

In the format provided by the authors and unedited.

# The genome of the jellyfish *Clytia hemisphaerica* and the evolution of the cnidarian life-cycle

Lucas Leclère<sup>1</sup>, Coralie Horin<sup>1</sup>, Sandra Chevalier<sup>1</sup>, Pascal Lapébie<sup>1,7</sup>, Philippe Dru<sup>1</sup>, Sophie Peron<sup>1</sup>, Muriel Jager<sup>2,3</sup>, Thomas Condamine<sup>2</sup>, Karen Pottin<sup>2,8</sup>, Séverine Romano<sup>1</sup>, Julia Steger<sup>1,4</sup>, Chiara Sinigaglia<sup>1,9</sup>, Carine Barreau<sup>1</sup>, Gonzalo Quiroga Artigas<sup>1,10</sup>, Antonella Ruggiero<sup>1,11</sup>, Cécile Fourrage<sup>1,12</sup>, Johanna E. M. Kraus<sup>4,13</sup>, Julie Poulain<sup>5</sup>, Jean-Marc Aury<sup>6</sup>, Patrick Wincker<sup>5</sup>, Eric Quéinnec<sup>2,3</sup>, Ulrich Technau<sup>4</sup>, Michaël Manuel<sup>2,3</sup>, Tsuyoshi Momose<sup>1</sup>, Evelyn Houliston<sup>1</sup> and Richard R. Copley<sup>1\*</sup>

<sup>1</sup>Laboratoire de Biologie du Développement de Villefranche-sur-mer, Sorbonne Université, CNRS, Villefranche-sur-mer, France. <sup>2</sup>Evolution Paris-Seine, Institut de Biologie Paris-Seine, Sorbonne Université, CNRS, Paris, France. <sup>3</sup>Institut de Systématique, Evolution, Biodiversité (ISYEB UMR 7205), Sorbonne Université, MNHN, CNRS, EPHE, Paris, France. <sup>4</sup>Department for Molecular Evolution and Development, Centre of Organismal Systems Biology, University of Vienna, Vienna, Austria. <sup>5</sup>Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, Evry, France. <sup>6</sup>Genoscope, Institut de Biologie François-Jacob, Commissariat à l'Energie Atomique, Université Paris-Saclay, Evry, France. <sup>7</sup>Present address: Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université, Marseille, France. <sup>8</sup>Present address: Laboratoire de Biologie du Développement (IBPS-LBD, UMR7622), Sorbonne Université, CNRS, Institut de Biologie Paris Seine, Paris, France. <sup>9</sup>Present address: Institut de Génomique Fonctionnelle de Lyon, École Normale Supérieure de Lyon, CNRS UMR 5242-INRA USC 1370, Lyon cedex 07, France. <sup>10</sup>Present address: The Whitney Laboratory for Marine Bioscience, University of Florida, St. Augustine, FL, USA. <sup>11</sup>Present address: Centre de Recherche de Biologie cellulaire de Montpellier, CNRS UMR 5237, Université de Montpellier, Montpellier Cedex 5, France. <sup>12</sup>Present address: Service de Génétique UMR 781, Hôpital Necker—APHP, Paris, France. <sup>13</sup>Present address: Sars International Centre for Marine Molecular Biology, University of Bergen, Bergen, Norway. \*e-mail: [copley@obs-vlfr.fr](mailto:copley@obs-vlfr.fr)

## **SUPPLEMENTARY INFORMATION**

### **The genome of the jellyfish *Clytia hemisphaerica* and the evolution of the cnidarian life-cycle**

|                                      |           |
|--------------------------------------|-----------|
| <b>Supplementary Methods .....</b>   | <b>2</b>  |
| <b>Supplementary Section 2 .....</b> | <b>4</b>  |
| <b>Supplementary Figures .....</b>   | <b>6</b>  |
| <b>Supplementary Tables.....</b>     | <b>20</b> |

## SUPPLEMENTARY METHODS

### Genome sequencing

#### Illumina Overlapping Paired-end library: AGLDOSE

30 ng of DNA was sonicated to a 100- to 800-bp size range using the E210 Covaris instrument (Covaris, Inc., USA). Fragments were end-repaired, then 3'-adenylated, and Illumina adapters were added by using NEBNext Sample Reagent Set (New England Biolabs). Ligation products were purified by Ampure XP (Beckmann Coulter) and DNA fragments (>200 pb) were PCR-amplified using Illumina adapter-specific primers and Platinum Pfx DNA polymerase (Invitrogen). Amplified library fragments were size selected on 3% agarose gel around 300 bp. After library profile analysis by Agilent 2100 Bioanalyzer (Agilent Technologies, USA) and qPCR quantification (MxPro, Agilent Technologies, USA), library was sequenced using 101 base-length read chemistry in a paired-end flow cell on the Illumina sequencer (Illumina, USA) in order to obtain overlapping reads and generate long reads of 180 bp.

#### Illumina Paired-end library: AGLDOSW

100 ng of DNA was sonicated to a 100- to 800-bp size range using the E210 Covaris instrument (Covaris, Inc., USA). Fragments were end-repaired, then 3'-adenylated, and Illumina adapters were added by using NEBNext Sample Reagent Set (New England Biolabs). Ligation products were purified by Ampure XP (Beckmann Coulter) and DNA fragments (>200 pb) were PCR-amplified using Illumina adapter-specific primers and Platinum Pfx DNA polymerase (Invitrogen). Amplified library fragments were size selected on 3% agarose gel around 500 bp. After library profile analysis by Agilent 2100 Bioanalyzer (Agilent Technologies, USA) and qPCR quantification (MxPro, Agilent Technologies, USA), library was sequenced using 101 base-length read chemistry in a paired-end flow cell on the Illumina sequencer (Illumina, USA)).

#### Illumina 3-5 kb Mate-Pair library: AGLDOSN

Mate pair library was prepared following Nextera protocol (Nextera Mate Pair sample preparation kit, Illumina). Briefly, genomic DNA was simultaneously enzymatically fragmented and tagged with a biotinylated adaptor. Fragments were size selected (3-5 Kb) through regular gel electrophoresis, and circularized overnight with a ligase. Linear, non-circularized fragments were digested and circularized. DNA was fragmented to 300-1000-bp size range using Covaris E210. Biotinylated DNA was immobilized on streptavidin beads, end-repaired, then 3'-adenylated, and Illumina adapters were added. DNA fragments were PCR-amplified using Illumina adapter-specific primers and then purified. Finally, library was quantified by qPCR and evaluated using an Agilent 2100 bioanalyzer (Agilent Technologies, USA). Library was sequenced using 100 base-length read chemistry on a paired-end flow cell on the Illumina sequencer (Illumina, USA).

#### 454 8 kb Mate-pairs library: AGLAOTA

15 µg of DNA were sheared to about 8kb, end repaired by using the END it Repair kit (Epicentre) and ligated to biotinylated loxP adaptors (Roche). After gel size selection of 8 kb bands and fill in, 300 ng DNA were circularized by Cre recombinase and remaining linear DNA digested by Plasmid Safe ATP

dependent DNase (Epicentre) and exonuclease I. Circular DNA was fragmented by nebulization and biotinylated fragments were immobilized on streptavidin beads. Ligation and then library was prepared according to 454 protocol without further modifications. Library was quantified and evaluated using a 2100 bioanalyzer (RNA 6000 PicoLabchip kit, Agilent Technologies, USA). Library was sequenced using Pico Titer Plate on 454 GSFlx instrument with Titanium chemistry (Roche Diagnostic, USA).

#### 454 20 kb Mate-pairs library: AGLDOTA

30 µg of DNA were sheared to about 20kb, end repaired by using the END it Repair kit (Epicentre) and ligated to). After gel size selection of 20 kb bands and fill in, 300 ng DNA were circularized by Cre recombinase and remaining linear DNA digested by Plasmid Safe ATP dependent DNase(Epicentre) and exonuclease I. Circular DNA was fragmented by nebulization and biotinylated fragments were immobilized on streptavidin beads. Ligation and then library was prepared according to 454 protocol without further modifications. Library was quantified and evaluated using a 2100 bioanalyzer (RNA 6000 PicoLabchip kit, Agilent Technologies, USA). Library was sequenced using Pico Titer Plate on 454 GSFlx instrument with Titanium chemistry (Roche Diagnostic, USA).

#### 454 40 kb Mate-pairs library: AGLFOTA

30 µg of DNA were sheared to about 40kb, end repaired by using the END it Repair kit (Epicentre) and ligated to). After gel size selection of 40 kb bands and fill in, 300 ng DNA were circularized by Cre recombinase and remaining linear DNA digested by Plasmid Safe ATP dependent DNase (Epicentre) and exonuclease I. Circular DNA was fragmented by nebulization and biotinylated fragments were immobilized on streptavidin beads. Ligation and then library was prepared according to 454 protocol without further modifications. Library was quantified and evaluated using a 2100 bioanalyzer (RNA 6000 PicoLabchip kit, Agilent Technologies, USA). Library was sequenced using Pico Titer Plate on 454 GSFlx instrument with Titanium chemistry (Roche Diagnostic, USA).

## SUPPLEMENTARY SECTION 2

### The Marimba database.

*Clytia hemisphaerica* sequence and expression data are accessible from MARIMBA (<http://marimba.obs-vlfr.fr/organism/Clytia/hemisphaerica> - Marine Invertebrate Models Database). This will be described more fully elsewhere. Data are queryable by gene or transcript identifiers (Menu Tools->Genome & Gene), browsable, or searchable via sequence similarity. Genes can be searched by identifier or human ortholog names (when available). Gene pages contain links to genomic sequence, predicted transcript and protein sequences, as well as expression profiles. Protein pages include Interpro database and GO term annotation.

