

1 **Main Manuscript for**
2 **Coevolution of the *Tlx* homeobox gene with medusa development**
3 **(Cnidaria: Medusozoa)**

4 Matthew Traver¹, Reed Boohar^{1*}, Steven M. Sanders², Matthew L. Nicotra² Lucas Leclère³,
5 Robert E. Steele⁴, Pauly Cartwright¹

6 ¹University of Kansas, ²University of Pittsburgh, ³Sorbonne Université, CNRS, Laboratoire de
7 Biologie du Développement de Villefranche-sur-Mer (LBDV), Villefranche-sur-Mer, France,

8 ⁴Department of Biological Chemistry, University of California, Irvine

9 *Current affiliation: Department of Biology, University of Miami, Coral Gables, Florida

10 Matthew Traver

11 **Email:** traver.matthew@gmail.com

12 **Author Contributions:** Pauly Cartwright, Steven Sanders, Robert Steele, Matthew Traver
13 designed research; Reed Boohar, Steven M. Sanders and Matthew Traver performed research;
14 Lucas Leclère, Matthew L. Nicotra, Steven M. Sanders and Matthew Traver contributed new
15 reagents/analytic tools; Pauly Cartwright, Lucas Leclère and Matthew Traver analyzed data;
16 Pauly Cartwright and Matthew Traver wrote the paper with input from the authors.

17 **Competing Interest Statement:** The authors declare no conflict of interest.

18 **Classification:** Biological Sciences, Evolution

19 **Keywords:** Hydrozoa, *Podocoryna*, life cycle, genome evolution, jellyfish

20 **This PDF file includes:**

21 Main Text

22 Figures 1 to 4

23 Tables 1 to 2

24

1 **Abstract**

2 The jellyfish, or medusa, is a life cycle stage characteristic of the cnidarian subphylum Medusozoa.
3 By contrast, the other cnidarian subphyla Anthozoa and Endocnidozoa lack a medusa stage. Of
4 the medusozoan classes, Hydrozoa is the most diverse in terms of species number and life cycle
5 variation. A notable pattern in hydrozoan evolution is that the medusa stage has been lost or
6 reduced several times independently. Although this loss of the jellyfish stage is thought to be due
7 to heterochrony, the precise developmental mechanisms underlying this complex pattern of
8 medusa evolution are unknown. We found that the presence of the homeobox gene *Tlx* in cnidarian
9 genomes is correlated with those medusozoans that have a medusa stage as part of their life cycle.
10 Although *Tlx* is conserved in Bilateria and Cnidaria, it is missing in the genomes of anthozoans,
11 endocnidozoans, and those hydrozoans that have lost the medusa stage. Selection analyses of *Tlx*
12 across medusozoans revealed that hydrozoans undergo relatively relaxed selection compared to
13 the other medusozoan classes, which may in part explain the pattern of multiple medusa losses.
14 Differential expression analyses on three distantly related medusozoan representatives indicate an
15 upregulation of *Tlx* during medusa development. In addition, *Tlx* expression is spatially restricted
16 to regions of active development in medusae of the hydrozoan *Podocoryna carnea*. Our results
17 suggest that *Tlx* plays a key role in medusa development and that the loss of this gene is likely
18 linked to the repeated loss of the medusa life cycle stage.

19 **Main Text**

20 **Introduction**

21 Jellyfish represent part of a complex life cycle that is characteristic of the cnidarian subphylum
22 Medusozoa, which includes true jellyfish (Scyphozoa), box jellyfish (Cubozoa), stalked jellyfish
23 (Staurozoa) and hydromedusae (Hydrozoa). Medusozoans possess a metagenetic life cycle
24 alternating between an asexual phase in the form of a sessile polyp and a sexually reproducing,
25 typically pelagic, phase called the medusa or jellyfish. The medusa exhibits several distinct features
26 including a bell-shaped morphology, striated muscles, gonads, and sensory organs. The parasitic
27 Endocnidozoa is sister to Medusozoa (Chang et al., 2015) and lacks a definitive polyp or medusa
28 stage. The other major cnidarian class, Anthozoa (corals, anemones and sea pens) also lacks a
29 medusa stage as well as all traits associated with this free-living stage, and instead possesses a
30 monogenetic life cycle, comprising only sexually and asexually reproducing sessile polyps. Given
31 that these other two major classes lack medusae characters, it can be assumed that medusae are
32 a derived feature of Medusozoa and not ancestral to Cnidaria.

33 Despite the medusa being characteristic of Medusozoa, losses of this life cycle stage within
34 Hydrozoa are omnipresent and most likely are due to developmental heterochrony (Boero and
35 Bouillon 1987; Boero et al., 1992; Kubota, 2000). The fully developed hydromedusa exhibits
36 distinctive features such as a muscular structure at the bell margin called a velum, striated muscles,
37 marginal tentacles and tentacle bulbs, a manubrium which contains the gut and mouth, and a
38 gastrovascular system composed of radial and circular canals. The truncation of the medusa stage
39 correlates with the loss or reduction of the aforementioned features, with the varying degrees of
40 medusa truncation often mirroring stages of medusa development (Hincks 1868, Allman 1872,
41 Goette 1916 and Namikawa 1991). The differing degrees of medusa truncation across species is
42 a type of paedomorphic progenesis (Gould 1985, Cunningham and Buss 1993), where somatic
43 development is truncated due to earlier sexual maturation.

44 While the development of the hydrozoan polyp has been well studied, particularly in the model
45 systems *Hydra* (Galliot, 2012) and *Hydractinia* (Frank et al., 2020), the molecular mechanisms
46 underlying the development of the hydromedusa remain poorly understood. While some studies

1 have indicated that medusa development co-opts key developmental pathways functioning in the
2 polyp (Sanders & Cartwright, 2015a, b; Kraus et al., 2015, Masuda-Nakagawa et al., 2000, Müller
3 et al., 1999), recent studies suggest that some medusa-specific transcription factors might act as
4 molecular switches that regulate aspects of medusa development (Leclère et al, 2019 and Kahlturn
5 et al, 2019). Amongst identified “medusa specific” genes a large proportion are homeobox genes,
6 suggesting that these genes may play a key role in the development and maintenance of the
7 medusa.

8 One purported “medusa specific gene”, is the T Cell Leukemia Homeobox gene (*Tlx*) (Leclère et
9 al, 2019). In vertebrates *Tlx* (also referred to as *HOX11*) is involved in spleen organogenesis, brain
10 and skeleton patterning and in *Drosophila melanogaster* it is involved in distal patterning of the leg
11 (*clawless*) (Lenti et al., 2016, Kojima et al., 2005). *Tlx* could not be found in the genome of the sea
12 anemone *Nematostella vectensis* (Kamm and Schierwater, 2006) and thus was thought to be
13 absent in cnidarians. Here we show that *Tlx* is indeed present in cnidarians. However, our survey
14 of this gene across cnidarian genomes found that it is present only in those cnidarian lineages
15 exhibiting a medusa stage. *Tlx* is absent in all anthozoans and endocnidozoan genomes surveyed.
16 In addition, an intact *Tlx* is absent in nearly all hydrozoans surveyed that have lost a medusa life
17 cycle stage. In several distantly related medusae-bearing species, we found that *Tlx* expression is
18 upregulated during medusa development and its expression is consistent with it playing a role in
19 medusae patterning.

20 Results

21 The cnidarian *Tlx* ortholog shares a highly conserved genomic structure 22 with bilaterian *Tlx*.

23 Cnidarian and bilaterian *Tlx* genes are remarkably conserved in structure, including an EH1 domain
24 near the N-terminus, the homeodomain, and its signature motif, RRIGHPY just upstream of the
25 homeodomain, called the N-terminal arm (Figure 1A). In our search of public databases for *Tlx*, the
26 EH1 domain, as well as the N-terminal arm were invariably found in complete sequences of *Tlx* in
27 both cnidarians and bilaterians. However, these conserved regions could not be found in the
28 putative *Tlx* orthologs of sponges and placozoans, and although ctenophore sequences had an
29 EH1 domain, they lacked the N-terminal arm. In addition to low sequence identity in the
30 homeodomain, a Bayesian phylogenetic analysis did not recover the putative *Tlx-like* gene
31 previously identified in ctenophores (Pang and Martindale, 2008 and Derelle and Manuel, 2007),
32 within the strongly supported bilaterian and cnidarian *Tlx* orthology group (Figure S1). This
33 suggests that *Tlx* arose in the last common ancestor of Cnidaria + Bilateria. Vertebrates possess
34 three *Tlx* paralogs, suggesting two duplication events in the last common ancestor of vertebrates.
35 A phylogenetic tree consisting of 38 cnidarian taxa and 11 vertebrate taxa including their three
36 paralogs is shown in Figure 1B. *Tlx* forms a well-supported orthology group (BS=83, PP=0.99).
37 Although vertebrate *Tlx* paralogy groups are respectively well supported, the relationship between
38 cnidarian *Tlx* to a specific vertebrate paralog could not be recovered (TLX1/3) with sufficient
39 support. A phylogenetic analysis including select protostome, cnidarian and vertebrate *Tlx* genes
40 also formed a well support *Tlx* orthology group in the Bayesian analysis (PP=0.98) but failed to
41 recover specific orthology relationships between major taxa (Figure S1).

