Comprehensive phylogenomic analyses resolve cnidarian relationships and the origins of key organismal traits

Ehsan Kayal[1,2], Bastian Bentlage[1,3], M. Sabrina Pankey[5], Aki H. Ohdera[4], Monica Medina[4], David C. Plachetzki[5]*, Allen G. Collins[1,6], Joseph F. Ryan[7,8]*

Authors Institutions:
1. Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution
2. UPMC, CNRS, FR2424, ABiMS, Station Biologique, 29680 Roscoff, France
3. Marine Laboratory, university of Guam, UOG Station, Mangilao, GU 96923, USA
4. Department of Biology, Pennsylvania State University, University Park, PA, USA
5. Department of Molecular, Cellular and Biomedical Sciences, University of New Hampshire, Durham, NH, USA
6. National Systematics Laboratory, NOAA Fisheries, National Museum of Natural History, Smithsonian Institution
7. Whitney Laboratory for Marine Bioscience, University of Florida, St Augustine, FL, USA
8. Department of Biology, University of Florida, Gainesville, FL, USA

Abstract

Background: The phylogeny of Cnidaria has been a source of debate for decades, during which nearly all-possible relationships among the major lineages have been proposed. The ecological success of Cnidaria is predicated on several fascinating organismal innovations including symbiosis, colonial body plans and elaborate life histories, however, understanding the origins and subsequent diversification of these traits remains difficult due to persistent uncertainty surrounding the evolutionary relationships within Cnidaria. While recent phylogenomic studies have advanced our knowledge of the cnidarian tree of life, no analysis to date has included genome scale data for each major cnidarian lineage.

Results: Here we describe a well-supported hypothesis for cnidarian phylogeny based on phylogenomic analyses of new and existing genome scale data that includes representatives of all cnidarian classes. Our results are robust to alternative modes of phylogenetic estimation and phylogenomic dataset construction. We show that two popular phylogenomic matrix construction pipelines yield profoundly different datasets, both in the identities and the functional classes of the loci they include, but resolve the same topology. We then leverage our phylogenetic resolution of Cnidaria to understand the character histories of several critical organismal traits. Ancestral state reconstruction analyses based on our phylogeny establish several notable organismal transitions in the evolutionary history of Cnidaria and depict the ancestral cnidarian as a solitary, non-symbiotic polyp that lacked a medusa stage. In addition, Bayes factor tests of multiple origins strongly suggest that symbiosis has evolved multiple times independently across the cnidarian radiation.

Conclusions: Cnidaria have experienced more than 600 million years of independent evolution and in the process generated an array of organismal innovations. Our results add significant clarification on the cnidarian tree of life and the histories of these innovations. Further, we confirm the existence of Acraspeda (staurozoans plus scyphozoans and cubozoans), thus reviving an evolutionary hypothesis put forward more than a century ago.

Background

Cnidaria is a diverse phylum of mostly marine species comprised of three major clades: Anthozoa, Endocnidozoa and Medusozoa (Collins 2009). Anthozoa encompasses more than half (7,200 of 13,300) of the known cnidarian species and consists of Octocorallia (sea pens, sea fans, and soft corals), Hexacorallia (stony corals, black corals, sea anemones, zoantharians and corallimorpharians), and Ceriantharia (tube anemones). Endocnidozoa is an entirely parasitic clade that includes about 2,200 species of Myxozoa (minute endoparasites of invertebrates and vertebrates with complex life cycles) and the monotypic Polypodiozoa (a parasite that infects the eggs of sturgeon and paddlefish). It was not until after a long line of evidence that it became clear that Myxozoa was a clade within Cnidaria (Foox and Siddall 2015). Finally, Medusozoa consists of Cubozoa (45 species of box jellyfish), Hydrozoa (3,600 species of hydroids, siphonophores and hydromedusae), Scyphozoa (200 conspicuous species of true jellyfishes), and Staurozoa (50 species of benthic stalked jellyfishes). The ecological success of Cnidaria is predicated on several fascinating organismal innovations including relationships with phototrophic endosymbiotic eukaryotes, colonial body plans, and the metagenetic life cycle that includes medusa (jellyfish) and polyp stages. However, understanding the origins and subsequent diversification of these critical innovations remains difficult due to persistent uncertainty surrounding the evolutionary relationships within Cnidaria.