### *Ectopleura larynx* orthologs of *Clytia* medusa specific TFs.

from SRA accessions SRR923510\_1, SRR923510\_2 (see methods and Supplementary Tab. 5).

```
>E_lar_DN34629_c0_g1 Nitrogen Regulatory Protein
MPKSI IKDSGDSQEKNKSKKFRHKICLHTHGLVSKFRRLRLFNKEDDFKRKLSWRSDWEA
SSSVYDEDSRSSFGDFKSDLSQVSSDNESLTDHCSWEVCSDSGFCEDSDCFSQERDNDKE
VLFTGIVKTASQMKPPVKRKKFTPLSKSTPKAQRKVRSPKCSSLVKEKLKSGSTGKSSKS
VMFKTKKITDYYQKTTT VKTIKSSIFPRHTSAITLSKPPQSLSQSLNVTTPRPSKLSPRKS
LTEPRKKRKLEDSPIRKCREVSKPSVKKSRSSELTIASLLDDNLSNFSFFDTIFRSRENT
LDLGVSPIQQKKECVSCGTTRTPMWRDADENTPLCNACGIRWKYRVRCTRCWYVPMKNE
RVPYCSNCGHFQTLKKVNSMRRYKSCIY*

>E_lar_DN34214_c0_g1 CREB-binding protein
MRKSYNLQRKKGSSFAYGLYDGDDEEFLPKIKNSSAGGRKKSCAKKELETTT KTLET
HHEILTRPAVEIAAVEPENFLTDSEALNHALDCIFTEAKCDIKKCYTAKEFLFKGMIHAN
RCEVANTDLEHNQKFLFCKSMKAILLHTSSCKNEDCSTHLFVLKLCQFHQETCKTADCKS
ELCMNQPRNRQEIDEIEEDPPKQQPFQPASSFDEAQQQYYPSSLTSIPPAYPANDWSHLY
QIPPPPPFNTTYQWQPFNTPLPSNFWLAKPNEKPLYDEDTTTSQVKDKTSLLEAPSEV
FKDNHLQTTTDES LFRPVADDQYFRPQPEVDVICGSQWANINFMSREIPHQADTETADE
SDSDGVKEDVVYDERPRITPVTPPIENYNHFNAAC*

>E_lar_DN33986_c0_g1 heat shock factor
MANYPGIPNFM LTHNFNIVQPFTFVSQDLPHFLPVSKVVYECTDLNTVDEDDYDGKLFPKR
LRGLLNEKSLNHIIRWSE DGKSLVIPCENAFYTSILHRRKHKKFFKTESLRSFYRQLNLYG
FNKIKASQLES DANETIDTAFHHPYFLRDKPELIDYVKRRIKAKKGGDEKSFQRLKKCNM
SPKPYMETLAGTEGRRSKFLRAGKRRGSGGTICELYEGQSFYRLYHLLAQNSPDICWSE
NGQFVEIPNHLSFYRNVLLRKSGKIFKTESLKSFVRQLNLYGFKKIHPEKEEYHVEMAKK
DEDTYSPESVFRHVSFMRDRDDLHLVKRRCNVKNSPTHYKTQRFSLQPNDEQLMQIYQS
TSTIHSPTNAVTKSLVHLPSPLLKIPPSIPNSPVLPPQRQEETSFPMATYQNCSTPQQQ
QYIlyEPLRTEQQYVSLVQPLLPTPALPPNKELVTGSSHETVNIVVEEEKSYASPEPVCM
DERQHSLS EDDSYNECHRLFLDLKIEAFAIPSFSFAS*

>E_lar_DN39961_c0_g1 TCF15/Paraxis
MFRSTDYHRFYASSTSSSSSRENT HLLDSLRSKDAKNTHYLNKPLTHTTLLYPHDYSYA
KAGKKT FHSSTKKLKGKERARSHCVNDGFSALRDLIPTDPNRLSKSKIETLRLATSYICH
LKSLVEKSAPVDSQYEELYYATCSSGGAENICTFCVTF LRSTSYGK*

>E_lar_DN32686_c0_g1 PDX/XLOX
MMEGSYGNNYC WQKKLSHPPVRYEPTASQHTPLTQCRYQQGGVKKSLFPPLDVGEVID
TLCNMSNVPGSEEKNEKKTNKESPTNDTKEERVPKSWQSAKVKKRTRTTYTRAQQVELEK
EYLFSRYISRARRIELAKTLALTEKHIKIWYQNRMMKEKRDEEEALRGENNAFCDIAS*

>E_lar_DN34776_c0_g1 DRGX
INLTEARVQVWFQNRRAKWRKTSNKNKSTSGACSPDPPPV CSTSEKPPAKPQPAKTKETK
SDEKKA EVSSLVEETGASYSSPKLTHPPYLFPNYQSYHSSYSGGLEPRAYPMGGKEHY
TRNHDNVE*
```

```

>E_lar_DN32752_c0_g1 TLX
MGKEFSFSIERILAQEDFRKERDDVIIKSDIVLPDERFSLEKEYSTCEEIKCSCHETLLE
SQLTPQWGPTARRIGHYPYQNRAPAKRTKNAQKIFTRAQVFQLEVLHFHEKKYLSSTERGKV
AANLRMTDSQLKTWYQNRRTKMKREFNEKLERFRKQYRANEMNLSSEKMFH*

>E_lar_DN30393_c0_g1 Six1/2
MERNNNYLDLCNYHPHTAAVSSHQQQHHLSFVSYTPEQVSCICEVLEQSGDIEHLAKFLW
SLPSYDNVYTNEAVVAKCVVAFDQGNLPELYRMIESNHFSPPGYHEKLQMLWLRSRMEV
EKVRGRPLGAVGKYRVRKYPLPRTIWDGEETSYCFKEKSRAILKEWYATNPYPSPREKK
ELAEGTGLSTTQVRNWFKNRRQRDRTAEYKEREELDKEKLVYRHHHQHASQEEDWGKEA
NIASTSYQFAKQHFSNAGDHLFNKAFQEOPYRELTHKQAYNVMTNI*

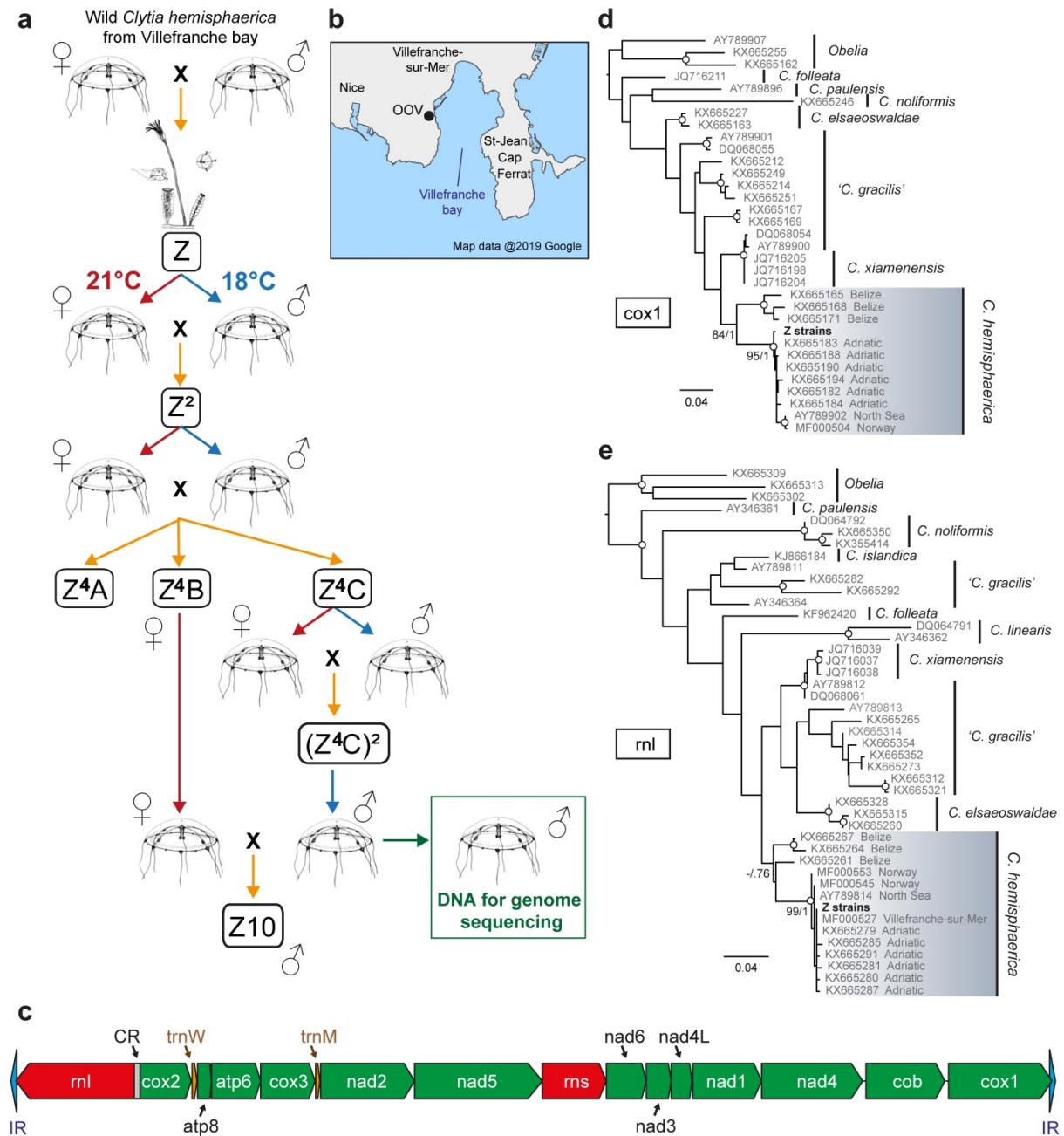
>E_lar_DN29073_c0_g1 FoxL2
MTVMTLLDQLTNQTSLSGHVTMKSECSEIVEGFEEPTYLCTTTATRGDDVDNDDNNNTE
DGCKTPEKSATNDKDPNLKPPYSYVALIAMAIKQTPDKKLTNLQIYQFIVEKFPFYKKNK
KGWQNSIRHNLSLNECFIKVPREGGERKGNWTLDPSCCEYMFDD

>E_lar_DN23123_c0_g1 CnoxA
RPSPPLLLMKRDQGVCRKVPPMIQSNCTKDCCKPRYGLSPPPAIYRTPYFYPNFYFGFDY
DCEYMKRFRFTSYNSTQLARLEKAFGENKYLTRKKRSELAIQLRLSEKQVKIWFQNKRMKW
KKQAQEEENNEDEHVEIMC*

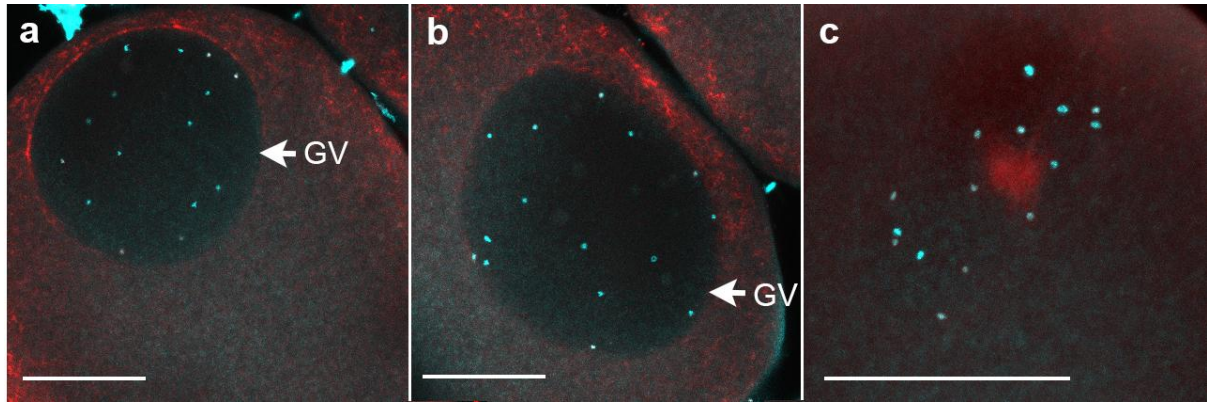
>E_lar_DN37286_c1_g1 CDX
MRRDIFRPASLLAEPETFDNERFNAMVDSALTMTFNWDPSEKIIIFPEKTKDEFDEAQRA
LNCDDALKTEPPSIKSLQQDQNALIQRIYDRTLGLRGMPSLHQPFQQQIGTQHQQQQQ
QQQQRFPSSLELNIQPTLMGQQMDMFNRPLVKNQQQTNNVDWVEYAQDPAMRSRPFSSQ
QTCELEREFRTLQYITRRRRRIELAYNLNLSEKQIKTWQNRVRVKERKKRKNNGDNVKEP
FENSEGIV*

```

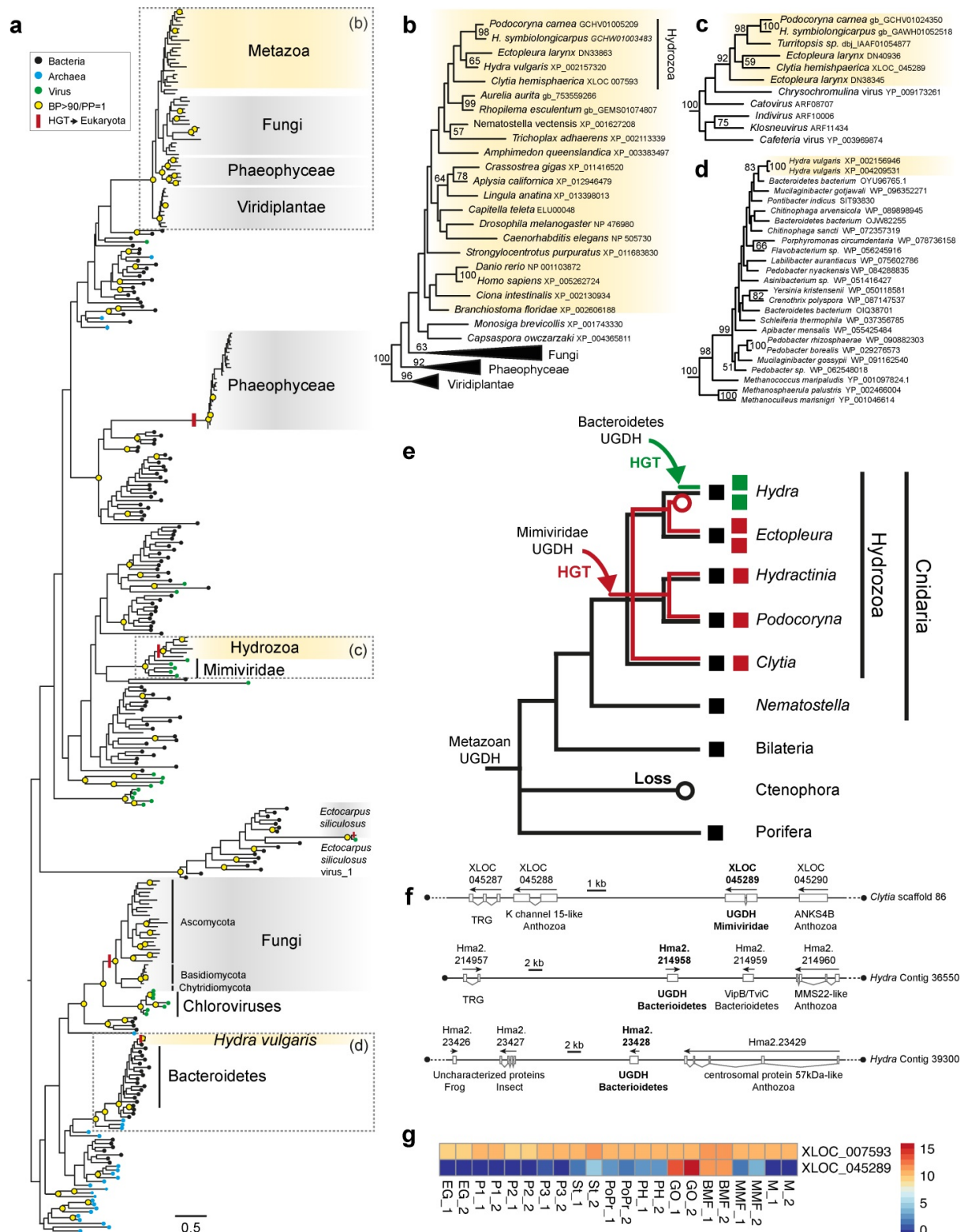
## SUPPLEMENTARY FIGURES



**Supplementary figure 1. Sequenced strains information.** (a) Relationships between *Clytia hemisphaerica* laboratory strains; An original colony Z was derived by metamorphosis of a founder planula larva, with subsequent self-crossing between male and female medusae budded from this colony. See methods for details. (b) The medusae used to generate the founder planula were collected in the bay of Villefranche-sur-Mer. (c) Organization of the linear mitochondrial genome of the sequenced *Clytia Z* strain showed the same organization as that inferred for the Hydroidolina common ancestor. (d,e) Maximum likelihood analyses (RaxML, GTR+G) using two mitochondrial phylogenetic makers (d: *cox1*/COI, e: *rnl*/16S RNA) unambiguously assign the Z strains to the species *Clytia hemisphaerica*. White circles indicate highly supported nodes with >90% bootstrap values (RaxML, 500 replicates) and a Bayesian posterior probability of 1 (MrBayes, model GTR+G). Drawings in (a) adapted from <sup>87</sup> with permission, in (b) adapted from Google Maps.