42 *Tlx* is absent from available genomes of cnidarians lacking a medusa.

43 We define the presence of a medusa by the presence of medusa specific characters. Reduced
44 medusae are often referred to as eumedusoids, cryptomedusoids or sporosacs (Bouillon et al.,
45 2006) depending on their degree of developmental arrest. Here we call reduced medusae
46 eumedusoids if they exhibit a gastrovascular system, velum and tentacles but lack discrete gonads
47 and a mouth, cryptomedusoids if they bear only radial canals and highly reduced tentacle
48 processes and sporosacs if they represent a fixed gonophore lacking any medusa features. Some
49 hydrozoans, such as *Hydra*, do not bear any gonophores and instead release their gametes directly

1 from the body column of the polyp. Given that eumedusoids possess many medusa-specific
2 characteristics, we consider those species bearing eumedusoids as having a medusa stage,
3 whereas those bearing cryptomedusoids, sporosacs or absence of any gonophore, we consider
4 lacking a medusa stage.

5 In our search for the *Tlx* gene in 70 publicly available cnidarian draft genome assemblies we found
6 that the presence of *Tlx* is invariably correlated with the presence of a medusa in the cnidarian life
7 cycle. Specifically, *Tlx* was found in all 27 of the draft genomes from species that have medusae
8 and not found in any of the 43 available draft genomes from species that lack medusae, including
9 all anthozoans and endocnidozoans and the six hydrozoans that lost the medusae stage (Table 1).
10 Although Ryan et al (2006) reported a *Tlx* gene in the sea anemone *Nematostella*, this gene lacks
11 the signature EH1 domain and N-terminal arm and was not recovered in the *Tlx* orthology group
12 with sufficient support in their analysis.

13 *Tlx* was also searched for in 188 cnidarian transcriptomes, wherein the gene was present in 47 of
14 75 transcriptomes from medusa-bearing species (Table. 1). The failure to recover *Tlx* in some of
15 the transcriptomes of medusa-bearing species does not necessarily indicate its absence from the
16 genome, as transcriptomes are a limited subset of expressed genes of the tissue sampled.
17 Therefore, the absence of *Tlx* in some of the sampled medusa-bearing species is likely the result
18 of the particular tissue and/or developmental stage from which the transcriptome was generated.
19 *Tlx* was not present in any available transcriptomes from species that lack a medusa, with three
20 exceptions (*Millepora squarrosa*, *Ectopleura larynx* and *Dynamena pumila*) (Table 1). *Millepora*
21 *squarrosa* *Tlx* is likely a pseudogene as it exhibits premature stop codons and the *Dynamena*
22 *pumila* sequence was a partial sequence lacking the EH1 domain and the C-terminus of the
23 homeodomain and thus the functionality cannot be inferred. *Ectopleura larynx* has the characteristic
24 *Tlx* domains. It does however exhibit a unique codon insertion, leading to a threonine in the highly
25 positively charged N-terminus of the homeodomain. Whether this change could affect *E. larynx* *Tlx*
26 function is unknown.

27 Due to the sampling restrictions imposed by screening *Tlx* from publicly available cnidarian
28 genomes, and the limitations of transcriptomes for estimating the presence of a gene as discussed
29 above, we used degenerate PCR to screen genomic DNA from 100 medusozoan taxa for the
30 presence of *Tlx*, including all medusozoan suborders, in order to span the breadth of medusozoan
31 diversity. The primers were designed to amplify the N-terminal arm and the entire homeobox region.
32 In the 69 taxa surveyed that have a medusa (or eumedusoid), a *Tlx* gene fragment was successfully
33 amplified in 58 (84%). The failure to amplify a *Tlx* fragment in the other 11 medusa-bearing taxa
34 could be due to the limitations of degenerate PCR, which is highly sensitive to DNA quality and
35 primer binding. An amplification product was not obtained in 28 out of 31 taxa (90%) that lack a
36 medusa (sporosac, cryptomedusoid or no gonophore). The three non-medusa bearing species for
37 which a *Tlx* fragment was recovered were the cryptomedusoid bearing *Ectopleura larynx*, also
38 found in the transcriptome above, as well as two species that bear sporosacs (*Amphisbetia minima*
39 and *Sertularia perpusilla*). The sequence from *Sertularia perpusilla*, like the sequence of *Millepora*
40 *squarrosa* discussed above, is likely a pseudogene as it contains several premature stop codons.
41 Thus, of the total of five TLX sequences isolated from non-medusae bearing species from
42 transcriptomes and/or PCR, only *Amphisbetia minima* has a typical TLX sequence.

43 **The absence of the *Tlx* gene is correlated with the absence of the medusa in** 44 **Hydrozoa.**

45 Using the medusozoan phylogenetic tree from Cartwright and Nawrocki (2010) we pruned the taxa
46 to match those samples for which degenerate PCR data were generated and reconstructed the
47 evolution of the medusa stage. Similar to findings of Cartwright and Nawrocki (2010) our analysis
48 inferred several independent instances of reduction and loss of the medusa stage. Out of the 14
49 such cases, seven are reductions to sporosacs, four to eumedusoid, one to cryptomedusoid and
50 two are complete losses of the gonophore (Figure 2).

1 A Bayesian correlation analysis of the presence of *Tlx* and the presence of medusa-like structures
2 (medusa and eumedusoid) shows a very strong correlation between the two traits (Log Bayes factor
3 = 19.535592). The same analysis also supports the presence of TLX together with the medusa
4 stage as ancestral in Medusozoa (PP=1) (Figure 2).

5 **TLX homeodomain shows evidence of relaxed selection in species lacking** 6 **medusa-like structures.**

7 To further investigate the apparent conservation of the TLX homeodomain amongst medusa
8 bearing and the few non-medusa bearing lineages, we tested for relaxation/intensification of
9 selection on the *Tlx* homeodomain in a codon-based phylogenetic framework. Using selection
10 analyses, we tested the four *Tlx* sequences that were found from non-medusa-bearing species
11 (*Millepora squarrosa*, *Ectopleura larynx*, *Dynamena pumila*, *Amphisbetia minima*) for relaxed
12 selection. *S. perpusilla* sequence was removed from the analysis as the aberrant sequence did not
13 allow for proper codon alignment. When testing these four sequences against the 46 medusa
14 bearing reference species sequences, we found strong evidence for a relaxation of selection on
15 *Tlx* in those medusa-less lineages (K=0.15, p=0.0000). While testing those four lineages
16 independently, the same evidence of relaxation of selection was found except for *E. larynx* (K=1.19,
17 p=0.1715) for which a non-significant intensification was inferred. By contrast, a significant
18 intensification of selection on the *Tlx* homeodomain was detected for Acraspeda (Scyphozoa,
19 Cubozoa and Staurozoa) (K=1.66, p=0.0000) and the hydrozoan order Siphonophorae (K=2.30,
20 p=0.0050), using medusa-bearing hydrozoan species as a reference. No significant trend in
21 selection was detected for the other hydrozoan lineages (see Table 2). This relaxation of selection
22 in Hydrozoa may in part explain the pattern of multiple medusa losses that is not found in the other
23 medusozoan classes.