Cnidarian phylogeny has been a source of debate for decades, with nearly every possible sister group relationship proposed among the major lineages of Medusozoa (Figure 1) and Anthozoa (Figure 2) (Collins 2009; Collins, et al. 2006b; Daly, et al. 2007; Stampar, et al. 2014). Whole mitochondrial phylogenomic analyses have supported paraphyletic Anthozoa and Scyphozoa (Kayal, et al. 2013; Park, et al. 2012), but subsequent work suggested that these findings were due to saturation bias (Pratlong, et al. 2017). More recent phylogenomic studies have supported the monophyly of Anthozoa and Scyphozoa (Zapata, et al. 2015) and placed Endocnidozoa as the sister group to Medusozoa (Chang, et al. 2015). However, these phylogenomic studies lacked several key taxa. For instance, Chang et al. (2015) did not include data from Staurozoa, Ceriantharia, or Coronatae (Scyphozoa), while Zapata et al. (2015) lacked data from Endocnidozoa and Rhizostomeae (Scyphozoa). In addition, data representation was sparse for Ceriantharia and Staurozoa in Zapata et al. (2015) with weak support for the positions of both taxa represented by single exemplar species. Nevertheless, the topologies from these two independent phylogenomic studies were otherwise largely congruent, providing some prospect that large datasets and increased taxon sampling may settle long-standing questions about the evolutionary history of Cnidaria.