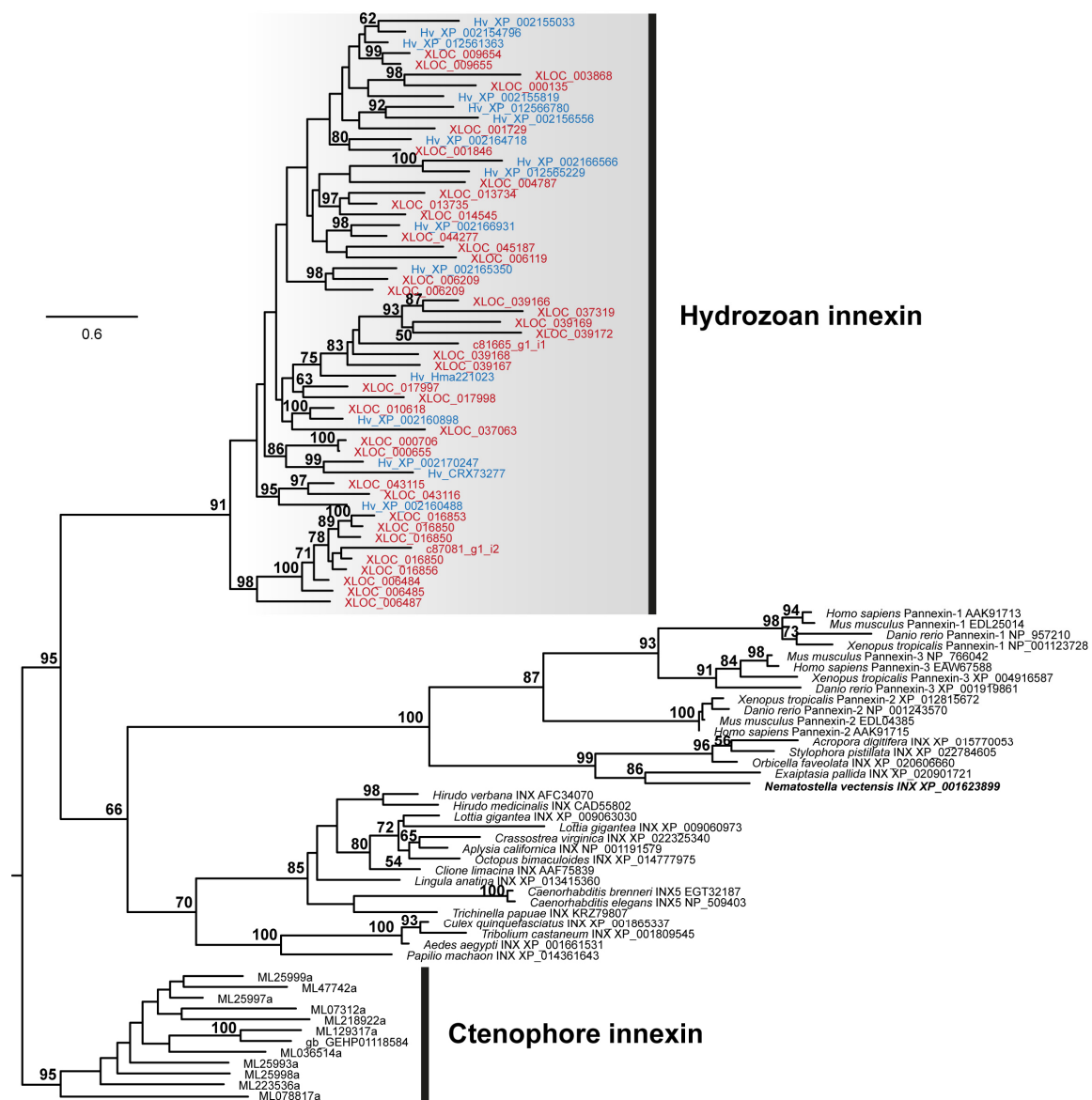


**Supplementary figure 2. *Clytia hemisphaerica* has 15 pairs of chromosomes.** Hoechst staining (blue) of DNA showing the paired chromosomes of *Clytia* fully grown oocytes dissected from isolated gonads. Anti-tubulin immunofluorescence is shown in red. In *Clytia*, the chromosomes in oocytes arrested in meiotic prophase I before maturation (a and b) are already strongly condensed and each pair of replicated chromosomes is clearly distinct in the intact oocyte nucleus (GV, Germinal Vesicle). The oocyte in c), fixed in cold methanol, has initiated meiotic maturation and the chromosome pairs are gathering on a microtubule aster. In each of 8 oocytes in which the Hoechst-stained chromosome packets were clearly distinct, 15 were counted. Additional small Hoechst-stained fragments were observed in some cases. Maximum projections of confocal z stacks covering the whole GV or chromosome cluster are shown. Scale bar 50 $\mu$ m.



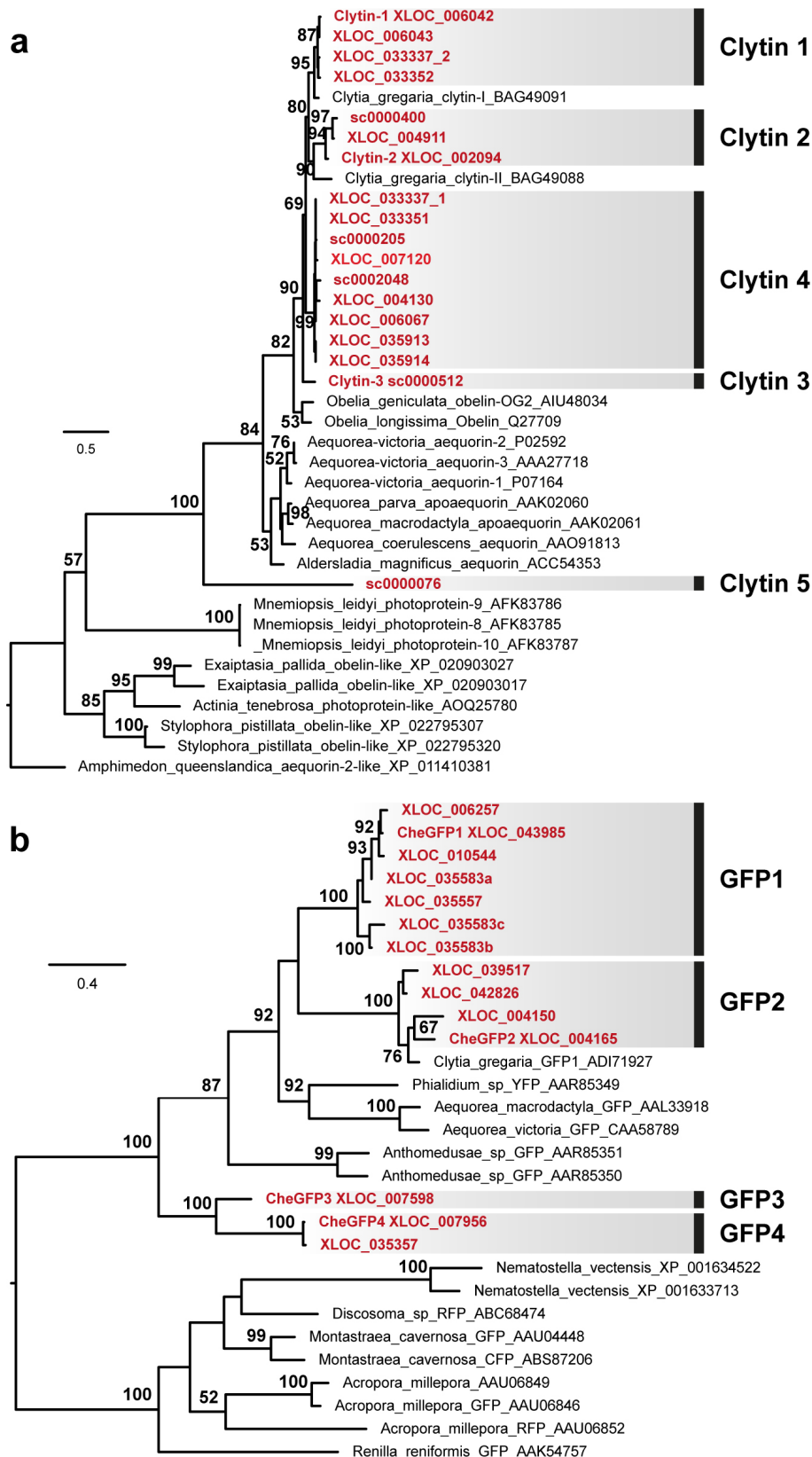
**Supplementary figure 3. UGDH-like phylogeny and evolution.** (a) Maximum likelihood phylogenetic reconstruction (RaxML, LG+G) of the two UGDH-like (UDP-glucose 6-dehydrogenase) genes found in the *Clytia* genome. Yellow circles indicate highly supported nodes with >90% bootstrap values (BP RaxML, 500 replicates) and a Bayesian posterior probability of 1 (MrBayes, model GRT+G). (b) Metazoa UGDH subtree including one *Clytia* ortholog. (c) Subtree showing that the second *Clytia* UGDH, together with several other hydrozoan sequences, are nested within UGDH genes from mimiviridae, strongly arguing for a Horizontal Gene Transfer (HGT), to a hydrozoan

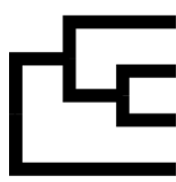
ancestor, from a giant virus of the Mimiviridae family. (d) Subtree showing that two UGDH-like genes of *Hydra* are nested within a strongly supported Bacteroidetes clade, strongly suggesting that they were acquired in the *Hydra* lineage by HGT from bacteria. In (b,c,d) values on nodes correspond to BP>50%. (e) Evolutionary scenario for the evolution of the UGDHs found in the hydrozoan lineage. (f) genomic location of *Clytia* and *Hydra* UGDH genes inherited by HGT. Presence of clear metazoan orthologs in these scaffolds argues against bacterial/viral contamination. Gene name and taxon of the BLASTp hit are indicated below the genes; TRG: Taxon Restricted Gene - no BLASTp hit. (g) expression level (rlog values) of the two *Clytia* UGDH genes across the life cycle (see legend of Fig. 1 for details on stage abbreviations). Note the ubiquitous expression of the “metazoan” UGDH gene (XLOC\_007593) and the Gonozooid (GO) and Baby Medusae (BMF) specific expression of the “Mimiviridae-related” UGDH (XLOC\_045289).



**Supplementary figure 4. Innexin Maximum Likelihood phylogeny.** Model LG+G RAxML.

Rooted with ctenophore innexin. Red: *Clytia* sequences, blue: *Hydra* sequences, black bold: *Nematostella* sequence. Values on nodes = rapid RAxML Bootstrap values (500 replicates) > 50%.

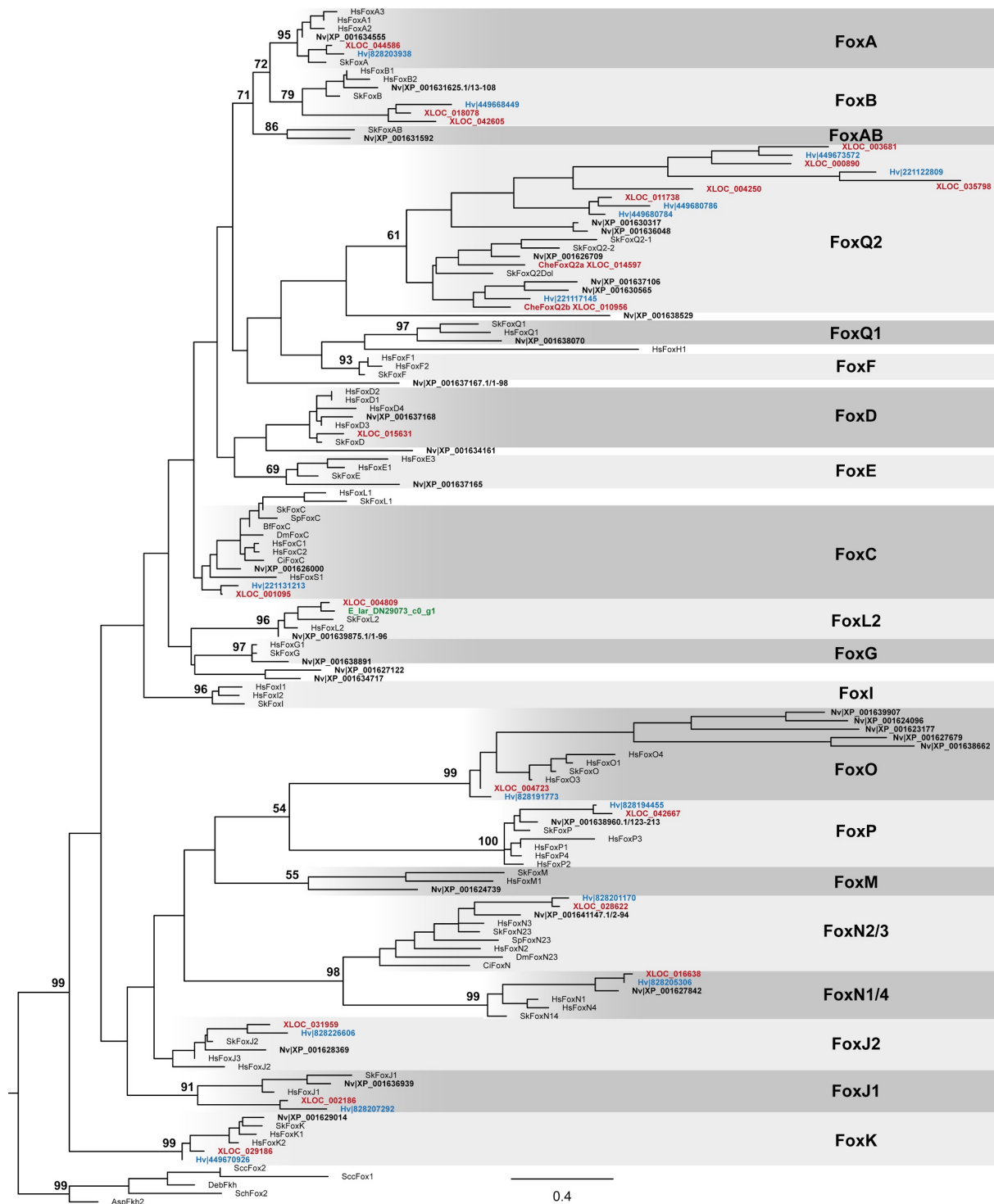




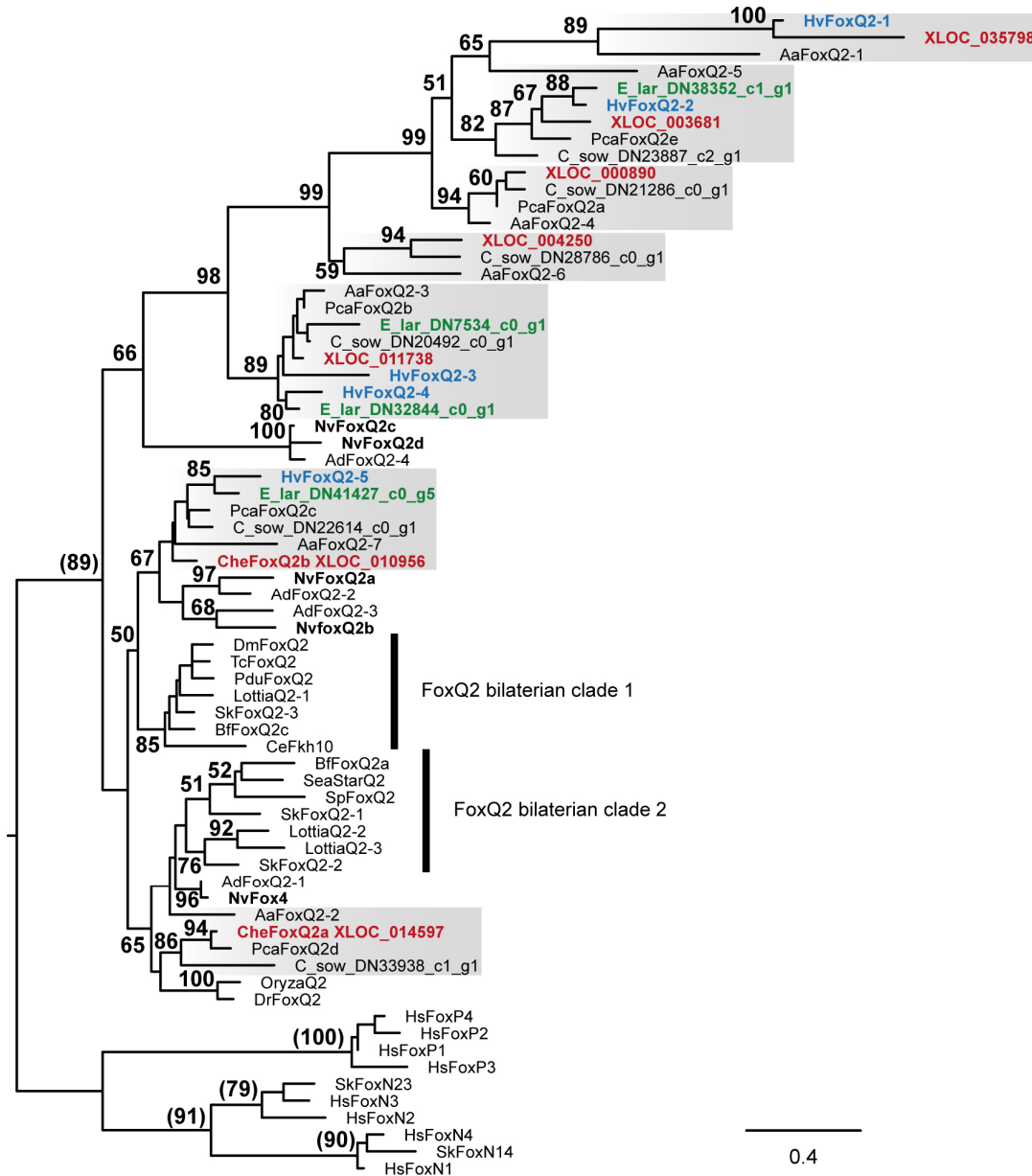
|  | <i>Clytia</i>     | TLX  | PDX   | DRGX  | CDX   | CnoxA | SIX1/2 | FOXL2 |
|--|-------------------|------|-------|-------|-------|-------|--------|-------|
|  | <i>Hydra</i>      | 69.4 | 64.8  | 85.8  | 52.8  | 62.8  | 245.0  | 146.5 |
|  | <i>Ectopleura</i> | 88.6 | 152.8 | 166.0 | 145.7 | 106.5 | 344.5  | 219.4 |
|  | <i>Aurelia</i>    | 87.3 | 156.3 | 144.3 | 107.3 | 111.8 | 344.0  | 257.2 |