24 Lastly a FUBAR test (Fast, Unconstrained Bayesian Approximation, Murrel et al., 2013) was
25 performed to identify site specific variation in the selection of cnidarian TLX. Unsurprisingly, 9
26 codons in the EH1 domain and 75 codons in the homeodomain (including the N-terminal arm and
27 flanking regions), respectively, showed evidence for pervasive purifying selection (PP>0.99), and
28 no phylum-wide diversifying selection was detected upon remaining sites. Interestingly, the
29 analysis detected smaller motifs flanking the homeodomain and its N-terminal arm. Amongst
30 scyphozoans, a significant (PP>0.99) episodic purifying selection was detected on sites flanking
31 the homeodomain (PWQILXK upstream of the N-terminal arm and TEEEEKEEQRHAL downstream
32 of the homeodomain), while a significant (PP>0.99) episodic purifying selection was detected in the
33 flanking regions of the homeodomain of hydrozoan TLX, corresponding to a highly conserved CXC
34 motif upstream of the N-terminal arm and a EINEMXEQQXR motif downstream of the
35 homeodomain. Flanking motifs found in hydrozoans and scyphozoans do not provide direct
36 information regarding the binding target of the homeodomain but could suggest that the target or
37 the affinity for the target of TLX might differ between these lineages.

38 **Expression of *Tlx* is up-regulated during medusa development.**

39 To investigate the expression profile of *Tlx* in the medusozoan life cycle, we performed a differential
40 expression (DE) analysis on distinct life cycle stages for the scyphozoan *Aurelia coerulea*, and two
41 medusae-bearing hydrozoan species *Podocoryna carnea* and *Clytia hemisphaerica*. For *Aurelia*,
42 *Tlx* expression is first detected in the polyp (scyphostoma), peaks when the polyp is producing
43 medusae (strobilating) and is then downregulated in the juvenile medusa (ephyra) and mature
44 medusa stages to expression levels comparable to the scyphostoma (Figure 3a). In *Clytia*
45 *hemisphaerica*, *Tlx* expression is first detected at low levels during the planula stage and is
46 maintained at low level in the feeding polyp (gastrozoid). *Tlx* expression is upregulated in the
47 reproductive polyp that buds medusae (gonozoid) and the newly released juvenile medusa. In the
48 adult medusa of *Clytia*, the expression of *Tlx* is downregulated to an intermediate level (Figure 3b).
49 In *Podocoryna carnea* no significant expression of *PcTlx* is detected at the planula stage, the non-
50 reproductive polyp nor in the reproductive polyp when the medusae buds are initially detected.

1 *PcTlx* expression is upregulated in reproductive polyps budding medusae, during later stages of
2 medusae development and remains at this expression level in the fully developed medusa after it
3 is released from the polyp (Figure 3c). Although DE patterns were significant for the three species
4 using estimated counts, other metrics, namely Transcripts per million (TPM) and Fragments Per
5 Kilobase Million (FPKM), were inconclusive, likely due to the overall low expression of *Tlx*. To
6 validate the RNA-Seq results in *Podocoryna*, we performed RT-qPCR on planulae, non-
7 reproductive polyps, budding polyps and released medusae of *P. carnea* and found a significant
8 difference in the relative expression of *Tlx* between the four life cycle stages (ANOVA, $p < 0.0001$),
9 as well as a higher expression of *Tlx* in released medusae compared to budding polyps (t-test,
10 $p < 0.0001$), (see Figure 3d).

11 ***Tlx* expression is spatially restricted during medusa development.**

12 *Tlx* expression was detected by whole mount *in situ* hybridization in juvenile and mature medusae
13 of *Podocoryna carnea* but not in feeding and reproductive polyps (Figure 4A-C), or planulae (not
14 shown). In *Podocoryna*, the gonads develop on the manubrium (the structure that contains the
15 mouth and gut). *Tlx* expression was detected in an endodermal cell subpopulation surrounding the
16 gonads of newly released medusae (Figure 4C, D) and additionally in the oocytes of female
17 medusae (Figure 4F). This expression is initially localized at the base of the manubrium (not shown)
18 in one day old medusae and expands mid-orally in the endoderm surrounding the gonads (Figure
19 4D). *Tlx* was sporadically detected in isolated ectodermal cells in the manubrium as well. In older
20 medusae, the expression was also detected in the endoderm of the tentacle bulbs (structure on the
21 bell margin proximal to the tentacles) and at the position of newly developing tentacles (Figure 4E).

22 **Discussion**

23 Our detailed analysis of phylogenetic distributions of the medusa stage and the *Tlx* sequence
24 reveals a striking correlation between the presence of an intact *Tlx* and the presence of the medusa
25 life cycle stage in cnidarians. The few occurrences of *Tlx* in non-medusa bearing species are
26 characterized by sequence alteration and/or relaxed selection in conserved regions, suggesting
27 conversion to pseudogenes and possible loss of TLX function in those species.

28 In addition, our RNA-seq analyses show striking upregulation of *Tlx* during medusa development
29 in three disparate medusozoans, suggesting the existence of a conserved role of *Tlx* in the
30 development of the medusa, despite having very distinct developmental trajectories. Specifically,
31 while the hydrozoan medusae relies on lateral budding from the polyp, *Clytia* possesses dedicated
32 polyps (gonozooids) for budding medusae, whereas *Podocoryna* transforms its feeding polyp to a
33 reproductive polyp upon onset of medusa budding. *Aurelia* does not bud medusae but undergoes
34 a process of transverse fission of its polyp, called polydisc strobilation.

35 *PcTlx* expression is localized in cell populations distinct from the germline in areas of active somatic
36 development, as well as mature oocytes in the released medusa. This suggests that *Tlx* plays a
37 role in the maintenance of somatic development in the developing medusa, and an additional role
38 in oocyte maturation or as a maternal effect gene. Progenetic hydrozoans such as *Podocoryna*'s
39 close relative, *Hydractinia*, exhibit a truncation of somatic development and a relatively early
40 maturation of the germline. Evolutionarily, lineages exhibiting progenesis might have undergone
41 relaxed selection on mechanisms controlling somatic development of the medusae, resulting in the
42 loss of *Tlx*.

43 The phylogenetic distribution and expression pattern suggests that *Tlx* is intimately tied to medusa
44 development. Given that the presence of *Tlx* appears to be ubiquitous in bilaterians, *Tlx* was likely
45 present in the last common ancestor of Bilateria and Cnidaria and was secondarily lost multiple
46 times independently in cnidarians, including at the base of Anthozoa, Endocnidozoa and
47 independently several times in Hydrozoa. The ancestral presence of the *Tlx* gene in Cnidaria, in

1 conjunction with the striking correlation of the gene *Tlx* with the presence of the medusa stage in
2 medusozoans, and its apparent role in medusa development, suggests the possibility of an
3 ancestral metagenetic life cycle in Cnidaria. That is, *Tlx* and medusae could have been present in
4 the ancestor of Cnidaria and lost together multiple times in cnidarian evolution, including at the
5 base of Anthozoa, Endocnidozoa and multiple times in Hydrozoa. An alternative explanation is that
6 *Tlx* could have had an unknown ancestral function in the last common ancestor of Cnidaria and
7 Bilateria and was rapidly exapted in medusozoans for medusa development and/or medusa
8 specific structures or cell types. In this scenario, the ancestral function is no longer necessary in
9 anthozoans and other cnidarians that lack *Tlx*. A third explanation is that *Tlx* has an ancestral
10 function in the development of medusa structures in a cnidarian ancestor that exhibits both polypoid
11 and medusoid features. In that scenario, the emergence of a discrete polyp stage in anthozoans,
12 reduction of the ancestral body plan in parasitic endocnidozoans, and the uncoupling of these
13 features into two discrete generations in medusozoan could have been responsible for the pattern
14 of loss and maintenance of *Tlx* respectively. Further investigations of *Tlx* function in medusozoans
15 and in early diverging bilaterians may help clarify the ancestral role of *Tlx* in Cnidaria.

16 **Materials and Methods**

17 **Phylogenetic Methods**

18 Amino acid sequence alignments were carried out with Muscle (Edgar, 2004) using default
19 parameters and manually refined on MEGA7 (Kumar et al., 2016). The EH1 domain, N-terminal
20 arm and homeodomain of TLX were aligned with the homeodomain of NK-L members NK6 and
21 HEX, totaling 82 taxa and 83 characters. Phylogenetic analyses were conducted respectively
22 under Mrbayes 3.2.7 (Ronquist and Huelsenbeck, 2003) and RaxML (Stamatakis, 2014), under
23 the model LG+I+G with 4 discrete gamma categories as selected by Modeltest-NG (based on
24 BIC), on CIPRES portal (Miller et al. 2011). For the Bayesian phylogenetic analysis, 4 runs and 6
25 Markov chains were generated, and the analysis was run for 5 million generations with a 25%
26 burn-in. The posterior probabilities, as well as the final topology come from a majority consensus
27 of the sampled trees. For the Maximum likelihood phylogenetic analysis, support values were
28 evaluated by non-parametric bootstraps (1000 replicates). Bootstrap support values were
29 reported when above 50 and posterior probabilities ranging from 0.90 to 1 are indicated at the
30 node in respect to the color coding (Figure 1b).