Staurozoa (Figure 1) is one of the most enigmatic and understudied cnidarian clades. These animals have unique life history attributes, including non-ciliated, creeping, larval planulae that develop into polyps, which later undergo metamorphosis (unique from strobilation or budding) into adult stauromedusae (Kikinger and Vonsalviniplawen 1995; Miranda, et al. 2013; Wietrzykowski 1912). In addition, the adult staurozoan body plan includes features common to both the polyp and medusa stages of other cnidarians, thus adding to their phylogenetic significance (Collins 2002, 2010; Miranda, et al. 2016a; Miranda, et al. 2013; Miranda, et al. 2016b). The phylogenetic position of staurozoans relative to other medusozoan lineages remains controversial and recent hypotheses (Kayal, et al. 2013; Simion, et al. 2017; Zapata, et al. 2015) present different implications for the evolution of medusozoan body plans (Figure 1F). Analyses of rDNA, life history and anatomical features have suggested that Staurozoa is the sister lineage to the remainder of Medusozoa (Collins 2002; Van Iten 1992). However, recent phylogenomic analyses (Zapata, et al. 2015) placed Staurozoa in a clade with Cubozoa and Scyphozoa (Figure 1F middle). Support for the placement of Staurozoa has been characteristically low in these prior studies, leaving open the question of their phylogenetic position within Cnidaria. Critically, prior to the present study, relatively little evidence has been
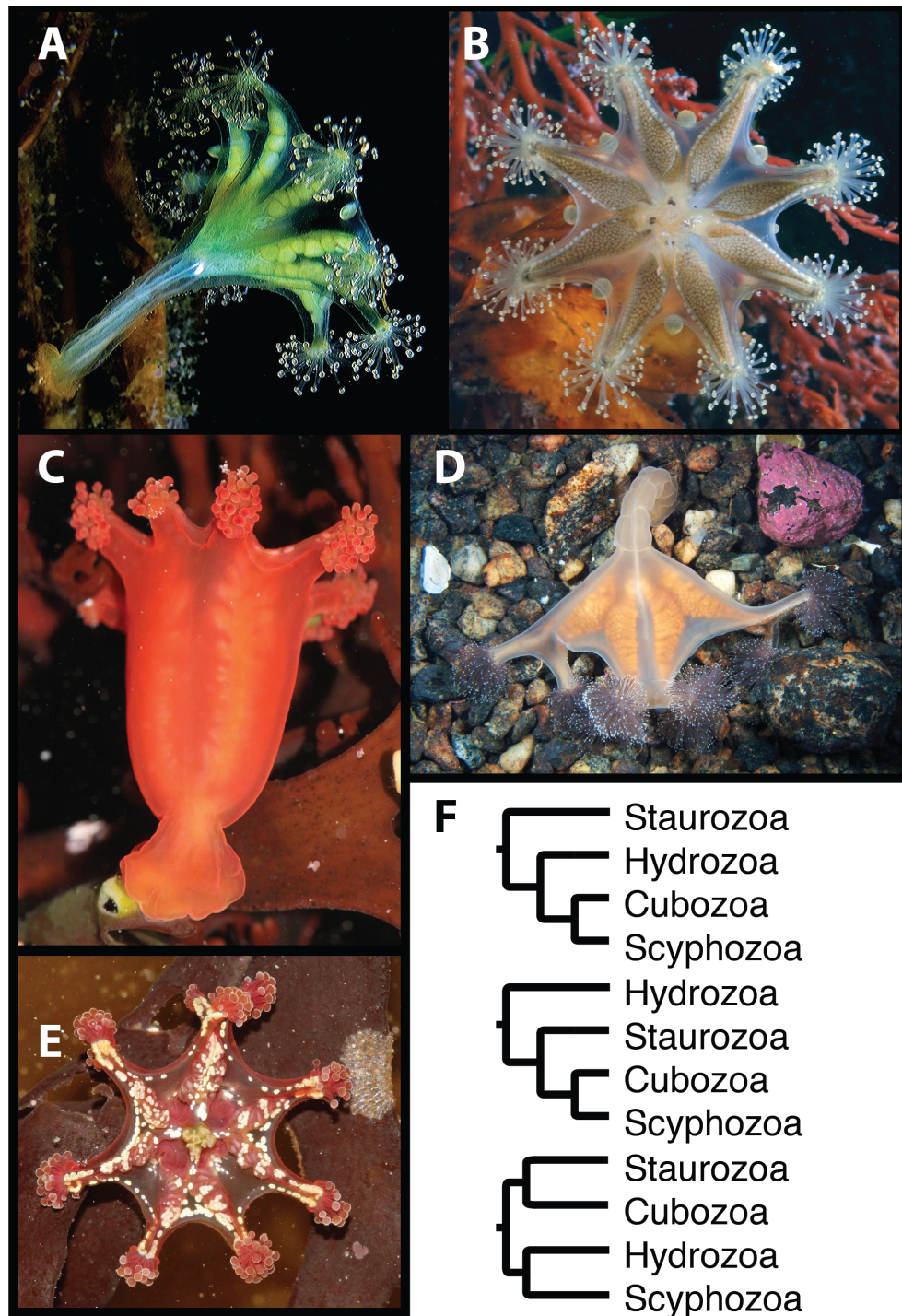
Figure 1. Exemplar staurozoans and competing phylogenetic hypotheses for the position of Stuarozoa within Medusozoa. **A.** *Haliclystus auricula*. Credit: Marco Faasse; cropped. **B.** *Calvadosia cruxmelitensis* Credit: Stauromedusae UK & David Fenwick; cropped. **C.** *Haliclystus* "sanjuanensis". Credit: Ron J. Larson. **D.** *Craterolophus convolvulus*. Credit: Stauromedusae UK & David Fenwick; cropped. **E.** *Lucernaria quadricornis*. Credit: Alexander Semenov. **F.** Competing hypotheses for the phylogenetic position of Staurozoa within Anthozoa. Top from Collins et al. (2006) and Van Iten et al. (2006); bottom from Zapata et al. (2015).

brought to bear on the phylogenetic position of Staurozoa, whether it be from morphology (Hyman 1940; Kiderlen 1937; Marques and Collins 2004; Van Iten 1992; Van Iten, et al. 2006), a small number of genes (Collins 2002; Collins, et al. 2006a), or sparse representation (in terms of both data and taxon sampling) in an otherwise, large phylogenomic study (Zapata, et al. 2015).