**Supplementary figure 6. Evidence for gene loss in *Hydra* of *Clytia* medusa-specific TFs.**

Phylogenetic analysis shows orthologs of the named *Clytia* genes in *Ectopleura* and *Aurelia* but not in *Hydra*, see Supplementary Fig. 6.1-6.6 (*Ectopleura larynx* sequences listed in Supplementary Section 2). The table gives the phmmer bit score<sup>75</sup> for the pairwise comparison of the best hit of the *Clytia* gene in these three medusozoan species. The *Hydra* best hits are non-orthologous and substantially lower scoring than the medusoid producing sister species (*Ectopleura larynx*), or more phylogenetically distant species that retain an ortholog and medusa stage (the scyphozoan *Aurelia aurita*).

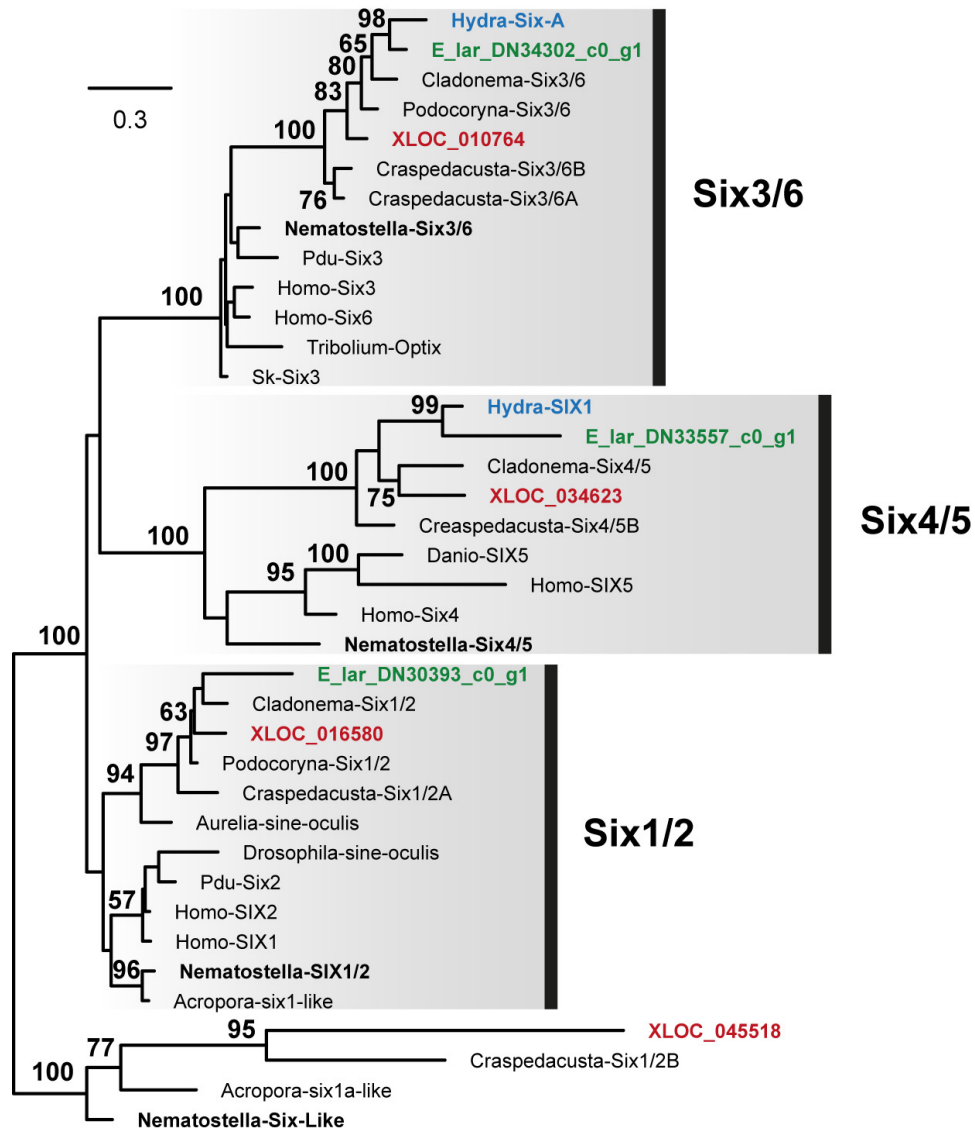


**Supplementary figure 6.1. Maximum likelihood phylogenetic analysis of *Clytia* and *Hydra* Fox-like genes.** Model LG+G RAXML. Rooted with non metazoan Forkhead domains. Values on nodes = rapid RAXML Bootstrap values (500 replicates). Only BP>50% for nodes outside Fox families are labeled. From this topology it can be inferred that FoxAB, FoxE, FoxM, FoxG, FoxQ1 were lost in the common ancestor of *Clytia* and *Hydra* and FoxL2, FoxD, and several FoxQ2 paralogs were lost in the *Hydra* lineage. Red: *Clytia hemisphaerica*, Green: *Ectopleura larynx* FoxL2, black bold: *Nematostella vectensis*, Blue: *Hydra* - Hv: *Hydra vulgaris*, Sk: *Saccoglossus kowalevskii*, Hs: *Homo sapiens*.

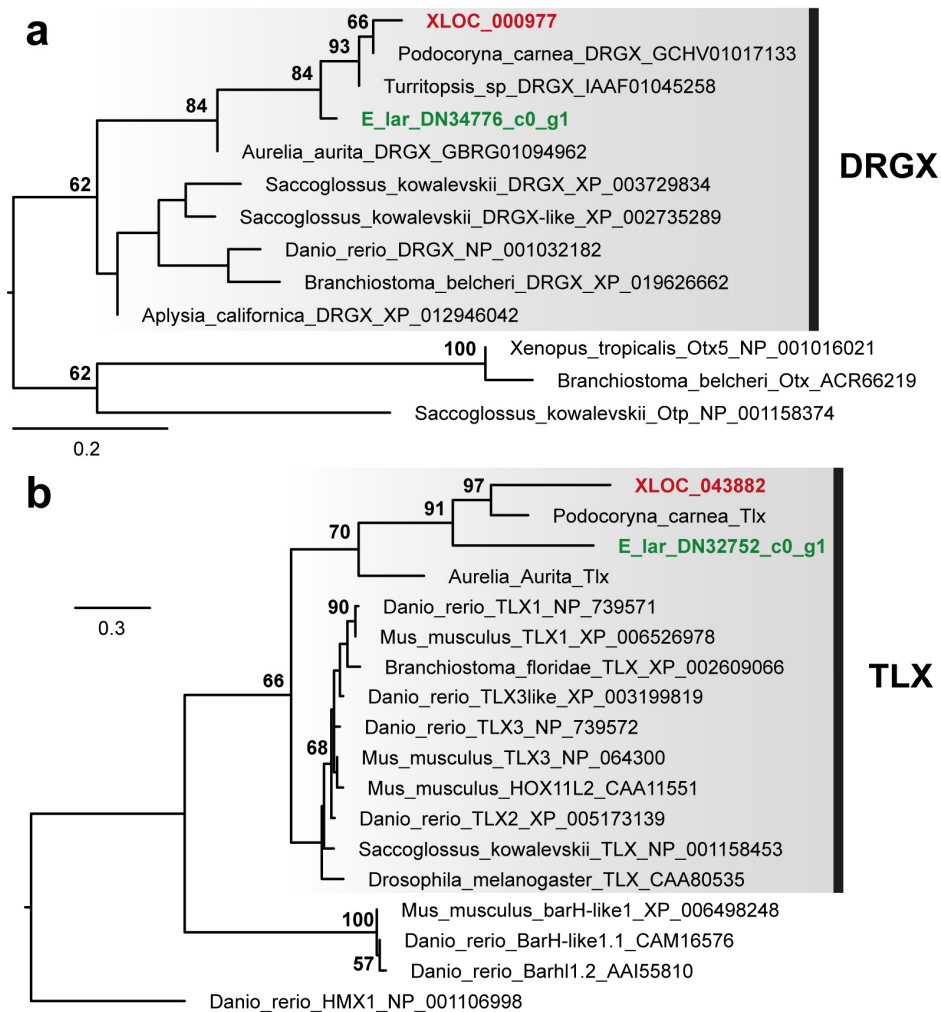


**Supplementary figure 6.2. Maximum likelihood phylogenetic analysis of *Clytia* and *Hydra* FoxQ2-like genes** (Model LG+G RAXML) rooted with FoxP/FoxN14/FoxN23. Rooted and unrooted ML analyses lead to the exact same ingroup topology. rapid RAXML Bootstrap values >50% from the ML analysis of the unrooted dataset are indicated next to the branches. BP values between brackets are from the rooted ML analysis. From this topology, it can be inferred that the cnidarian ancestor had at least 3 FoxQ2-like genes, two of each were also present in the bilaterian ancestor. Expansion of one of those 3 ancestral FoxQ2 gene in the common ancestor of Medusozoa lead to the formation of 7 FoxQ2 medusozoan orthologous groups (highlighted in grey), all still present in the *Clytia* genome. The *Hydra* lineage lost 3 of those FoxQ2 paralogs. Note that the aborally expressed *Clytia* FoxQ2a does not belong to the same FoxQ2 orthologous group as the aborally expressed *Nematostella* FoxQ2a. Highlighted in grey: Medusozoa FoxQ2 paralog groups. Red: *Clytia hemisphaerica* sequences, Green: *Ectopleura larynx* sequences, black bold: *Nematostella vectensis*, Blue: *Hydra* sequences, Hv: *Hydra vulgaris*, Pca: *Podocoryna carnea*, C\_sow: *Craspedacusta sowerbyi*, Aa: *Aurelia aurita*, Sk: *Saccoglossus kowalevskii*, Hs: *Homo sapiens*, Lottia: *Lottia gigantea*, Sp: *Strongylocentrotus*

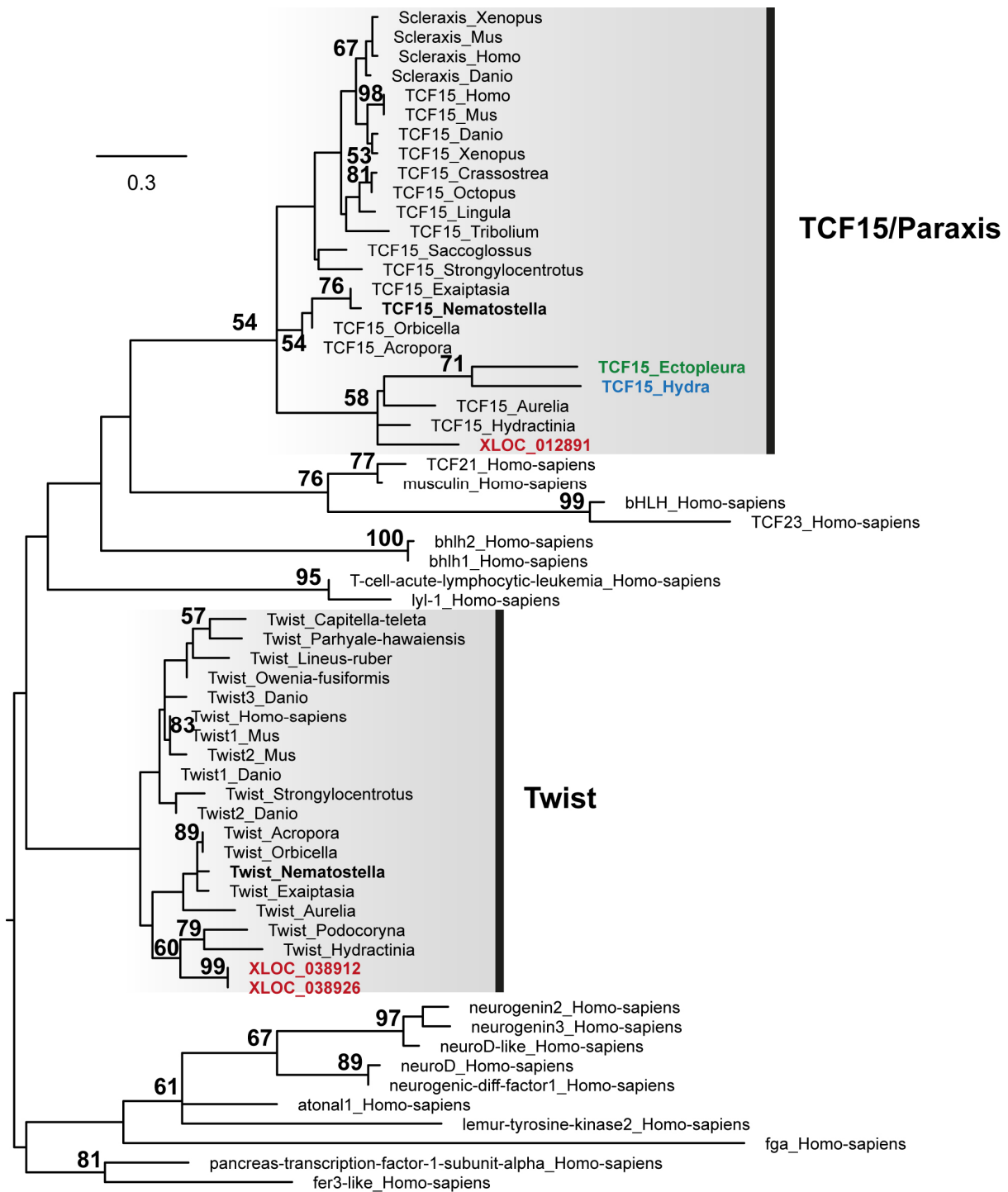
*purpuratus*, Bf: *Branchiostoma floridae*, Pdu: *Platynereis dumerilii*, Dm: *Drosophila melanogaster*, Tc: *Tribolium castaneum*, Ce: *Caenorhabditis elegans*.