31 **Ancestral character state reconstruction of *Tlx* and the gonophore in** 32 **Medusozoa**

33 The MP ancestral character state reconstruction of the presence of *Tlx* in genomic samples and
34 the gonophore were conducted under Mesquite v3.61 (Maddison and Maddison, 2007) on a
35 pruned tree from Cartwright and Nawrocki, 2010, re-rooted with Acraspeda. Dollo parsimony was
36 imposed for *Tlx* ancestral character reconstruction as convergent evolution or horizontal gene
37 transfer seemed unlikely. The reconstruction of the gonophore ancestral state was conducted
38 under Wagner parsimony; characters were unweighted and unordered. Character states were
39 coded as such: no gonophore, sporosacs, cryptomedusoids, eumedusoids, and medusa. In
40 *Ectopleura larynx*, *Tlx* could not be amplified by degenerate PCR but was found in publicly
41 available transcriptomic data, and thus was coded as present (only one case, *Ectopleura larynx*).
42 Within siphonophores three different structures have been proposed to be homologous to the
43 hydromedusa. The eudoxid is a sexual free-living individual, the nectophore is an asexual
44 swimming zooid exhibiting a velum, radial and circular canals, and the gonophore is the sexual
45 zooid that may or not exhibit medusa-like structures depending on lineages (see Dunn et al.,
46 2005). Here the presence of either nectophores or eudoxids was coded as eumedusoid.

47 **Cnidarian genomes and transcriptomes assembly.**

1 Most of the genome and transcriptome sequences were obtained from the NCBI sequence read
2 archive (SRA) (Table S1) and required assembly in order to screen for *Tlx*. Each library was
3 trimmed of low-quality reads and adapters using fastp (Chen et al., 2018). For those transcriptomes
4 from different libraries, filtered reads were combined into a single dataset followed by *de novo*
5 transcriptome assembly using Trinity v2.8.5 (Grabherr et al., 2011). Genome assemblies were
6 carried out using Spades v3.13.1 (Bankevich et al., 2012). Genomes that did not require assembly
7 were obtained from NCBI or from unpublished work shared by collaborators. The source of all the
8 genomes and transcriptomes used in this study can be found in Table S1.

9 ***In silico* search of *Tlx* and phylogenetic analyses.**

10 Potential orthologs of *Tlx* were identified through reciprocal blasts of TLX amino acid sequences
11 (tblastn, e-value cut-off set to 10^{-80} and 10^{-10}) from *Chironex fleckeri*, *Calvadosia cruxmelitensis*,
12 *Aurelia aurita*, *Clytia hemisphaerica*, *Podocoryna carnea*, *Agalma elegans*, and *Craspedacusta*
13 *sowerbii* against cnidarian transcriptomes and genomes. Phylogenetic methods are outlined in the
14 Supplementary Information.

15 **Degenerate PCR screening of *Tlx* from cnidarian genomic DNA.**

16 Degenerate primers in the second exon of *Tlx* (Forward 5'-GGNCAYCCNTAYCAVAGC/MGNGC-
17 3', Reverse 5'-GTKCKHCKRTTYTGAWACCA-3') were designed from 8 medusozoan species:
18 *Chironex fleckeri*, *Calvadosia cruxmelitensis*, *Aurelia aurita*, *Clytia hemisphaerica*, *Podocoryna*
19 *carnea*, *Agalma elegans* and *Craspedacusta sowerbii*. The primers span the N-terminal arm to the
20 C-terminal end of TLX homeodomain, for a total expected amplicon length of 196 bp. Degenerate
21 PCR were carried out using OneTaq 2X Master Mix Standard Buffer, according to manufacturer
22 instructions and with an annealing temperature of 40°C. Amplification products were
23 electrophoresed in a 1.2% agarose gel to assess the presence of the amplicon. The amplifiability
24 of the genomic samples was assessed using 16S degenerate primers. Ten amplicons were
25 selected at random, cloned into the pCR4-TOPO plasmid and sequenced to confirm TLX identity.

26 **Bayesian correlation analysis of the presence of *Tlx* and the medusa stage.**

27 Character coding and ancestral state reconstruction methods are outlined in the Supplementary
28 Information. Correlation analyses between the presence of *Tlx* and medusa stage was performed
29 using BayesTraits v3 (Pagel et al., 2004), imposing irreversibility for *Tlx* by setting transition rate
30 for regains of *Tlx* to zero ($q_{13}=0$, $q_{24}=0$). The presence of *Tlx* and the medusa stage were coded as
31 binary characters. The presence of sporosacs and cryptomedusoids as well as no gonophores
32 were coded as absent and the eumedusoids and fully developed medusae were coded as present.
33 Although eumedusoids do not feed, they have nearly all other medusa components and thus were
34 treated as medusae. The statistical support for the correlation analysis was carried out by
35 computing the marginal likelihood of the two alternative models, independence and dependence of
36 *Tlx* with the medusa stage. The calculation of the Log Bayes Factor was performed and interpreted
37 as recommended by the BayesTraits manual. According to the BayesTraits manual logBF can be
38 interpreted as such: logBF <2 Weak evidence, logBF >2 Positive evidence, $5 \leq \logBF \leq 10$ Strong
39 evidence, logBF >10 Very strong evidence.

40 **Animal care.**

41 *P. carnea* colonies were grown on microscope slides contained in slide racks and kept in artificial
42 seawater (REEF CRYSTALS, Aquarium Systems) in a 7L Kreisel tank at room temperature (~18°C)
43 with a salinity of 29 ppt. Male and female colonies were kept in separate tanks. Colonies were fed
44 two-day old *Artemia* nauplii twice a week and blended mussels once a week. Unfed one and three
45 day old released medusae were collected. Prior to every experiment, *P.carnea* colonies were
46 starved for four days. Animals were relaxed for 30min by addition of menthol crystals (1mg/ml) to
47 the medium and fixed after two medium changes.

48 **Probe synthesis and *In situ* hybridization of *Tlx* in *P.carnea*.**

1 The sequence for *Tlx* transcript was recovered from a newly assembled transcriptome of *P. carnea*.
2 *Tlx* was amplified from medusae cDNA using the following PCR primers: *P. carnea* forward 5'-
3 GAAAGATAAACACGAAAAAGAAACGG-3' and reverse 5'-
4 TCCGGAACCTCATTACTCGCTGTTGC-3' for an expected amplicon length of 528 bp. Amplicons
5 were cloned using the Invitrogen pCR4-TOPO-TA Cloning Kit and sequenced using M13 forward
6 and reverse primers. Sense and antisense DIG labeled riboprobes were synthesized from clones
7 using the Invitrogen T7/T3 Megascript kit. *In situ* hybridization (ISH) protocol was adapted from
8 Gajewsky et al, 1996. Animals were fixed in ice cold fix (3.7%PFA and 0.25% glutaraldehyde in 1X
9 PBS). Hybridization was carried out at 50°C for 18 hours with a probe concentration of 1 ng/μl. DIG
10 labeled riboprobes localization was detected by immunostaining with anti-DIG-Fab-AP (ROCHE)
11 and NBT/BCIP.