Ceriantharia is another cnidarian clade whose uncertain phylogenetic placement has major evolutionary implications (Figure 2). Ceriantharians are burrowing sea anemones with larval and adult stages that are morphologically distinct from the other anthozoan lineages (i.e., Octocorallia and Hexacorallia) (Daly, et al. 2007; Stampar, et al. 2014). Ceriantharians differ from other anthozoans in that they possess a distinct planktonic larval stage known as the cerinula (Molodtsova 2003), a secreted composite tube into which they retract when disturbed, a novel type of structural cnidocyte called a ptychocyte that provides support for their tube dwellings (Mariscal, et al. 1977) and two whorls of tentacles that surround the oral opening. Ceriantharia has been placed in several conflicting phylogenetic positions (Chen, et al. 1995; Song 1997) including in a recent study (Stampar, et al. 2014) as the sister lineage to the two main anthozoan lineages Hexacorallia and Octocorallia (Figure 2F bottom). More commonly, Ceriantharia has been recovered as the sister to Hexacorallia, with which they share spirocytes, a common nematocyst type absent from octocorals (Figure 2F top) (Berntson, et al. 1999; Daly, et al. 2003; France, et al. 1996; Kayal, et al. 2013; Zapata, et al. 2015). As with Staurozoa, all previous analyses of the phylogenetic position of Ceriantharia have been based on morphology or limited sequence data, and support for its position has been consistently low (Stampar, et al. 2014; Zapata, et al. 2015).

Phylogenomic analyses of genome-scale datasets (i.e., whole-genome-derived gene models or RNAseq-derived transcripts) have recently been exploited to resolve a host of longstanding phylogenetic issues (Dunn, et al. 2008; Hampl, et al. 2009; Qiu, et al. 2006). A critical step common to these analyses is the identification of one-to-one orthologs from genome-scale datasets for each taxon, which are then used as data partitions in large super-matrices. Various methods are available for the identification of such data partitions and methodological differences among them have been shown to impact phylogenetic inference (Borowiec, et al. 2015; Shen, et al. 2017; Tekaia 2016). Yet, most phylogenomic analyses rely on a single method for data matrix construction and do not examine the impact of alternative approaches on phylogenetic reconstruction.

Here, we apply extensive new phylogenomic data for Staurozoa, Ceriantharia and other under-sampled cnidarian clades to the construction and analyses of independent phylogenomic datasets for Cnidaria using two popular approaches: *1)* Agalma (Dunn, et al. 2013) and *2)* a custom pipeline based on Orthofinder (Emms and Kelly 2015) and PhyloTreePruner (Kocot, et al. 2013). We show that both procedures produce datasets with surprisingly little overlap in terms of data partitioned, but resolve the same topology under robust phylogenetic methods. We then leverage our highly resolved cnidarian phylogeny to address questions surrounding the origins and evolutionary histories of several key organismal innovations in Cnidaria. Our character mapping studies, based on explicit statistical models, identify key evolutionary transitions within Cnidaria and strongly suggest that the ancestral cnidarian was a solitary polyp that lacked a medusa stage or a photosynthetic endosymbiont. Further, our analyses strongly suggest that symbiosis with photosynthetic eukaryotes has evolved on multiple occasions in Cnidaria.