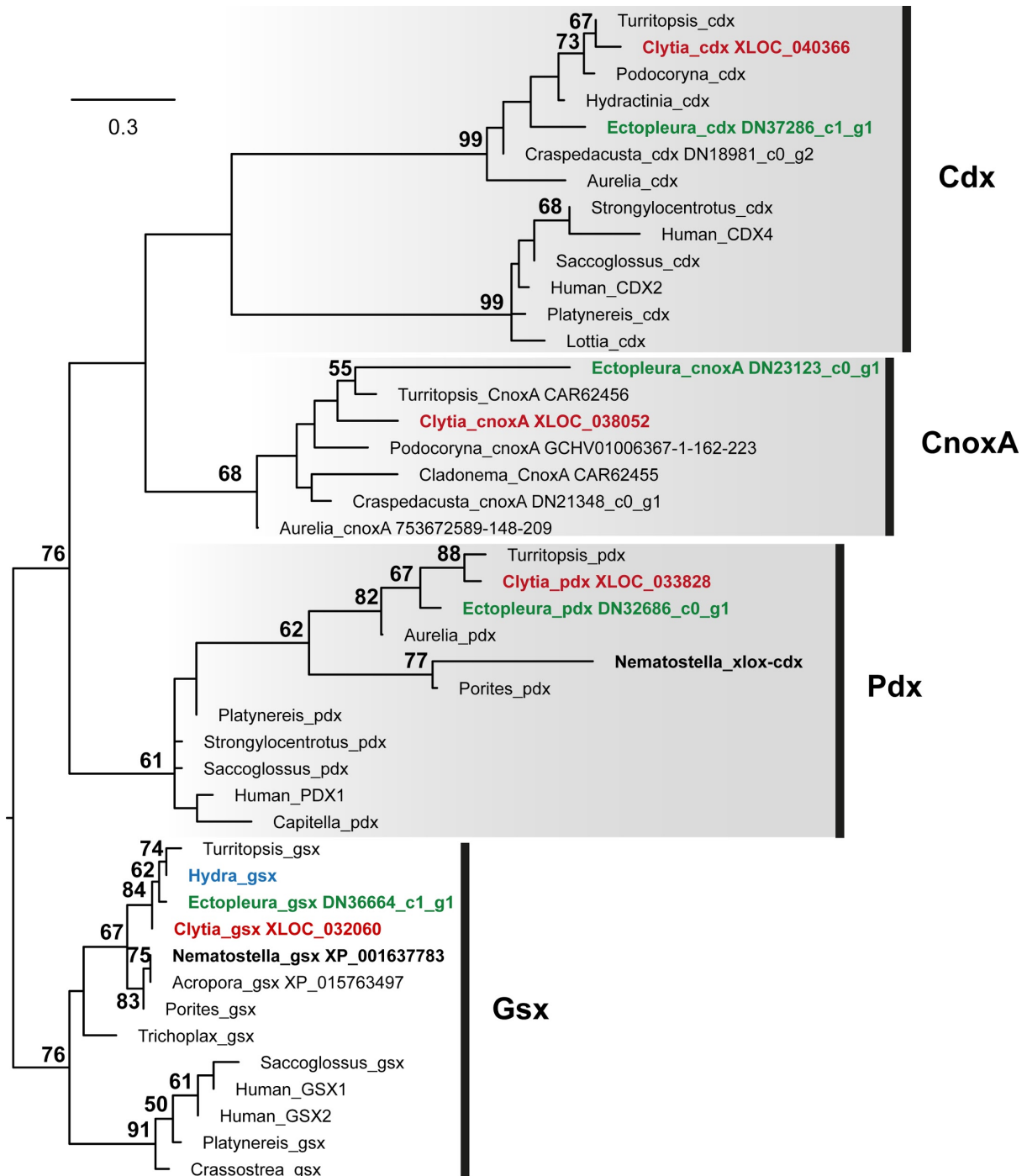
**Supplementary figure 6.3. Maximum likelihood phylogenetic analysis of *Clytia* and *Hydra* Six genes** (Model LG+G RAXML) rooted with distantly related Six-like genes found in cnidarians. Values on nodes = rapid RAXML Bootstrap values (500 replicates). From this topology it can be inferred that Six1/2 was lost in the *Hydra* lineage after its divergence with the *Ectopleura* lineage. Red: *Clytia hemisphaerica*, black bold: *Nematostella vectensis*, blue: *Hydra vulgaris*, green: *Ectopleura larynx*. Pdu: *Platynereis dumerilii*, Sk: *Saccoglossus kowalevskii*.



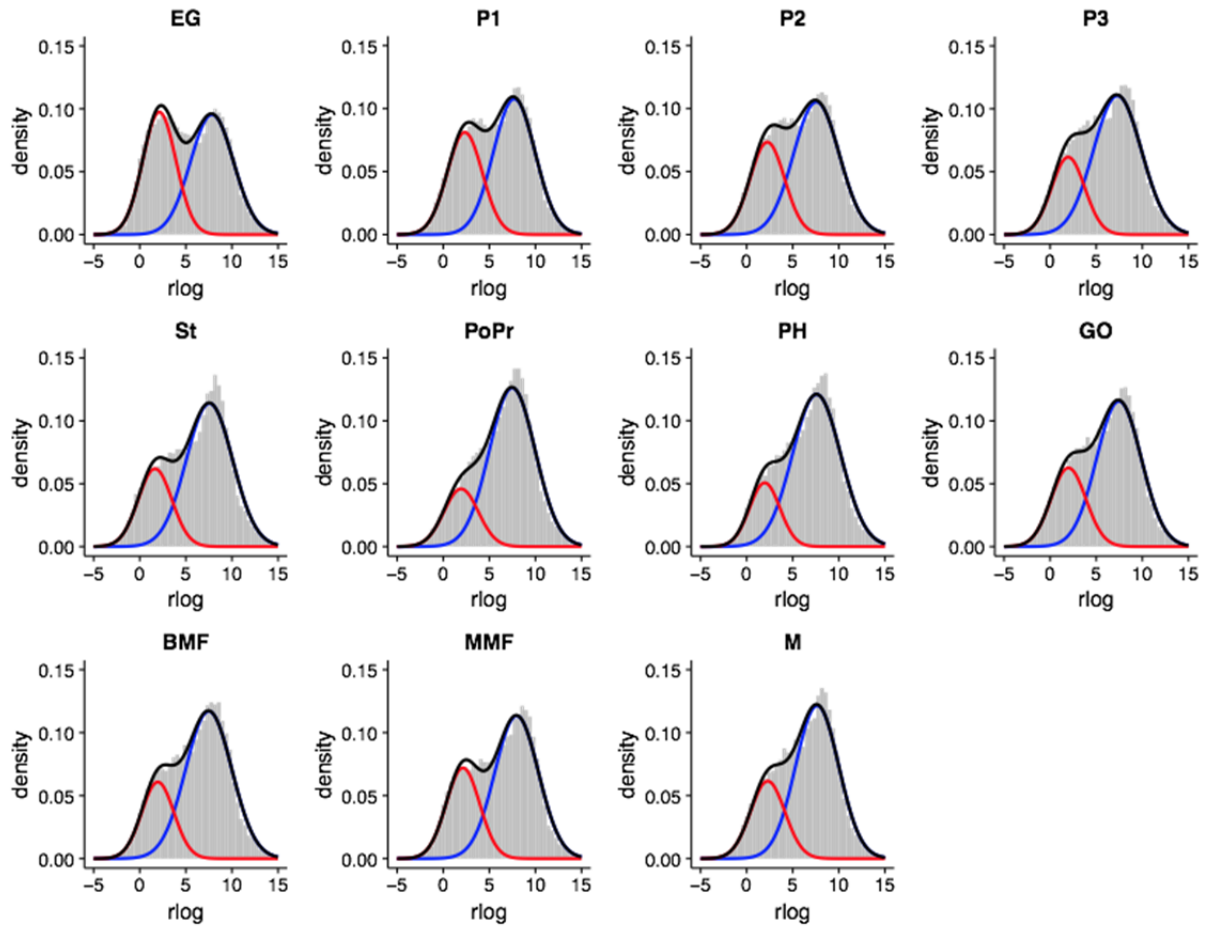
**Supplementary figure 6.4. Maximum likelihood phylogenetic analysis of *Clytia* and *Ectopleura* and DRGX (a) and TLX (b) genes.** Model LG+G RAXML. Values on nodes = rapid RAXML Bootstrap values (500 replicates). Only BP>50% are labeled. Red: *Clytia* sequences, green: *Ectopleura larynx*.



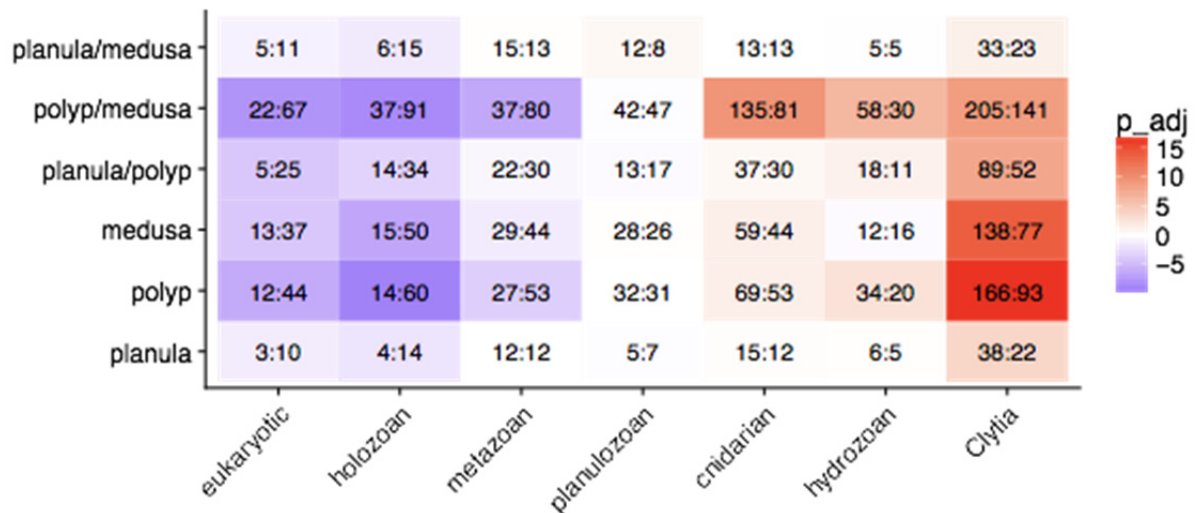
**Supplementary figure 6.5. Maximum likelihood phylogenetic analysis of *Clytia* Twist and TCF15/Paraxis genes** (Model LG+G RAXML) rooted with a selection of human bHLH genes. Values on nodes = rapid RAXML Bootstrap values (500 replicates). From this topology it can be inferred that Twist was lost in a Hydra ancestor. Red: *Clytia hemisphaerica*, black bold: *Nematostella vectensis*, blue: *Hydra vulgaris*, green: *Ectopleura larynx*.



**Supplementary figure 6.6. Maximum likelihood phylogenetic analysis of *Clytia* Cdx, Pdx, CnoxA and Gsx genes** (Model WAG+G RAXML) rooted with Gsx genes. Values on nodes = rapid RAXML Bootstrap values (500 replicates). From this topology it can be inferred that Cdx, Pdx and CnoxA were lost in the *Hydra* lineage after the divergence with *Ectopleura*. Red: *Clytia hemisphaerica*, black bold: *Nematostella vectensis*, blue: *Hydra vulgaris*, green: *Ectopleura larynx*.



**Supplementary figure 7. Gene expression for all life-cycle stages partitioned into ‘on’ and ‘off’ distributions.** Red and blue lines show fitted log-normal distributions and the black line their sum. The blue lines represents genes that we describe as ‘on’, and the red ‘off’. Grey bars correspond to the empirically observed distribution of expression levels. Library names are as for Fig. 1a.



**Supplementary figure 8. Enrichment of OMA HOG phylogenetic class across stage specific genes.** Each cell contains the ratio of observed to expected genes for that stage specificity (vertical axis) and phylogenetic scope, as determined by OMA Hierarchical Orthologous Group composition (horizontal axis). Red cells show an enrichment of the phylogenetic class, blue cells a depletion. Effect significance was determined by a Chi-squared test for each cell, corrected for multiple testing using the Benjamini and Hochberg (FDR) method of the p.adjust R function. These values were used to determine the cell color after  $\log_{10}$  rescaling, and under-represented cells being made negative.

## SUPPLEMENTARY TABLES

Supplementary table 1. Comparison between *Clytia*, *Hydra* and *Nematostella* genomic features.

| Species                       | <i>Clytia hemisphaerica</i> | <i>Hydra vulgaris</i> * | <i>Nematostella vectensis</i> ** |
|-------------------------------|-----------------------------|-------------------------|----------------------------------|
| Clade                         | Hydrozoa                    | Hydrozoa                | Anthozoa                         |
| Presence of jellyfish         | yes                         | no                      | no                               |
| Haploid chromosome number (n) | 15                          | 15                      | 15                               |
| Sequenced genome size (Mb)    | 445                         | 1,050                   | 457                              |
| Contig N50 (kb)               | 3.8                         | 9.7                     | 19.8                             |
| Scaffold N50 (kb)             | 366                         | 92.5                    | 472                              |
| No. scaffolds                 | 7644                        | 39356                   | 10804                            |
| Gap bases (% N)               | 16.6                        | 7.8                     | 16.6                             |
| GC content (%)                | 35                          | 29                      | 39                               |
| Repeats (%)                   | 41                          | 57                      | 26                               |
| Number of genes               | 26,727                      | 31,452                  | 27,273                           |
| BUSCO (genome) (%)            | 86.4                        | 76.8***                 | 91.4***                          |

\* from <sup>12</sup>, RP genome assembly

\*\* from <sup>9</sup>, JGI V1.0 *Nematostella* genome assembly

\*\*\* Numbers recalculated.