12 **De novo transcriptome assembly for *P. carnea*.**

13 For the *Podocoryna carnea* *de novo* transcriptome assembly used in this study, planula and early
14 budding polyp libraries were sequenced and the sequences submitted to NCBI Sequence Read
15 Archive (BioProject ID PRJNA744579) in addition to updated libraries for non-reproductive polyps,
16 budding polyps and medusae (BioProject ID PRJNA245897), and used along with previously
17 generated *P. carnea* libraries described in Sanders and Cartwright, 2015. Stage 2 planulae
18 (elongated, swimming, non-competent) and stage 1-2 early budding polyps, see Frey (1968), were
19 flash frozen and sent to KUMC-GSF for RNA library preparation using the TruSeq RNA Sample
20 Preparation Kit (BoxA). All libraries were 100 bp paired-end with an average insert size of 170 bp.
21 Libraries were then barcoded, pooled, and multiplexed on a single lane of an Illumina NovaSeq
22 6000 S1 flow cell at KUMC-GSF. Low -quality reads were trimmed and adapters using fastp (Chen
23 et al., 2018). Reads from all libraries but the planula ones were mapped to the draft genomes of
24 their respective strains of *Podocoryna carnea* (Chang and Baxevanis, pers. comm.) using STAR
25 (Dobin et al, 2013). Uniquely mapped reads (61.28%) and reads mapping on multiple loci (20.05%)
26 were kept for *de novo* assembly. The *de novo* transcriptome assembly was produced with Trinity
27 v2.8.5 (Grabherr et al., 2011), yielding 472366 transcripts total with an average length of 714.62
28 and a G+C content of 36.76%. Transcripts were blasted against a *Mus musculus* transcriptome
29 dataset (GCA_000001635.9 GRCm39) with an e-value threshold of $1.e^{-100}$, and best hits were
30 removed from the assembly. The longest ORFs from the transcriptomes were predicted using
31 Transdecoder version 5.5.0 (<http://transdecoder.github.io>), duplicate sequences and isoforms were
32 removed by clustering sequences with a 95% identity threshold using CD-HIT version 4.8.1 (Fu et
33 al, 2012) and a BUSCO analysis (Simão et al, 2015) against the metazoan database
34 (metazoa_odb10) was performed to assess the completeness of the transcriptome, estimating a
35 96.2% completeness of the transcriptome (82.9% single copy Buscos and 13.3% duplicated
36 Buscos), 1.5% of fragmented Buscos and 2.3% of missing Buscos on a total of 954 BUSCO
37 markers.

38 **Differential expression analysis in the life cycle of *P. carnea*, *C. hemisphaerica* and *A. coerulea* and *Tlx* qPCR validation in *P. carnea*.**

39 The differential expression analyses were carried out on the transcriptome of *Clytia hemisphaerica*
40 (<http://marimba.obs-vlfr.fr>), *Aurelia coerulea* (<https://davidadlergold.faculty.ucdavis.edu>) and the *de*
41 *nov* transcriptome assembly of *Podocoryna carnea*. The reads quantification was performed at
42 the gene level, quantification combining isoforms, using RSEM (Li and Dewey, 2011). Estimated
43 counts, Fragments per Kilobase Million (FPKM) and Transcripts per Million (TPM) values were
44 generated and differential expression of *Tlx* for the three species was analyzed on these three
45 metrics using Ebsseq (Leng et al. 2013). The differential expression analysis was performed on the
46 main developmental stages of the life cycle for the three species, planulae (binning the three
47 planulae stages for *Clytia*), non-reproductive polyp (respectively, gastrozoid, scyphostoma and
48 non-reproductive polyp), reproductive polyp (respectively, gonozoid, early and late budding
49 polyps, early and late strobila) and medusa (respectively, ephyra, juvenile and mature). The relative
50 expression of *Tlx* was validated in *Podocoryna carnea* through RT-qPCR. Tissues from planulae
51

1 (7 biological replicates) non-reproductive polyps (8 biological replicates), budding polyps (8
2 biological replicates) and released medusae (8 biological replicates) were homogenized in Trizol
3 and incubated for 15 min. Samples were then combined with 0.5 volume of chloroform, mixed,
4 incubated for 3 min and then spun down at 4C for 15min. The supernatant was mixed with one
5 volume of 70% ethanol and RNA extraction was carried out using a Qiagen RNeasy Micro Kit.
6 cDNA synthesis was carried out using Superscript IV and a primer mixture of random hexamers
7 and oligo-dT. cDNA was quantified with Qubit and the qPCR was performed using PowerUP SYBR
8 Green, using the following primers (*Ef1* forward 5'-TTGCCACCTCAACGACCATC-3', *Ef1* reverse
9 5'-TACCGACTGGCACTGTTCCA-3' and *Tlx* forward 5'-CAGAGCCCCACCGAAAAGAA-3', *Tlx*
10 reverse 5'-ATTCCTTGCCACACGCAAT-3').

11 **Acknowledgments**

12 We thank A. Baxevanis and E.S Chang for providing draft genome sequences, A. Klompen for help
13 in data collection and R. Helm and Catriona Munro for providing access to transcriptome
14 sequences. Genome sequencing services was provided by the University of Kansas Medical
15 School Genome Sequencing Center. Computing support was provided by the University of Kansas
16 Center for Research Computing. This work was supported by the National Science Foundation
17 (NSF) Division of Environmental Biology 095357 the University of Kansas General Research Fund,
18 and the University of Kansas Research Excellence Initiative (to PC) and support to PC while serving
19 at the National Science Foundation MT was supported by the Chancellor's Doctoral Fellowship at
20 the University of Kansas and received research support from the KU Graduate Research Fund.

