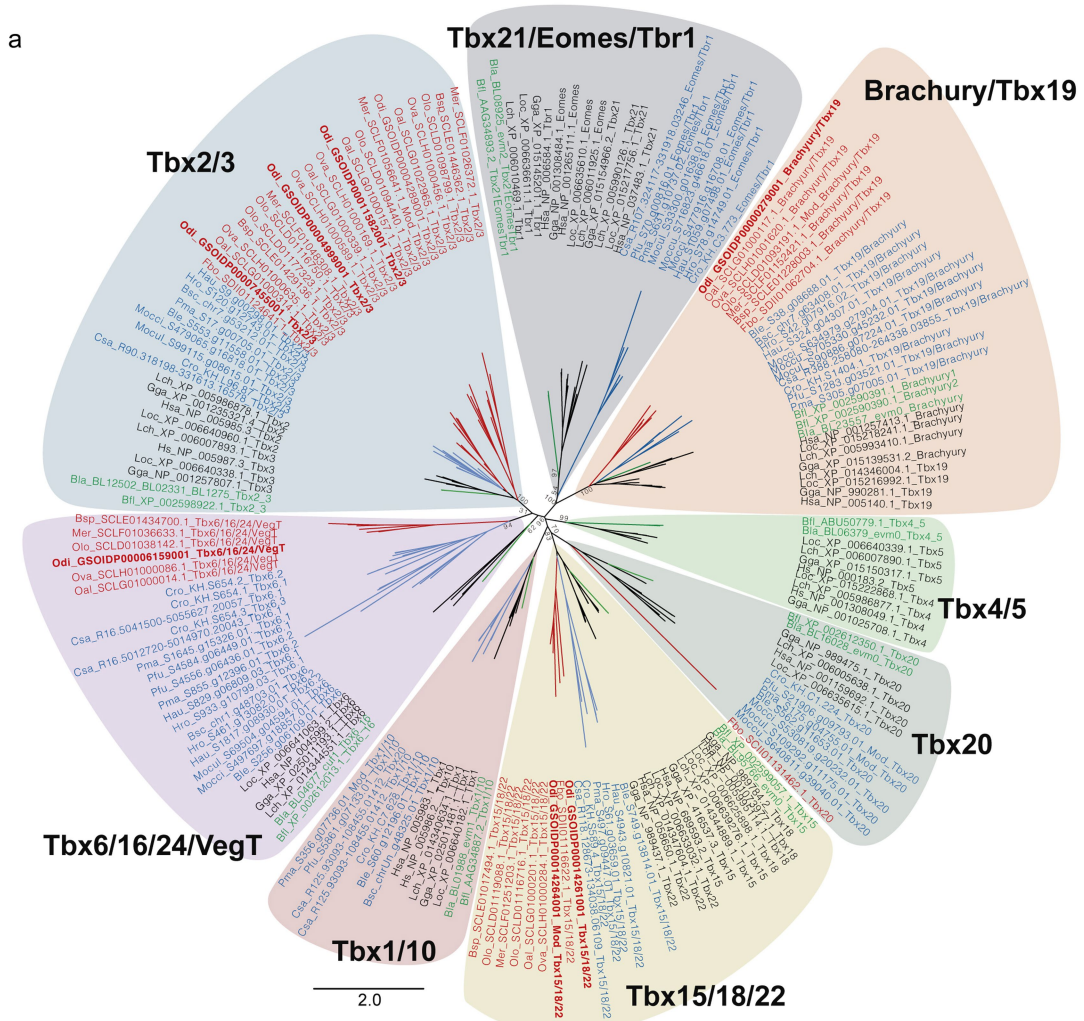
Extended Data Fig. 10 | See next page for caption.

Article

Extended Data Fig. 10 | Tbx ML phylogenetic tree and *Islet, Ebf, MyoD* and *Dach* expression. a, ML phylogenetic tree of the Tbx subfamilies in chordates reveals the loss of *Tbx1/10* and *Tbx21/Eomes/Tbr1* subfamilies in appendicularians and the ancestral loss of *Tbx4/5* subfamily in tunicates. Scale bar indicates amino acid substitutions. Bootstrap values are shown in the nodes. Vertebrates (black): *Gallus gallus* (Gga), *Homo sapiens* (Hsa), *Latimeria chalumnae* (Lch), *Lepisosteus oculatus* (Loc); Ascidian tunicates (blue): *Botrylloides leachii* (Ble), *Botrylloides schlosseri* (Bsc), *Ciona robusta* (Cro), *Ciona savignyi* (Csa), *Halocynthia aurantium* (Hau), *Halocynthia roretzi* (Hro), *Molgula occidentalis* (Moocci), *Molgula occulta* (Mooccu), *Molgula oculata* (Moocul), *Phallusia fumigata* (Pfu), *Phallusia mammillata* (Pma); Ascidian appendicularians (red): *Bathochordaeus sp.* (Bsp), *Fritillaria borealis* (Fbo), *Mesochordaeus erythrocephalus* (Mer), *Oikopleura albicans* (Oal), *Oikopleura dioica* (Odi), *Oikopleura longicauda* (Olo), *Oikopleura vanhoeffeni* (Ova); Cephalochordates (green): *Branchiostoma floridae* (Bfl), *Branchiostoma lanceolatum* (Bla). **b–w**, Whole mount in situ hybridization of *O. dioica Islet, Ebf, MyoD* and *Dach* homologs. 64-cell embryos did not showed expression of *Islet* (**b**) which was only detected in the developing nervous system from tailbud to hatchling embryos (**c–f**). *Ebf* (COE) did not show expression in early stages (**g**, **h**) but we detected expression in the nervous system from tailbud to mid-hatchling stage (**i–k**) and in the oikoplastic epithelium of late-hatchling embryos (**l**). We did not detect expression of *MyoD* from 32-cell to hatchling embryos (**m–p**). In late-hatchling embryos *MyoD* was expressed in the oikoplastic epithelium (**q**). *Dach* expression started at the 64-cell stage in the developing nervous system (pink arrowheads) and continued until late-tailbud stage (**r–t**). In tailbud stages, *Dach* started expressing in the trunk epidermis (blue arrowheads) which was maintained until late-hatchling stages when it was expressed in the whole oikoplastic epithelium (blue arrowheads) (**s–v**). In mid-hatchling stage, beside the epidermis, *Dach* expression was also detected in the endostyle (green arrowheads) (**w**). Large images from tailbud in advance correspond to left lateral views oriented anterior towards the left and dorsal towards the top. Inset images are dorsal views of optical cross sections at the levels of dashed lines. Pink arrowheads indicate the developing nervous system. Blue arrowheads indicate the oikoplastic epithelium.