### Full BUSCO results for *Clytia hemisphaerica*:

Genome: Complete 86.7% (83.7% single copy, 3% duplicates); Fragments 3.8%; missing 9.5%.

Protein: Complete 90.7% (54.7% single copy, 36.3% duplicates); Fragmented 3.7%; missing 5.6%.

**Supplementary table 2. Genomic location of all predicted homeodomain-containing, Fox and Wnt genes in *Clytia hemisphaerica* and *Hydra vulgaris*.** Genes on the same line are orthologous between *Hydra* and *Clytia*. Two or three genes on the same line as one gene in the second species correspond to species specific paralogs. Genes on the same scaffold in a given species are similarly colored, in bold and labeled in the right column “Cl.”: cluster. Note that none of the gene clusters found in either *Clytia* or *Hydra* can be found in the second species or has been described in Anthozoans. *Hydra* gene ID and genomic location information refer to the 2.0 *Hydra* genome version <https://arussha.nhgri.nih.gov/hydra/>.

**Supplementary table 2.1. Homeodomain containing genes.**

| Orthology group | <i>Clytia hemisphaerica</i> |                                      |          | <i>Hydra vulgaris</i>          |                                       |          |
|-----------------|-----------------------------|--------------------------------------|----------|--------------------------------|---------------------------------------|----------|
|                 | Gene ID                     | Genomic location                     | Cl.      | Gene ID (genome v.2.0)         | Genomic location                      | Cl.      |
| Barhl1          | XLOC_014264                 | sc0000533:122014..127957             |          | Not present                    |                                       |          |
| Barhl2          | XLOC_029225                 | scaffold_100:256460..258860          |          | Not present                    |                                       |          |
| Cdx             | XLOC_040366                 | scaffold_412:34146..38219            |          | Not present                    |                                       |          |
| CnoxA           | <b>XLOC_038052</b>          | <b>scaffold_343:133428..137605</b>   | <b>A</b> | Not present                    |                                       |          |
| Dbx             | Not present                 |                                      |          | Not present                    |                                       |          |
| Dlxa            | <b>XLOC_000289</b>          | <b>sc0000002:1018188..1026889</b>    | <b>B</b> | <b>Sc4wPfr_8.g18245.t1</b>     | <b>Sc4wPfr_8:322182..328724</b>       | <b>A</b> |
| Dlxb            | XLOC_009976                 | sc0000276:159367..168110             |          | Sc4wPfr_330.g6880.t1           | Sc4wPfr_330:825610..831376            |          |
| Dmbx            | XLOC_014679                 | sc0000562:196158..202708             |          | Sc4wPfr_945.g15387.t1          | Sc4wPfr_945:540436..541737            |          |
| DRGX            | <b>XLOC_000977</b>          | <b>sc0000012:938502..948601</b>      | <b>C</b> | Not present                    |                                       |          |
| Ems             | XLOC_034200                 | scaffold_242:587461..593047          |          | Not present                    |                                       |          |
| Emx1            | XLOC_036673                 | scaffold_308:111137..116469          |          | Not present                    |                                       |          |
| Eve             | XLOC_032906                 | scaffold_208:1..42218                |          | Not present                    |                                       |          |
| Gbx             | Not present                 |                                      |          | Not present                    |                                       |          |
| Gsc             | <b>XLOC_010754</b>          | <b>sc0000320:160388..166688</b>      | <b>D</b> | Sc4wPfr_802.g11765.t1          | Sc4wPfr_802:97808..101470             |          |
| Gsx             | XLOC_032060                 | scaffold_18:133511..136155           |          | Sc4wPfr_165.g10051.t1          | Sc4wPfr_165:1632789..1634340          |          |
| Hbn             | XLOC_003372                 | sc0000061:471651..489633             |          | Sc4wPfr_654.g18227.t1          | Sc4wPfr_654:590461..592016            |          |
| Hhex            | XLOC_023193                 | sc0001897:33048..35972               |          | Sc4wPfr_391.g24677.t1          | Sc4wPfr_391:1967642..1971182          |          |
| Hmx1/Nk5        | XLOC_030138                 | scaffold_125:69368..89190            |          | Not present                    |                                       |          |
| Hmx2/Nk5        | XLOC_043454                 | scaffold_508:131557..139297          |          | Not present                    |                                       |          |
| Hox1            | <b>XLOC_038841</b>          | <b>scaffold_361:1685165..1689203</b> | <b>E</b> | Sc4wPfr_717.g8768.t1           | Sc4wPfr_717:840449..841903            |          |
| Hox9-14a        | Not predicted               | sc0000309:278071..281544             |          | Sc4wPfr_62.g4503.t1            | Sc4wPfr_62:3086285..3087453           |          |
| Hox9-14b        | XLOC_004917                 | sc0000087:265188..268577             |          | Sc4wPfr_230.g26228.t1          | Sc4wPfr_230:414963..418609            |          |
| Hox9-14c        | <b>XLOC_038055</b>          | <b>scaffold_343:225948..229139</b>   | <b>A</b> | Sc4wPfr_212.1.g3831.t1         | Sc4wPfr_212.1:378594..379911          |          |
|                 |                             |                                      |          | Sc4wPfr_547.1.g24998.t1        | Sc4wPfr_547.1:1166112..1189029        |          |
| Hox-like        | XLOC_004823                 | sc0000086:517516..523926             |          | Not present                    |                                       |          |
| Irx             | XLOC_009244                 | sc0000241:343654..349240             |          | <b>Sc4wPfr_439.g20769.t1</b>   | <b>Sc4wPfr_439:802583..806014</b>     | <b>B</b> |
| Islet           | XLOC_004661                 | sc0000084:610644..646448             |          | not present                    |                                       |          |
| Lhx1/5 lin-11   | XLOC_043220                 | scaffold_498:276695..314892          |          | Sc4wPfr_706.2.g8000.t1         | Sc4wPfr_706.2:1233632..1258094        |          |
| Lhx2/9 Apterous | XLOC_002692                 | sc0000048:411418..432961             |          | <b>Sc4wPfr_47.1.g2187.t1</b>   | <b>Sc4wPfr_47.1:188277..201641</b>    | <b>C</b> |
| Lmx             | XLOC_007074                 | sc0000146:2998..43912                |          | <b>Sc4wPfr_47.1.g2192.t1</b>   | <b>Sc4wPfr_47.1:237406..263850</b>    | <b>C</b> |
| Lhx6/8_Awh-like | XLOC_012557                 | sc0000414:60351..66533               |          | Sc4wPfr_504.g10379.t1          | Sc4wPfr_504:1539023..1539878          |          |
| PKNOX           | XLOC_033395                 | scaffold_220:475373..491505          |          | Not present                    |                                       |          |
| Meis1-like      | XLOC_002564                 | sc0000045:638000..714333             |          | Sc4wPfr_305.g4616.t1           | Sc4wPfr_305:1267981..1282322          |          |
| Mnx             | Not present                 |                                      |          | Not present                    |                                       |          |
| Mox             | <b>XLOC_033055</b>          | <b>scaffold_213:11832..70934</b>     | <b>F</b> | Sc4wPfr_293.g24401.t1          | Sc4wPfr_293:190902..192159            |          |
| Msx             | XLOC_035505                 | scaffold_273:15027..20135            |          | Sc4wPfr_87.g16557.t1           | Sc4wPfr_87:1610036..1613443           |          |
| Msx1x           | XLOC_013455                 | sc0000465:136354..146659             |          | Not present                    |                                       |          |
| nk2a            | XLOC_009058                 | sc0000231:31623..65306               |          | Sc4wPfr_884.g31954.t1          | Sc4wPfr_884:250152..258271            |          |
| nk2b            | XLOC_043051                 | scaffold_493:213978..217686          |          | Not present                    |                                       |          |
| nk2c            | XLOC_007385                 | sc0000163:313515..360252             |          | Not present                    |                                       |          |
| nk2d            | XLOC_043740                 | scaffold_516:395696..405514          |          | Sc4wPfr_910.1.g9205.t1         | Sc4wPfr_910.1:781492..784692          |          |
| Nkx1            | XLOC_016820                 | sc0000749:123829..130805             |          | Sc4wPfr_350.g2253.t1           | Sc4wPfr_350:20535..29166              |          |
| Nkx6            | Not predicted               | scaffold_348:560533..562015          |          | Not present                    |                                       |          |
| Otp             | XLOC_016543                 | sc0000724:61496..91324               |          | <b>Sc4wPfr_1357.1.g3228.t1</b> | <b>Sc4wPfr_1357.1:277787..291177</b>  | <b>D</b> |
|                 |                             |                                      |          | <b>Sc4wPfr_1357.1.g3222.t1</b> | <b>Sc4wPfr_1357.1:152198..166787</b>  | <b>D</b> |
| OtxA            | <b>XLOC_000267</b>          | <b>sc0000002:15611..25685</b>        | <b>B</b> | <b>Sc4wPfr_224.1.g33440.t1</b> | <b>Sc4wPfr_224.1:1795345..1813005</b> | <b>E</b> |
| OtxB            | XLOC_040405                 | scaffold_415:222146..255072          |          | <b>Sc4wPfr_224.1.g33396.t1</b> | <b>Sc4wPfr_224.1:752585..765688</b>   | <b>E</b> |
| OtxC            | XLOC_039817                 | scaffold_391:77706..106174           |          | Sc4wPfr_804.g24114.t1          | Sc4wPfr_804:956580..960357            |          |
| PaxB            | XLOC_010541                 | sc0000309:150119..165463             |          | <b>Sc4wPfr_439.g20916.t1</b>   | <b>Sc4wPfr_439:4086426..4113494</b>   | <b>B</b> |
| PaxD/3-7        | Not present                 |                                      |          | Not present                    |                                       |          |
| PaxE            | XLOC_023354                 | sc0001955:22030..33480               |          | Not present                    |                                       |          |

|                   |                    |                                      |                            |                                 |                              |
|-------------------|--------------------|--------------------------------------|----------------------------|---------------------------------|------------------------------|
| Pbx1              | XLOC_036148        | scaffold_288:478403..491550          | Sc4wPfr_1241.g21489.t1     | Sc4wPfr_1241:404054..461013     |                              |
|                   |                    |                                      | Sc4wPfr_2124.g26585.t1     | Sc4wPfr_2124:21188..89317       |                              |
| Pdx/Xlox          | XLOC_033828        | scaffold_233:1429381..1436630        | Not present                |                                 |                              |
| Pitx              | XLOC_033274        | scaffold_218:307560..335047          | Sc4wPfr_390.g5621.t1       | Sc4wPfr_390:1472582..1479119    |                              |
| POU class 6       | <b>XLOC_038904</b> | <b>scaffold_361:1691815..1711422</b> | <b>E</b>                   | Sc4wPfr_78.g10424.t1            | Sc4wPfr_78:265168..398986    |
| POU class 4       | <b>XLOC_001017</b> | <b>sc0000012:1097269..1104976</b>    | <b>C</b>                   | Sc4wPfr_287.g9045.t1            | Sc4wPfr_287:206879..208405   |
| POU class 2/3     | XLOC_033684        | scaffold_228:401920..404369          | Not present                |                                 |                              |
| Prdl-a            | XLOC_030117        | scaffold_124:61129..73405            | Sc4wPfr_1080.g15226.t1     | Sc4wPfr_1080:482234..487811     |                              |
| Prdl-b            | XLOC_004378        | sc0000079:13871..24278               | Sc4wPfr_372.g27997.t1      | Sc4wPfr_372:979316..993911      |                              |
| Prdl-d            | XLOC_014392        | sc0000541:164178..190096             | Sc4wPfr_106.g23541.t1      | Sc4wPfr_106:1131009..1132435    |                              |
| Rax               | Not present        |                                      | Not present                |                                 |                              |
| Ro/Rough          | Not present        |                                      | Not present                |                                 |                              |
| SHOX              | XLOC_013191        | sc0000454:157585..166343             | Sc4wPfr_113.1.g28302.t1    | Sc4wPfr_113.1:82385..83889      |                              |
| Six1/2            | XLOC_016580        | sc0000726:95120..103003              | Not present                |                                 |                              |
| Six3/6            | <b>XLOC_010764</b> | <b>sc0000320:121240..125619</b>      | <b>D</b>                   | Sc4wPfr_226.1.g26815.t1         | Sc4wPfr_226.1:911649..916550 |
| Six4/5            | XLOC_034623        | scaffold_247:25902..33180            | Sc4wPfr_275.3.g1976.t1     | Sc4wPfr_275.3:182138..184881    |                              |
| Six-like1         | XLOC_045518        | scaffold_90:214078..216417           | Not present                |                                 |                              |
| Six-like2         | XLOC_006968        | sc0000144:219276..236334             | Not present                |                                 |                              |
| TLX               | XLOC_043882        | scaffold_56:784543..787295           | Not present                |                                 |                              |
| Unc-4             | XLOC_039146        | scaffold_365:143528..147099          | Sc4wPfr_1362.g13664.t1     | Sc4wPfr_1362:490042..556968     |                              |
| Vsx               | <b>XLOC_010760</b> | <b>sc0000320:37791..41165</b>        | <b>D</b>                   | Sc4wPfr_1010.g13528.t1          | Sc4wPfr_1010:512602..515395  |
| Vax-like          | <b>XLOC_033108</b> | <b>scaffold_213:810437..812436</b>   | <b>F</b>                   | Not present                     |                              |
| No clear ortholog | XLOC_002261        | sc0000039:323558..333581             | Sc4wPfr_552.g10518.t1      | Sc4wPfr_552:884081..887330      |                              |
|                   | XLOC_044663        | scaffold_65:23906..25473             |                            |                                 |                              |
| No clear ortholog | XLOC_041744        | scaffold_46:39918..42065             | Sc4wPfr_144.1.g11736.t1    | Sc4wPfr_144.1:138290..139249    |                              |
| No clear ortholog | XLOC_018123        | sc0000903:8898..62945                | Sc4wPfr_198.g23296.t1      | Sc4wPfr_198:1486542..1489140    |                              |
| No clear ortholog | XLOC_044613        | scaffold_64:15267..19115             | Sc4wPfr_363.2.g30386.t1    | Sc4wPfr_363.2:1655171..1659929  |                              |
| No clear ortholog | Not present        |                                      | <b>Sc4wPfr_8.g18251.t3</b> | <b>Sc4wPfr_8:439689..452031</b> | <b>A</b>                     |
| No clear ortholog | XLOC_010598        | sc0000313:27555..29151               | Not present                |                                 |                              |
| No clear ortholog | XLOC_042134        | scaffold_47:496995..503800           | Not present                |                                 |                              |
| No clear ortholog | XLOC_015070        | sc0000591:172855..175172             | Not present                |                                 |                              |
| No clear ortholog | XLOC_018560        | sc0000952:11158..52548               | Not present                |                                 |                              |
| No clear ortholog | XLOC_042894        | scaffold_49:411948..414094           | Not present                |                                 |                              |