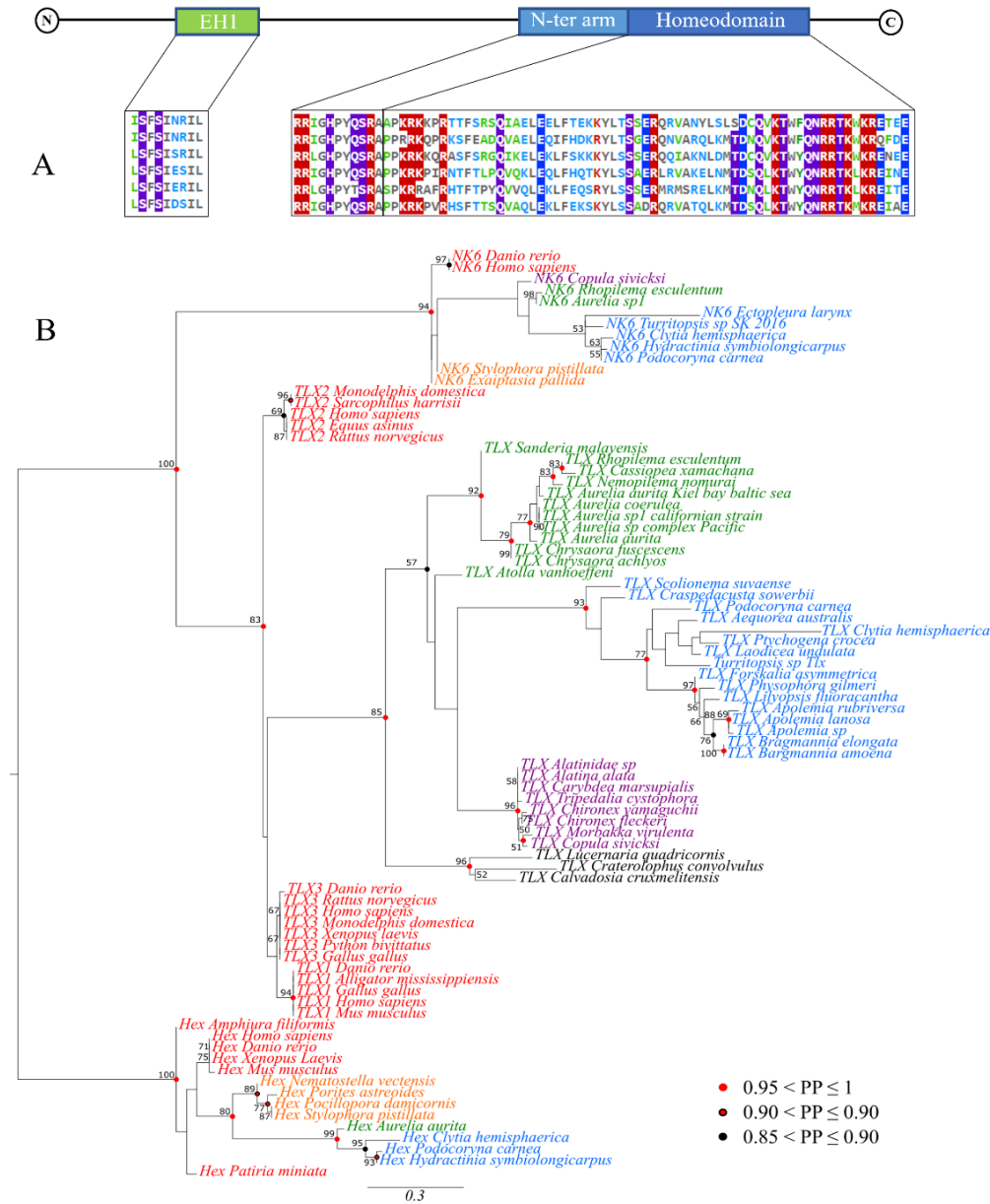
1 References

- 2 1. Allman, G. J. (1871). *A Monograph of the Gymnoblasic Or Tubularian Hydroids*. Ray
3 Society.
- 4 2. Bankevich, A et al. (2012). SPAdes: A New Genome Assembly Algorithm and Its
5 Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 19(5), 455–
6 477.
- 7 3. Boero, F., Bouillon, J., & Piraino, S. (1992). On the origins and evolution of
8 hydromedusan life cycles (Cnidaria, Hydrozoa). In R. Dallai (Ed.) *Sex Origin and*
9 *Evolution., Selected Symposia and Monograph UZI 6. Mucchi*, 59–68.
- 10 4. Boero, F., Bouillon, J., Cicogna, F., & Cornelius, P. (1987). Inconsistent evolution and
11 paedomorphosis among the hydroids and medusae of the Athecatae/Anthomedusae and
12 the Thecatae/Leptomedusae (Cnidaria, Hydrozoa). *Modern trends in the systematics,*
13 *ecology, and evolution of hydroids and hydromedusae*, 229-250.
- 14 5. Bouillon, J., Gravili, C., Gili, J. M., & Boero, F. (2006). An Introduction to Hydrozoa.
15 *Mémoires Du Muséum National d'Histoire Naturelle*, 194, 1–591.
- 16 6. Cartwright, P., & Nawrocki, A. M. (2010). Character Evolution in Hydrozoa (phylum
17 Cnidaria). *Integrative and Comparative Biology*, 50(3), 456–472.
- 18 7. Chang, E.S et al. (2015) Genomic insights into the evolutionary origin of Myxozoa within
19 Cnidaria. *Proc Natl Acad Sci USA* 112(48):14912–14917.
- 20 8. Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ
21 preprocessor. *Bioinformatics*, 34(17), i884–i890.
- 22 9. Cunningham, C. W., & Buss, L. W. (1993). Molecular evidence for multiple episodes of
23 paedomorphosis in the family Hydractiniidae. *Biochemical Systematics and Ecology*,
24 21(1), 57–69.
- 25 10. Derelle, R., & Manuel, M. (2007). Ancient connection between NKL genes and the
26 mesoderm? Insights from Tlx expression in a ctenophore. *Development Genes and*
27 *Evolution*, 217(4), 253–261.
- 28 11. Dobin, A. et al. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1),
29 15–21.
- 30 12. Dunn, C. W. (n.d.). *The colony-level evolution and development of the siphonophora*
31 *(Cnidaria, Hydrozoa)* [Ph.D., Yale University].
- 32 13. Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high
33 throughput. *Nucleic Acids Research*, 32(5), 1792–1797.
- 34 14. Frank, U., Nicotra, M. L., & Schnitzler, C. E. (2020). The colonial cnidarian *Hydractinia*.
35 *EvoDevo*, 11(1), 7.
- 36 15. Frey, J. (1968). Die Entwicklungsleistungen der Medusenknospen und Medusen
37 von *Podocoryne carnea* M. Sars nach Isolation und Dissoziation. *Wilhelm Roux' Archiv fur*
38 *Entwicklungsmechanik der Organismen*, 160(4), 428–464.
- 39 16. Fu, L et al. (2012). CD-HIT: Accelerated for clustering the next-generation sequencing
40 data. *Bioinformatics*, 28(23), 3150–3152.
- 41 17. Gajewski, M., Leitz, T., Schloßherr, J., & Plickert, G. (1996). LWamides from Cnidaria
42 constitute a novel family of neuropeptides with morphogenetic activity. *Roux's Archives of*
43 *Developmental Biology*, 205(5), 232–242.
- 44 18. Galliot, B. (2012). *Hydra*, a fruitful model system for 270 years. *International Journal of*
45 *Developmental Biology*, 56(6-7-8), 411–423.
- 46 19. Goette, A. (1916). Die Gattungen *Podocoryne*, *Stylactis* und *Hydractinia*. *Zoologische*
47 *Jahrbücher*, 39, 443-510.
- 48 20. Gould, S. J. (1985). *Ontogeny and Phylogeny*. Harvard University Press.
- 49 21. Grabherr, M. G et al. (2011). Trinity: Reconstructing a full-length transcriptome without a
50 genome from RNA-Seq data. *Nature Biotechnology*, 29(7), 644–652.
- 51 22. Hincks, T. (1868). *A History of the British Hydroid Zoophytes: 1: Text*. John Van Voorst.
- 52 23. Kamm, K., & Schierwater, B. (2006). Ancient complexity of the non-Hox ANTP gene
53 complement in the anthozoan *Nematostella vectensis*. Implications for the evolution of