Results

We generated transcriptomic data from five staurozoans (*Calvadosia cruxmelitensis*, *Craterolophus convolvulus*, *Haliclystus auricula*, *Haliclystus "sanjuanensis"*, and *Lucernaria quadricornis*), one scyphozoan *Cassiopea xamachana*, and the cerianthid *Cerianthus borealis*.
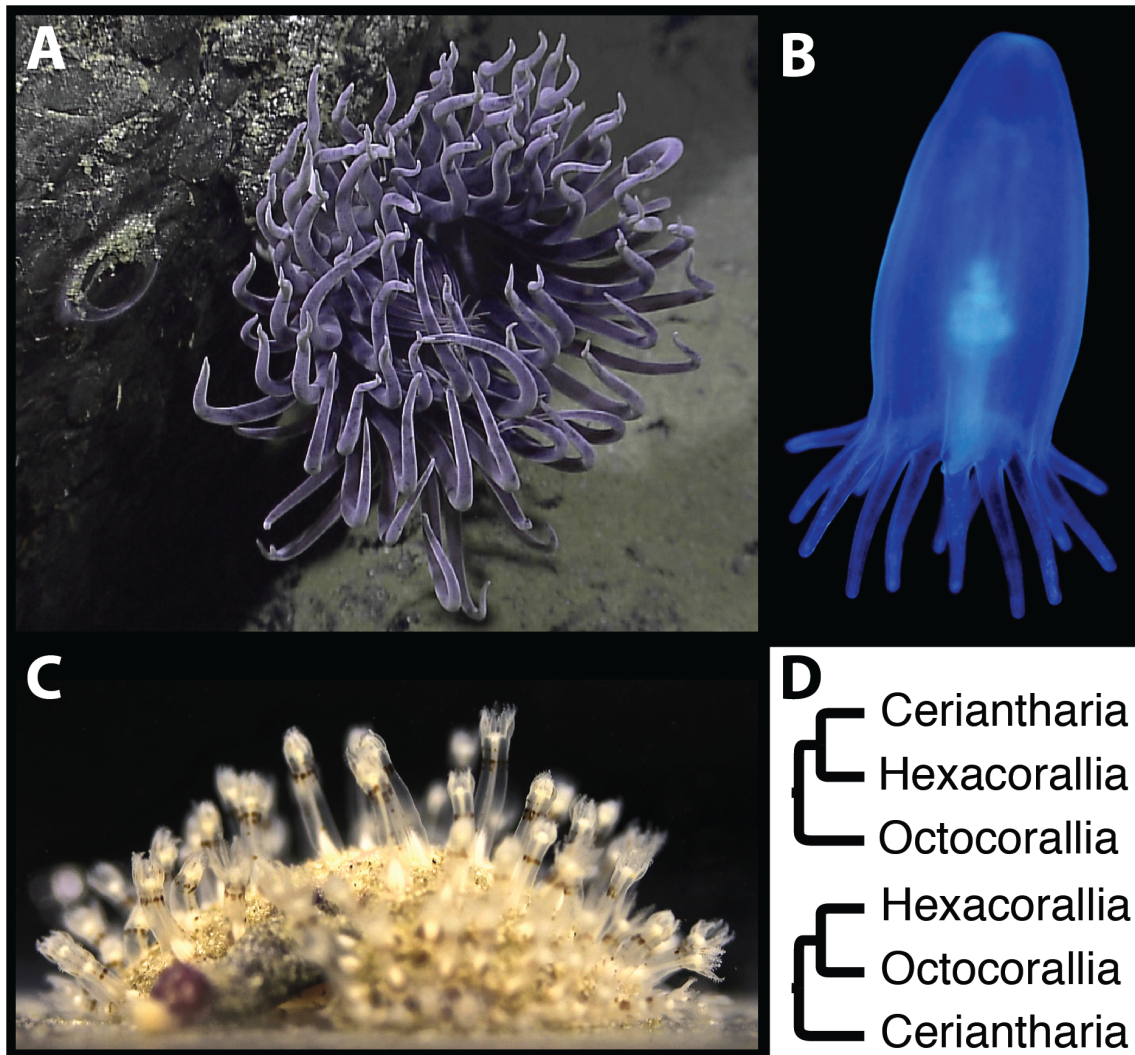
Figure 2. Exemplar ceriantharians and octocorals and competing hypotheses for the position of Ceriantharia within Anthozoa. **A.** Ceriantharian tube anemone adult. Credit: NOAA Okeanos Explorer Program, Galapagos Rift Expedition, 2011; cropped. **B.** Ceriantharian tube anemone pelagic larva. Credit: A. G. Collins, Bonaire, 2016. **C.** Octocorallian, the sea pansy, *Renilla reniformis*. Credit: J. Ryan. **D.** Competing hypotheses for the phylogenetic position of Ceriantharia within Anthozoa. Top from Daly et al. (2003) and Zapata et al. (2015); bottom from Stampar et al. (2014).