**Supplementary table 2.2. Fox genes.**

| Orthology group | <i>Clytia hemisphaerica</i> |                             | <i>Hydra vulgaris</i>       |                                   |          |
|-----------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------------|----------|
|                 | Gene ID                     | Genomic location            | Gene ID (genome v.2.0)      | Genomic location                  | Cl.      |
| FoxA            | XLOC_044586                 | scaffold_63:651015..654137  | Sc4wPfr_804.g24126.t1       | Sc4wPfr_804:1448171..1449522      |          |
| FoxB            | XLOC_018078                 | sc0000898:35537..52055      | Sc4wPfr_27.g12217.t1        | Sc4wPfr_27:1189606..1246921       |          |
|                 | XLOC_042605                 | scaffold_482:235994..258030 |                             |                                   |          |
| FoxAB           | Not present                 |                             | Not present                 |                                   |          |
| FoxC            | XLOC_001095                 | sc0000018:926903..939017    | Sc4wPfr_48.g30219.t1        | Sc4wPfr_48:480863..482828         |          |
| FoxD            | XLOC_015631                 | sc0000639:134797..138263    | Not present                 |                                   |          |
| FoxE            | Not present                 |                             | Not present                 |                                   |          |
| FoxF            | Not present                 |                             | Not present                 |                                   |          |
| FoxG            | Not present                 |                             | Not present                 |                                   |          |
| FoxI            | Not present                 |                             | Not present                 |                                   |          |
| FoxJ1           | XLOC_002186                 | sc0000038:481628..513842    | Sc4wPfr_719.g8279.t1        | Sc4wPfr_719:261216..265675        |          |
| FoxJ3           | XLOC_031959                 | scaffold_172:86480..101564  | Sc4wPfr_275.3.g1979.t1      | Sc4wPfr_275.3:282531..326443      |          |
| FoxK            | XLOC_029186                 | scaffold_100:113458..129970 | Sc4wPfr_441.g7675.t1        | Sc4wPfr_441:228395..241398        |          |
| FoxL2           | XLOC_004809                 | sc0000086:104787..109541    | Not present                 |                                   |          |
| FoxM            | Not present                 |                             | Not present                 |                                   |          |
| FoxN1/4         | XLOC_016638                 | sc0000730:71689..78690      | Sc4wPfr_546.g25835.t1       | Sc4wPfr_546:225758..232799        |          |
| FoxN2/3         | XLOC_028622                 | sc0006737:1..1506           | Sc4wPfr_921.g15144.t1       | Sc4wPfr_921:293012..295797        |          |
| FoxO            | XLOC_004723                 | sc0000085:8496..16763       | Sc4wPfr_909.g33493.t1       | Sc4wPfr_909:388907..391788        |          |
| FoxP            | XLOC_042667                 | scaffold_483:74611..97485   | Sc4wPfr_396.g3075.t1        | Sc4wPfr_396:1599761..1632651      |          |
| FoxQ1           | Not present                 |                             | Not present                 |                                   |          |
| FoxQ2a          | XLOC_014597                 | sc0000555:136278..138727    | Not present                 |                                   |          |
| FoxQ2b          | XLOC_010956                 | sc0000328:114974..117572    | Sc4wPfr_1172.g28449.t1      | Sc4wPfr_1172:62339..63456         |          |
| FoxQ2-like      | XLOC_003681                 | sc0000067:179592..195284    | Sc4wPfr_802.g11764.t1       | Sc4wPfr_802:83295..84541          |          |
|                 | XLOC_035798                 | scaffold_283:70696..74554   | Sc4wPfr_381.1.g32304.t1     | Sc4wPfr_381.1:3641170..3642287    |          |
|                 | XLOC_011738                 | sc0000373:214400..217632    | <b>Sc4wPfr_390.g5577.t1</b> | <b>Sc4wPfr_390:78606..79818</b>   | <b>A</b> |
|                 | XLOC_000890                 | sc0000011:52464..56290      | <b>Sc4wPfr_390.g5578.t1</b> | <b>Sc4wPfr_390:102917..103709</b> | <b>A</b> |
|                 | XLOC_004250                 | sc0000077:443293..445454    |                             |                                   |          |

**Supplementary table 2.3. Wnt genes.**

| Orthology group* | <i>Clytia hemisphaerica</i> |                             | <i>Hydra vulgaris</i>               |                                           |                   |
|------------------|-----------------------------|-----------------------------|-------------------------------------|-------------------------------------------|-------------------|
|                  | Gene ID                     | Genomic location            | Gene ID (genome v.2.0)              | Genomic location                          | Cl.               |
| Wnt1             | XLOC_015404                 | sc0000619:174979..186143    | Sc4wPfr_804.g24094.t1               | Sc4wPfr_804:309815..324652                |                   |
|                  | XLOC_015382                 | sc0000616:43035..76017      |                                     |                                           |                   |
| Wnt2             | XLOC_031686                 | scaffold_167:172713..185233 | Sc4wPfr_439.g20853.t1               | Sc4wPfr_439:2641798..2690844              |                   |
| Wnt3             | XLOC_001931                 | sc0000028:165051..178177    | Sc4wPfr_399.g28064.t1               | Sc4wPfr_399:439060..443036                |                   |
| Wnt4             | XLOC_041060                 | scaffold_44:356205..369174  | Sc4wPfr_275.3.g1992.t1              | Sc4wPfr_275.3:647975..683454              |                   |
| Wnt5             | XLOC_000650                 | sc0000007:944727..949345    | Sc4wPfr_287.2.g28262.t1             | Sc4wPfr_287.2:1208089..1215135            |                   |
|                  | XLOC_030003                 | scaffold_120:225773..240660 |                                     |                                           |                   |
| Wnt6             | XLOC_035748                 | scaffold_282:294654..301844 | Sc4wPfr_224.1.g33373.t1             | Sc4wPfr_224.1:64500..66643                |                   |
|                  |                             |                             | <a href="#">Sc4wPfr_68.g2320.t1</a> | <a href="#">Sc4wPfr_68:487676..489656</a> | <a href="#">A</a> |
|                  |                             |                             | <a href="#">Sc4wPfr_68.g2322.t1</a> | <a href="#">Sc4wPfr_68:512888..516766</a> | <a href="#">A</a> |
| Wnt7             | XLOC_034538                 | scaffold_245:573924..608832 | Sc4wPfr_1214.g29750.t1              | Sc4wPfr_1214:1020242..1023228             |                   |
| Wnt8             | XLOC_045775                 | scaffold_97:521345..540707  | Sc4wPfr_440.g6790.t1                | Sc4wPfr_440:137806..138763                |                   |
| Wnt11            | Unpredicted                 | scaffold_17:47802..49174    | Sc4wPfr_243.g15582.t1               | Sc4wPfr_243:55736..225710                 |                   |
|                  | XLOC_002938                 | sc0000051:670555..674222    |                                     |                                           |                   |
| WntA             | XLOC_006909                 | sc0000143:348385..370576    | Not present                         |                                           |                   |

\* Revised *Clytia* and *Hydra* Wnt names are used, based on revised phylogeny and orthology : TC & MM manuscript in preparation.

**Supplementary table 3. Significant gene ontology enrichments at major life-cycle stages relative to the genome as a whole.**

| Planula    | Adjusted P | Ontology Term                                      |
|------------|------------|----------------------------------------------------|
| GO:0004872 | 9.6E-07    | receptor activity                                  |
| GO:0004930 | 3.2E-05    | G-protein coupled receptor activity                |
| GO:0005886 | 2.4E-04    | plasma membrane                                    |
| GO:0007186 | 2.0E-03    | G-protein coupled receptor signaling pathway       |
| GO:0007154 | 6.2E-03    | cell communication                                 |
|            |            |                                                    |
| Polyp      | Adjusted P | Ontology Term                                      |
| GO:0004872 | 3.6E-11    | receptor activity                                  |
| GO:0005576 | 1.1E-10    | extracellular region                               |
| GO:0005102 | 5.7E-10    | receptor binding                                   |
| GO:0005886 | 8.5E-09    | plasma membrane                                    |
| GO:0007154 | 6.3E-08    | cell communication                                 |
| GO:0022610 | 1.3E-05    | biological adhesion                                |
| GO:0008081 | 1.5E-05    | phosphoric diester hydrolase activity              |
| GO:0016298 | 2.6E-05    | lipase activity                                    |
| GO:0019210 | 3.8E-04    | kinase inhibitor activity                          |
| GO:0050896 | 3.2E-03    | response to stimulus                               |
| GO:0032501 | 9.6E-03    | multicellular organismal process                   |
|            |            |                                                    |
| Medusa     | Adjusted P | Ontology Term                                      |
| GO:0022610 | 4.1E-07    | biological adhesion                                |
| GO:0003677 | 1.4E-03    | DNA binding                                        |
| GO:0001071 | 7.9E-03    | nucleic acid binding transcription factor activity |
| GO:0005576 | 8.7E-03    | extracellular region                               |

**Supplementary table 4. Expression patterns and genome location for all transcription factors (4.1) and for all predicted genes (4.2). See supplementary xlsx file.**

**Supplementary table 5. Classification of medusa specific transcription factors.**

List of *Clytia* medusa specific transcription factor and presence/absence of orthologs in *Hydra* (draft genome) and *Ectopleura* (transcriptome), its closest relative with a medusoid stage. Note that 7 *Clytia* medusa specific TF are absent in *Hydra* but present in *Ectopleura*, suggesting the loss of these TF in the *Hydra* lineage was related to loss of the medusa stage. Ortholog for only one *Clytia* medusa specific TF not sex specific, TCF15/Paraxis, was found in *Hydra*. na: non-applicable. *Ectopleura* sequences listed at the end of the supplementary material.

| <i>Clytia</i> gene ID | Sex specific    | Status                                                              | Gene family                 | <i>Hydra</i> ortholog | <i>Ectopleura</i> ortholog |
|-----------------------|-----------------|---------------------------------------------------------------------|-----------------------------|-----------------------|----------------------------|
| XLOC_033897           | Female enriched | present in <i>Hydra</i>                                             | Max-like                    | XP_012559560          | ?                          |
| XLOC_033152           | Male enriched   | present in <i>Hydra</i>                                             | Nitrogen Regulatory Protein | Sc4wPfr_703.g8940.t1  | E_lar_DN34629_c0_g1        |
| XLOC_010965           | Male enriched   | present in <i>Hydra</i>                                             | CREB-binding protein        | XP_004207175          | E_lar_DN34214_c0_g1        |
| XLOC_038135           | Male enriched   | present in <i>Hydra</i>                                             | heat shock factor           | XP_012558243          | E_lar_DN33986_c0_g1        |
| XLOC_012891           | no              | present in <i>Hydra</i>                                             | TCF15/Paraxis               | XP_012565575          | E_lar_DN39961_c0_g1        |
| XLOC_038912           | no              | lost in <i>Hydra</i>                                                | Twist(a)                    | no                    | ?                          |
| XLOC_038926           | no              | lost in <i>Hydra</i>                                                | Twist(b)                    | no                    | ?                          |
| XLOC_014264           | no              | lost in <i>Hydra</i>                                                | barH-like                   | no                    | ?                          |
| XLOC_004250           | no              | lost in <i>Hydra</i>                                                | FoxQ2 paralog               | no                    | ?                          |
| XLOC_000890           | no              | lost in <i>Hydra</i>                                                | FoxQ2 paralog               | no                    | ?                          |
| XLOC_036673           | no              | lost in <i>Hydra</i>                                                | Emx1                        | no                    | ?                          |
| XLOC_033828           | no              | lost in the <i>Hydra</i> lineage since <i>Ectopleura</i> divergence | pdx/xlox                    | no                    | E_lar_DN32686_c0_g1        |
| XLOC_000977           | no              | lost in the <i>Hydra</i> lineage since <i>Ectopleura</i> divergence | DRGX                        | no                    | E_lar_DN34776_c0_g1        |
| XLOC_043882           | no              | lost in the <i>Hydra</i> lineage since <i>Ectopleura</i> divergence | TLX                         | no                    | E_lar_DN32752_c0_g1        |
| XLOC_016580           | no              | lost in the <i>Hydra</i> lineage since <i>Ectopleura</i> divergence | Six1/2                      | no                    | E_lar_DN30393_c0_g1        |
| XLOC_004809           | no              | lost in the <i>Hydra</i> lineage since <i>Ectopleura</i> divergence | FoxL2                       | no                    | E_lar_DN29073_c0_g1        |
| XLOC_038052           | no              | lost in the <i>Hydra</i> lineage since <i>Ectopleura</i> divergence | CnoxA                       | no                    | E_lar_DN23123_c0_g1        |
| XLOC_040366           | no              | lost in the <i>Hydra</i> lineage since <i>Ectopleura</i> divergence | CDX                         | no                    | E_lar_DN32786_c1_g1        |
| XLOC_011741           | no              | lost in <i>Hydra</i> or <i>Clytia</i> specific gene                 | brachyury 3                 | no                    | ?                          |
| XLOC_032901           | no              | lost in <i>Hydra</i> or <i>Clytia</i> specific gene                 | zinc finger protein         | no                    | ?                          |
| XLOC_013346           | Male enriched   | <i>Clytia</i> specific gene                                         | Sox orphan                  | no                    | ?                          |
| XLOC_042894           | Male enriched   | <i>Clytia</i> specific gene                                         | Homeobox orphan             | no                    | ?                          |
| XLOC_006210           | no              | <i>Clytia</i> specific gene                                         | Tbox orphan                 | no                    | ?                          |
| XLOC_011841           | no              | Paralog of XLOC_030196 not medusa specific                          | HARBI1 paralog              | na                    | na                         |
| XLOC_035039           | no              | Paralog of XLOC_035034 not medusa specific                          | zinc finger protein         | na                    | na                         |
| XLOC_028449           | no              | Paralog of XLOC_029182 not medusa specific                          | zinc finger protein         | na                    | na                         |
| XLOC_035033           | no              | Paralog of XLOC_029182 not medusa specific                          | zinc finger protein         | na                    | na                         |
| XLOC_035038           | no              | Paralog of XLOC_029160 not medusa specific                          | zinc finger protein         | na                    | na                         |
| XLOC_035044           | no              | Paralog of XLOC_029160 not medusa specific                          | zinc finger protein         | na                    | na                         |
| XLOC_035032           | Female enriched | Paralog of XLOC_029182 not medusa specific                          | zinc finger protein         | na                    | na                         |
| XLOC_023898           | Female enriched | Paralog of XLOC_036398 not medusa specific                          | zinc finger protein         | na                    | na                         |
| XLOC_033110           | Male enriched   | Paralog of XLOC_038551 not medusa specific                          | Prdm9-like                  | na                    | na                         |
| XLOC_003858           | Male enriched   | Paralog of XLOC_038551 not medusa specific                          | Prdm9-like                  | na                    | na                         |
| XLOC_017073           | Male enriched   | Paralog of XLOC_038551 not medusa specific                          | Prdm9-like                  | na                    | na                         |