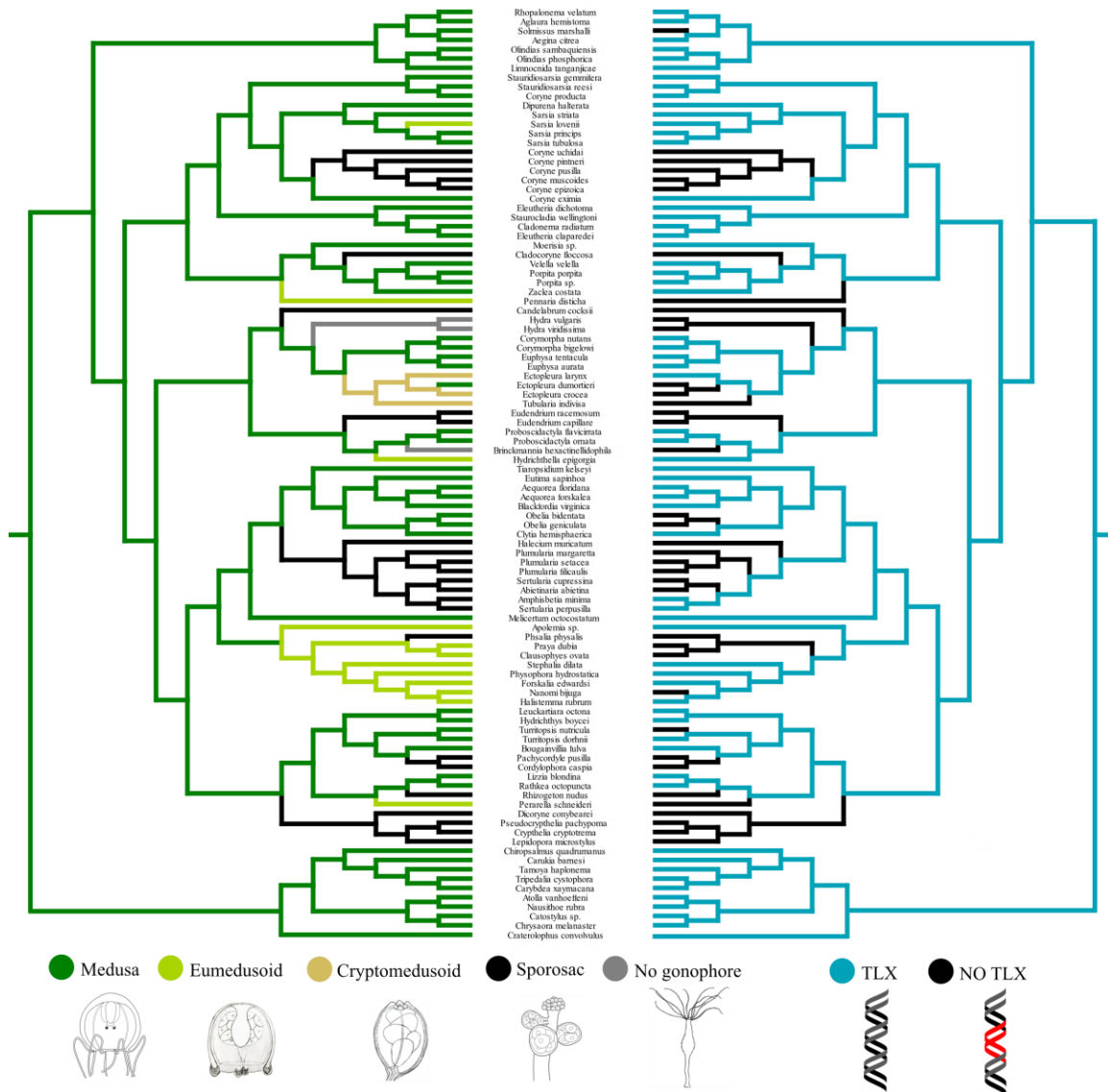
- 1 the ANTP superclass. *Journal of Experimental Zoology Part B: Molecular and*
2 *Developmental Evolution*, 306B(6), 589–596.
- 3 24. Khalturin, K et al. (2019). Medusozoan genomes inform the evolution of the jellyfish body
4 plan. *Nature Ecology & Evolution*, 3(5), 811–822.
- 5 25. Kojima, T., Tsuji, T., & Saigo, K. (2005). A concerted action of a paired-type homeobox
6 gene, *aristaleless*, and a homolog of *Hox11/tlx* homeobox gene, *clawless*, is essential for
7 the distal tip development of the *Drosophila* leg. *Developmental Biology*, 279(2), 434–
8 445.
- 9 26. Kraus, J. E et al. (2015). Adoption of conserved developmental genes in development
10 and origin of the medusa body plan. *EvoDevo*, 6(1), 23.
- 11 27. Kubota, S. (2000). Parallel, paedomorphic evolutionary processes of the bivalve-
12 inhabiting hydrozoans (Leptomedusae, Eirenidae) deduced from the morphology, life
13 cycle and biogeography, with special reference to taxonomic treatment of
14 *Eugymnanthea*. *Scientia Marina*, 64(S1), 241–247.
- 15 28. Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics
16 Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7), 1870–
17 1874.
- 18 29. Leclère, L et al. (2019). The genome of the jellyfish *Clytia hemisphaerica* and the
19 evolution of the cnidarian life-cycle. *Nature Ecology & Evolution*, 3(5), 801–810.
- 20 30. Leng, N et al. (2013). EBSeg: An empirical Bayes hierarchical model for inference in
21 RNA-seq experiments. *Bioinformatics*, 29(8), 1035–1043.
- 22 31. Lenti, E et al. (2016). Transcription factor TLX1 controls retinoic acid signaling to ensure
23 spleen development. *The Journal of Clinical Investigation*, 126(7), 2452–2464.
- 24 32. Li, B., & Dewey, C. N. (2011). RSEM: Accurate transcript quantification from RNA-Seq
25 data with or without a reference genome. *BMC Bioinformatics*, 12(1), 323.
- 26 33. Maddison, W., & Maddison, D. (2007). Mesquite 2. *A Modular System for Evolutionary*
27 *Analysis*. Available online at: <http://www.mesquiteproject.org>.
- 28 34. Masuda-Nakagawa, L. M., Gröer, H., Aerne, B. L., & Schmid, V. (2000). The HOX-like
29 gene *Cnox2-Pc* is expressed at the anterior region in all life cycle stages of the jellyfish
30 *Podocoryne carnea*. *Development Genes and Evolution*, 210(3), 151–156.
- 31 35. Miller, M. A., Pfeiffer, W., & Schwartz, T. (2011, July). The CIPRES science gateway: a
32 community resource for phylogenetic analyses. In *Proceedings of the 2011 TeraGrid*
33 *Conference: extreme digital discovery* (pp. 1-8).
- 34 36. Müller, P., Yanze, N., Schmid, V., & Spring, J. (1999). The Homeobox Gene *Otx* of the
35 Jellyfish *Podocoryne carnea*: Role of a Head Gene in Striated Muscle and Evolution.
36 *Developmental Biology*, 216(2), 582–594.
- 37 37. Murrell, B et al. (2013). FUBAR: A Fast, Unconstrained Bayesian AppRoximation for
38 Inferring Selection. *Molecular Biology and Evolution*, 30(5), 1196–1205.
- 39 38. Namikawa H. (1991). A New Species of the Genus *Stylactaria* (Cnidaria, Hydrozoa) from
40 Hokkaido, Japan. *Zoological Science*, 8(4), p805-812.
- 41 39. Pagel, M., Meade, A., & Barker, D. (2004). Bayesian estimation of ancestral character
42 states on phylogenies. *Systematic biology*, 53(5), 673-684.
- 43 40. Pang, K., & Martindale, M. Q. (2008). Developmental expression of homeobox genes in
44 the ctenophore *Mnemiopsis leidyi*. *Development Genes and Evolution*, 218(6), 307–319.
- 45 41. Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference
46 under mixed models. *Bioinformatics*, 19(12), 1572–1574.
- 47 42. Ryan, J. F et al. (2006). The cnidarian-bilaterian ancestor possessed at least 56
48 homeoboxes: Evidence from the starlet sea anemone, *Nematostella vectensis*. *Genome*
49 *Biology*, 7(7), R64.
- 50 43. Sanders, S. M., & Cartwright, P. (2015a). Patterns of Wnt signaling in the life cycle of
51 *Podocoryna carnea* and its implications for medusae evolution in Hydrozoa (Cnidaria).
52 *Evolution & Development*, 17(6), 325–336.

- 1 44. Sanders, S. M., & Cartwright, P. (2015b). Interspecific Differential Expression Analysis of
2 RNA-Seq Data Yields Insight into Life Cycle Variation in Hydractiniid Hydrozoans.
3 *Genome Biology and Evolution*, 7(8), 2417–2431.
- 4 45. Simão, F et al. (2015). BUSCO: Assessing genome assembly and annotation
5 completeness with single-copy orthologs. *Bioinformatics*, 31(19), 3210–3212.
- 6 46. Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-
7 analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313.