In addition, we sequenced and generated a rough-draft assembly of the nuclear genome of *Renilla reniformis*. The genome assembly had an N50 of 1,843 base pairs. We predicted 12,689 protein-coding genes, many of which are likely partial, but sufficient for downstream phylogenomic analyses. We also used the highest quality transcriptomic data from Zapata et al. (2015), to which we added genomic and transcriptomic data from several taxa that were underrepresented in previous studies, including most Endocnidozoa taxa from Chang et al. (2015). After an initial round of matrix construction and phylogenetic analyses, several new cnidarian transcriptome datasets became available, and we incorporated an additional 13 taxa into our final data matrix OF-PTP_75tx (Supplementary Tables S1 and S2).

Potential contamination identified in cnidarian transcriptome data

We applied a strict filter to all datasets to remove potential contaminants. In total, we removed less than 5% of sequences from most datasets except for the following taxa: *Alatina alata* (7.9%), *Anemonia viridis* (6%), *Anthopleura elegantissima* (7%), *Gorgonia ventalina* (6.8%), *Hydractinia polyclina* (6.8%), *Platygyra carnosus* (6.7%), and *Seriatopora hystrix* (6.9%). In addition, many sequences from the myxozoans *Kudoa iwatai* (39.8%), *Myxobolus cerebralis* (25.6%), *M. pendula* (40.5%), and *Thelohanellus kitauei* (21.4%), as well as the filiferan hydrozoan *Podocoryna carnea* (26.7%) had best matches to bilaterian sequences and were subsequently removed (Supplementary Table 1). Following the removal of these putative contaminants, preliminary phylogenetic analyses showed that the myxozoan sequences procured by the Agalma pipeline still contained many contaminants, as these species were resolved within the vertebrates. Further analysis of myxozoan-bearing partitions from both datasets showed that the Agalma pipeline was prone to include partitions with a single myxozoan species present and that these partitions were more likely to be comprised of contaminants (Figure 3), a situation not encountered in the OF-PTP dataset. Further, when Agalma partitions with greater than three myxozoan species were selected for phylogenetic analysis (47 partitions), the myxozoan species were resolved in their expected position within the Endocnidozoa, and the remaining topology was consistent with all other results (Supplementary Figure S2).

Phylogenomic matrix generation pipelines produced contrasting data matrices

We built two preliminary, independent phylogenomic data matrices with Agalma (AG_62tx), and OF-PTP (OF-PTP_62tx). After selecting orthologous partition alignments that exceeded 50% taxon occupancy, the Agalma pipeline incorporated roughly three times as many genes and four times as much data (962 single-gene partitions, 233,568 data positions) as OF-PTP (352 single-gene partitions, 53,389 data positions) (Figure 4). Furthermore, the average partition length was longer for the AG_62tx dataset than OF-PTP_62tx (Figure 4A). Comparisons of the across-partition *N. vectensis* complements of AG_62tx and OF-PTP_62tx revealed that only 52 loci are shared between the two datasets (Figure 4C).