**Supplementary table 6. RNA preparations for individual libraries.**

| Stage                | Name                     | Strain    | Nb individuals | Total RNA (ng) | Illumina   | cycles |
|----------------------|--------------------------|-----------|----------------|----------------|------------|--------|
| Early gastrula       | EG-1                     | Z4B * Z10 | 150            | 1220           | NextSeq    | 75 SR  |
| Early gastrula       | EG-1                     | Z4B * Z10 | 150            | 1260           | NextSeq    | 75 SR  |
| Planula 24hpf        | P1-1                     | Z4B * Z10 | 150            | 2390           | NextSeq    | 75 SR  |
| Planula 24hpf        | P1-2                     | Z4B * Z10 | 150            | 1760           | NextSeq    | 75 SR  |
| Planula 48hpf        | P2-1                     | Z4B * Z10 | 150            | 1040           | NextSeq    | 75 SR  |
| Planula 48hpf        | P2-2                     | Z4B * Z10 | 150            | 520            | NextSeq    | 75 SR  |
| Planula 72hpf        | P3-1                     | Z4B * Z10 | 150            | 520            | NextSeq    | 75 SR  |
| Planula 72hpf        | P3-2                     | Z4B * Z10 | 150            | 1105           | NextSeq    | 75 SR  |
| Primary polyp        | PoPr-1                   | Z4B * Z10 | 100            | 494            | NextSeq    | 75 SR  |
| Primary polyp        | PoPr-2                   | Z4B * Z10 | 100            | 450            | NextSeq    | 75 SR  |
| Gastrozoid           | PH-1                     | Z4B       | 60             | 224            | NextSeq    | 75 SR  |
| Gastrozoid           | PH-2                     | Z4B       | 60             | 936            | NextSeq    | 75 SR  |
| Gonozoid             | GO-1                     | Z4B       | 60             | 400            | NextSeq    | 75 SR  |
| Gonozoid             | GO-2                     | Z4B       | 60             | 1365           | NextSeq    | 75 SR  |
| Stolon               | St-1                     | Z4B       | 60             | 1131           | NextSeq    | 75 SR  |
| Stolon               | St-2                     | Z4B       | 60             | 780            | NextSeq    | 75 SR  |
| Baby medusa          | BMF-R1                   | Z4B       | 100            | 1523           | HiSeq 2500 | 50 SR  |
| Baby medusa          | BMF-R2                   | Z4B       | 100            | 1708           | HiSeq 2500 | 50 SR  |
| Mature female medusa | MMF-R1                   | Z4B       | 1              | 2607           | HiSeq 2500 | 50 SR  |
| Mature female medusa | MMF-R2                   | Z4B       | 1              | 2319           | HiSeq 2500 | 50 SR  |
| Mature male medusa   | M1                       | Z10       | 1              | 1370           | NextSeq    | 75 SR  |
| Mature male medusa   | M2                       | Z10       | 1              | 2370           | NextSeq    | 75 SR  |
| Mixed-stage          | See Supplementary Tab. 8 |           |                |                | HiSeq 2500 | 100 PE |
| Vienna Medusa        | See Material and Methods |           |                |                | HiSeq 2000 | 100 PE |

**Supplementary table 7. Mixed-stage RNA preparation.**

| Stage                    | Strain    | individuals | Total RNA (ng) |
|--------------------------|-----------|-------------|----------------|
| Planula P1               | Z4B * Z10 | 150         | 250            |
| Planula P2               | Z4B * Z10 | 150         | 250            |
| Planula P3               | Z4B * Z10 | 150         | 250            |
| Primary Polyp PoPr       | Z4B * Z10 | 20          | 250            |
| Gastrozoid/Polyp head PH | Z4B       | 60          | 250            |
| Gonozooid GO             | Z4B       | 60          | 250            |
| Stolon St                | Z4B       | 60          | 250            |
| Mature Female Medusa MMF | Z4B       | 1           | 250            |
| Mature Male Medusa M     | Z10       | 1           | 250            |
| <b>MIX*</b>              |           |             | <b>2250</b>    |

\* 250ng RNA extracted for each stage were combined to give a total of 2250ng RNA, sequenced by GATC.

**Supplementary table 8. Sequences generated for RNA libraries.**

| Stage                 |      | Replicate 1 | % mapped | Replicate 2 | % mapped |
|-----------------------|------|-------------|----------|-------------|----------|
| Early Gastrula        | EG   | 39331160    | 86.4     | 43845659    | 84.4     |
| One day old Planula   | P1   | 38554115    | 84.8     | 42356552    | 85.7     |
| Two day old Planula   | P2   | 55900254    | 82.9     | 48155906    | 83.0     |
| Three day old Planula | P3   | 47246549    | 82.1     | 70793282    | 82.8     |
| Stolon                | St   | 80487282    | 80.0     | 90277892    | 71.8     |
| Primary Polyp         | PoPr | 72275162    | 84.2     | 70676083    | 82.1     |
| Polyp Head            | PH   | 45393098    | 80.9     | 45627654    | 82.0     |
| Gonozooid             | GO   | 65897403    | 79.8     | 63831768    | 83.7     |
| Baby Medusa Female    | BMF  | 51835910    | 83.3     | 35779461    | 83.2     |
| Mature Medusa Female  | MMF  | 44120673    | 83.0     | 49731226    | 83.1     |
| Male Medusa           | M    | 45793597    | 84.8     | 50284626    | 85.3     |
| Mixed-stage           | -    | 143397155   | 68.5     | -           | -        |
| Vienna Medusa         | -    | 106245738   | 76.5     | -           | -        |

‘% mapped’ refers to uniquely mapped reads.

**Supplementary table 9. Genomic DNA sequencing libraries.**

| Library insert size | Technology       | Read length | No. of reads |
|---------------------|------------------|-------------|--------------|
| 180                 | Illumina PE      | 101         | 688126703    |
| 500                 | Illumina PE      | 101         | 203497401    |
| 4000                | Illumina Nextera | 101         | 136386432    |
| 8000                | 454              | -           | 11802050     |
| 20000               | 454              | -           | 9505979      |
| 40000               | 454              | -           | 502197       |

**Supplementary table 10. Protein datasets used in this study.**

|                                      | download source                                                                                     | GFF                                                | proteins                                           |
|--------------------------------------|-----------------------------------------------------------------------------------------------------|----------------------------------------------------|----------------------------------------------------|
| <i>Saccharomyces cerevisiae</i>      | www.yeastgenome.org                                                                                 | -                                                  | cerpep (13-Jan-2015)                               |
| <i>Capsaspora owczazarki</i>         | BROAD origins of multicellularity*                                                                  | -                                                  | salpingoeca_rosetta_1_proteins.fasta               |
| <i>Monosiga brevicollis</i>          | BROAD origins of multicellularity*                                                                  | -                                                  | monosiga_brevicollis_mx1_1_proteins.fasta          |
| <i>Mnemiopsis leidyi</i>             | <a href="https://research.nhgri.nih.gov/mnemiopsis/">https://research.nhgri.nih.gov/mnemiopsis/</a> | -                                                  | ML2.2.aa                                           |
| <i>Nematostella vectensis</i>        | www.cnidariangenomes.org                                                                            | nveGenes.vienna130208.nemVec1.gtf                  | nveGenes.vienna130208.protein.fasta                |
| <i>Amphimedon queenslandica</i>      | NCBI                                                                                                | GCF_000090795.1_v1.0_genomic.gff                   | GCF_000090795.1_v1.0_protein.faa                   |
| <i>Trichoplax adhaerens</i>          | NCBI                                                                                                | GCF_000150275.1_v1.0_genomic.gff                   | GCF_000150275.1_v1.0_protein.faa                   |
| <i>Hydra vulgaris</i>                | NCBI                                                                                                | ref_Hydra_RP_1.0_top_level.gff3                    | 102**                                              |
| <i>Exaiptasia</i>                    | NCBI                                                                                                | ref_Aiptasia_genome_1.1_top_level.gff3             | 100**                                              |
| <i>Orbicella faveolata</i>           | NCBI                                                                                                | ref_ofav_dov_v1_top_level.gff3                     | 100**                                              |
| <i>Acropora digitifera</i>           | NCBI                                                                                                | GCF_000222465.1_Adig_1.1_genomic.gff               | GCF_000222465.1_Adig_1.1_protein.faa               |
| <i>Strongylocentrotus purpuratus</i> | NCBI                                                                                                | ref_Spur_4.2_top_level.gff3                        | GCF_000222465.1_Adig_1.1_protein.faa               |
| <i>Saccoglossus kowalevskii</i>      | NCBI                                                                                                | ref_Skow_1.1_top_level.gff3                        | 101**                                              |
| <i>Homo sapiens</i>                  | NCBI                                                                                                | ref_GRCh38.p2_top_level.gff3                       | 107**                                              |
| <i>Ciona intestinalis</i>            | NCBI                                                                                                | ref_KH_top_level.gff3                              | 102**                                              |
| <i>Branchiostoma floridae</i>        | NCBI                                                                                                | GCF_000003815.1_Version_2_genomic.gff              | GCF_000003815.1_Version_2_protein.faa              |
| <i>Drosophila melanogaster</i>       | NCBI                                                                                                | GCF_000001215.4_Release_6_plus_ISO1_MT_genomic.gff | GCF_000001215.4_Release_6_plus_ISO1_MT_protein.faa |
| <i>Caenorhabditis elegans</i>        | NCBI                                                                                                | GCF_000002985.6_WBcel235_genomic.gff               | GCF_000002985.6_WBcel235_protein.faa               |
| <i>Capitella teleta</i>              | NCBI                                                                                                | GCA_000328365.1_Capca1_genomic.gff                 | GCA_000328365.1_Capca1_protein.faa                 |
| <i>Lingula anatina</i>               | NCBI                                                                                                | ref_LinAna1.0_top_level.gff3                       | 100**                                              |
| <i>Crassostrea gigas</i>             | NCBI                                                                                                | ref_oyster_v9_top_level.gff3                       | 100**                                              |
| <i>Aplysia californica</i>           | NCBI                                                                                                | ref_AplCal3.0_top_level.gff3                       | 101**                                              |

\* The original Broad institute ‘Origins of multicellularity’ website is no longer accessible - data downloads are redirected to the NCBI.

\*\* NCBI Annotation Release number.

**Supplementary table 11. Primer sequences and EST clones used for generating the matrix for the 14 genes studied by in situ hybridization.**

| Gene          | ID          | ISH Probe matrix          | Primer sequences/EST clones                                                                                                                                  |
|---------------|-------------|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Cdx</b>    | XLOC_040366 | EST                       | SA0AAB185YO17                                                                                                                                                |
| <b>CnoxA</b>  | XLOC_038052 | pGEM-T easy clone         | CnoxA_F_PCR1: TGTCGACCAGCAGTAAACCA<br>CnoxA_F_PCR2: GGCACAATCAACCGTCGAAA<br>CnoxA_R_PCR1: GAAAGTAGGTGGGTGGGGGAA<br>CnoxA_R_PCR2: TAGGGCTGGTAAGGGTGGTA        |
| <b>DRGX</b>   | XLOC_000977 | EST                       | SA0AAB15YB07                                                                                                                                                 |
| <b>FezF</b>   | XLOC_034937 | PCR including T7 promoter | FezF_F: CTCACCTACCTATCCGAAACAAGC<br>Fez_R_T7: ACCAATGATCACTGGTGCCTTCGTTTTCTAATACGACTCACTATAGG                                                                |
| <b>FoxD</b>   | XLOC_015631 | PCR including T7 promoter | FoxD_F: GCTTGTTCCCATTTAGCCG<br>FoxD_R GAAGCAGGATCGAGTGTCCA                                                                                                   |
| <b>Msx</b>    | XLOC_035505 | PCR including T7 promoter | Msx_F: GACCTTCAAGAACGAGAATTCCC<br>Msx_R_T7: GAACCTTACATGAGGCCACCTTACGTTTTCTAATACGACTCACTATAGG                                                                |
| <b>Neurog</b> | XLOC_009631 | pGEM-T easy clone         | Neurog_F_PCR1: AGGAAGTGGAACCGTAACGA<br>Neurog_F_PCR2: AGCAGGGATAAACAAGGCAAC<br>Neurog_R_PCR1: TTACCGTCCGTCATAGCAGC<br>Neurog_F_PCR2: TCCATTCACGCTATTATGGTCT  |
| <b>Nkx2</b>   | XLOC_009058 | EST                       | SA0AAB134YL05                                                                                                                                                |
| <b>Nkx6</b>   | 23_g670.t1  | pGEM-T easy clone         | Nkx6_F_PCR1: AGATCCGAAAACAGACAAATCG<br>Nkx6_F_PCR2: AAAGCGTTCACCATTTCTCGG<br>Nkx6_R_PCR1: TTTTCCGCCATTTTGTCCGA<br>Nkx6_R_PCR2: CTCTTTCAGGACCGGCTAAA          |
| <b>OtxA</b>   | XLOC_000267 | EST                       | SA0AAB124YI16                                                                                                                                                |
| <b>OtxB</b>   | XLOC_040405 | pGEM-T easy clone         | OtxB_F: CATCACACCTTCCAAAGCAACGAT<br>OtxB_F: AAGAACCAGATCTTCACGCAACAT                                                                                         |
| <b>OtxC</b>   | XLOC_039817 | PCR including T7 promoter | OtxC_F: GATAAACTCTCGGGAAACACCAG<br>OtxC_F_T7: TTTACCTGTACTCGAGCTTCAGCCGTTTTCTAATACGACTCACTATAGG                                                              |
| <b>Six3/6</b> | XLOC_010764 | PCR including T7 promoter | Six3/6_F: CGATTATTCCTAACGACGACGACG<br>Six3/6_R_T7: GGTGTCAAGATGGGTACAGACTTGGTTTTCTAATACGACTCACTATAGG                                                         |
| <b>Twistb</b> | XLOC_038912 | pGEM-T easy clone         | Twistb_F_PCR1: TCTGCTCTTTCACTTTCACTGT<br>Twistb_R_PCR2: TCCTGTTTTGTTATGGGCGG<br>Twistb_R_PCR1: AGACTGACTAAAGCCTGGAGA<br>Twistb_R_PCR2: TGGTGGAATCGGAGCATATCA |