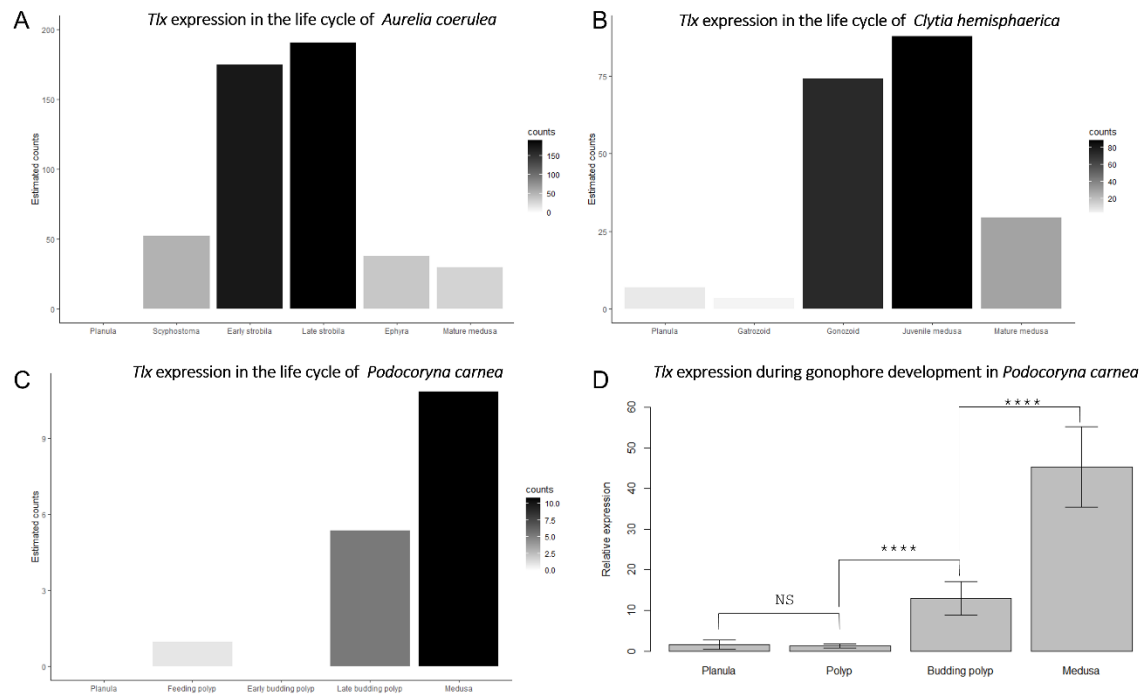
Figures and Tables



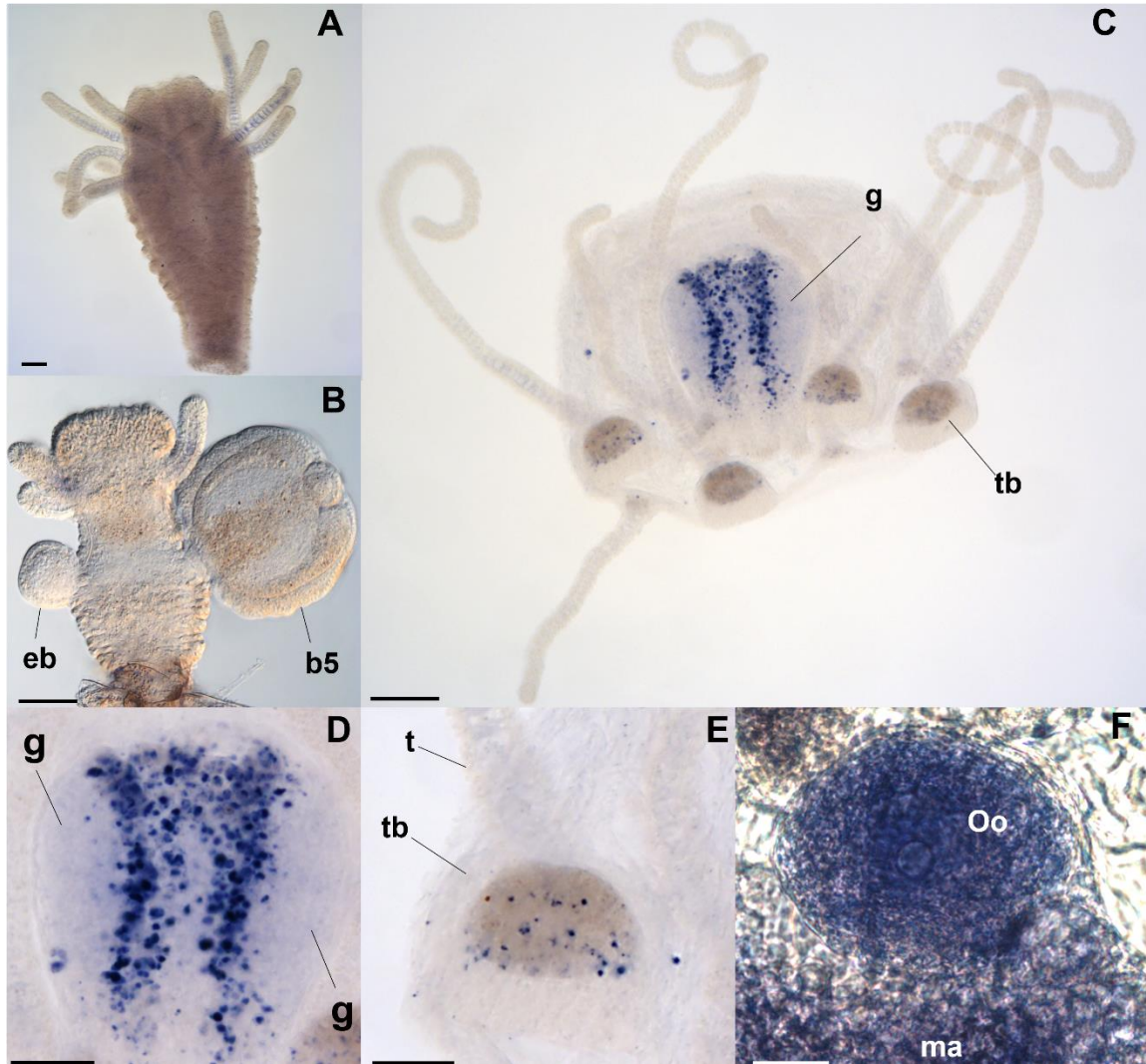
1 **Figure 1.** A) Schematic of TLX genomic structure and the corresponding amino acids sequence
 2 alignment for TLX conserved domains from six representative medusozoan taxa. Highly conserved
 3 positions are highlighted (>70% identity). Colors represent features of the position, purple (polar
 4 uncharged), red (positively charged), blue (negatively charged) and green (hydrophobic) B)
 5 Phylogram from maximum likelihood analysis, with three NK-L representatives (*Tlx*, *Nk6* and *Hex*).
 6 Vertebrate sequences are in red, hydrozoan sequences in blue, staurozoan sequences in black,
 7 scyphozoan sequences in green, cubozoan sequences in purple and anthozoan sequences in
 8 orange. Bootstrap values greater than 50% are indicated (1000 bootstraps) next to the nodes.
 9 Bayesian posterior probability greater than or equal to 85% are reported on the nodes with colored
 10 circles (color code on the figure). *Hex* sequences are used as the outgroup. Scale bar = number of
 11 inferred substitutions per position in the alignment.



1 **Figure 2.** Wagner maximum parsimony ancestral character state reconstructions of medusozoan
 2 reproductive systems (left) against the Dollo maximum parsimony ancestral character state
 3 reconstructions of the presence of *Tlx* (right). Characters are unweighted and unordered, no
 4 optimization model was applied. Branches are colored in terms of degree of medusa reduction and
 5 presence of *Tlx* (shown in legend). Phylogeny pruned from Cartwright and Nawrocki, 2010 and
 6 rooted with Acraspeda.



1 **Figure 3.** Normalized expression of *Tlx* in corrected counts for life cycle developmental stages (A-
2 C). A) *Aurelia coerulea*, B) *Clytia hemisphaerica*, C) *Podocoryna carnea*. The gradient charts
3 indicate the breadth of *Tlx* corrected counts values for each species. D) RT-qPCR in *Podocoryna*
4 *carnea*. (****) indicates a two tailed p-value <0.0001 from t-tests, (NS) indicates a non-significant
5 two tailed p-value (p-value = 0.4638) from unpaired t-test.



1 **Figure 4.** *In situ* localization of *PcTlx* transcripts on *Podocoryna carnea*. A) in male non-
2 reproductive polyp, B) in a male budding polyp, C) in a male medusa, D) higher magnification of
3 (A) of the manubrium, E) higher magnification of (A) of a tentacle bulb, F) and a mature oocyte in
4 a female medusa, showing *Tlx* expression,. Samples are presented in radial view except the
5 tentacle bulb that is presented in oral view. The manubrium (D) is presented oral down, aboral up.
6 The tentacle bulb (G) is presented proximal left, distal right. Abbreviations: b5, medusa bud stage
7 5; eb, early bud (stage 1); g, gonad; t, tentacle; tb, tentacle bulb. Scale bar: 200 μ m (A-C), 100 μ m
8 (D and E), 20 μ m (F).

Table 1. *Tlx* presence in publicly available cnidarian genomes and transcriptomes

	<i>TLX in genomes</i>	<i>TLX in transcriptome</i>
<i>Anthozoa</i>	0 out of 31	0 out of 97
<i>Hexacorallia</i>	0 out of 29	0 out of 75
<i>Octocorallia</i>	0 out of 2	0 out of 20
<i>Ceriantharia</i>	N/A	0 out of 6
<i>Endocnidozoa</i>	0 out of 6	0 out of 4
<i>Hydroidolina without medusaeⁱ</i>	0 out of 6	1 out of 12
<i>'Anthoathecata'</i>	0 out of 6	0 out of 11
<i>Leptothecata</i>	N/A	1 out of 1
<i>Siphonophorae</i>	N/A	0 out of 2
<i>Hydroidolina with medusae¹</i>	4 out of 4	29 out of 51
<i>Siphonophorae</i>	N/A	14 out of 30
<i>Leptothecata</i>	1 out of 1	10 out of 10
<i>'Anthoathecata'</i>	1 out of 1	6 out of 7
<i>Trachylinae</i>	2 out of 2	1 out of 4
<i>Trachymedusae</i>	N/A	0 out of 2
<i>Narcomedusae</i>	N/A	0 out of 1
<i>Limnomedusae</i>	2/2	1 out of 1
<i>Scyphozoa</i>	10 out of 10	8 out of 10
<i>Discomedusa</i>	10 out of 10	7 out of 8
<i>Coronata</i>	N/A	1 out of 2
<i>Staurozoa</i>	2 out of 2	4 out of 5
<i>Cubozoa</i>	3 out of 3	5 out of 5

¹ Medusae defined as any medusa-like structure (nectophore, medusoid, medusa)

Table 2. Analyses of the intensity of the selection on the TLX homeodomain within medusozoans

Reference branches Test branches	Medusozoans with medusae		Hydrozoans with medusae	
Hydrozoans without medusae	K=0.15, p<0.0001	Acraspida	K=1.66, p<0.0001	
<i>Amphisbetia minima</i>	K=0.00, p=0.0010	Trachylina	K=0.78, p=0.413	
<i>Dynamena pumila</i>	K=0.16, p<0.0001	Siphonophorae	K=2.30, p=0.0050	
<i>Ectopleura larynx</i>	K=1.19, p=0.7710	Leptothecata	K=1.26, p=0.521	
<i>Millepora squarrosa</i>	K=0.66, p=0.0249	Anthoathecata	K=0.72, p=0.195	
Amplicons from degenerate PCR	K=0.71, p=0.1715	Amplicons from degenerate PCR	K=1.04, p=0.917	