We also detected significant differences in taxon occupancy between the datasets produced by Agalma and OF-PTP. AG_62tx had, on average, significantly lower taxon occupancy (481/972 or 49%) compared to OF-PTP_62tx (232/372 or 62%; Fisher's Exact Test *P*=0.022) or OF-PTP75tx. In addition, the AG_62tx dataset had exceedingly low coverage for several key groups (see below and also Supplementary Figures S1, S3 and S4). Similarly, the OF-PTP datasets were significantly denser, but smaller in size (Figure 4A-B) than the dataset reported by Zapata et al. (2015), which was also constructed using the Agalma pipeline (Dunn, et al. 2013) (Figure 4B, Supplementary Figures S3 and S4). The data matrix reported by Chang et al. (2015), a manually curated dataset, had the highest density, but the lowest number of ingroup taxa present (Figure 4B). Overall, our findings suggested significant differences
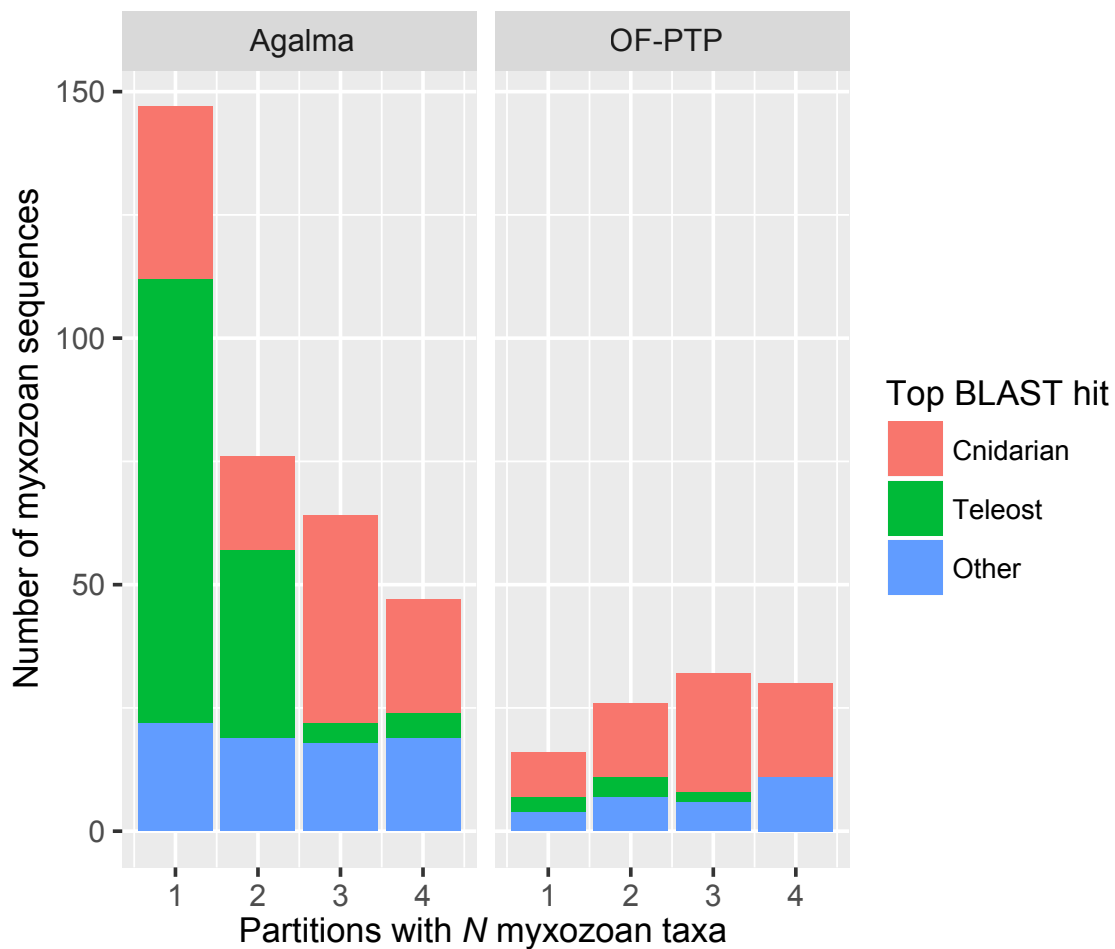
Figure 3. Relationship between sparse data representation and the likelihood of retaining contaminated sequences in phylogenomic data matrices as illustrated by myxozoan species. We conducted BLAST similarity searches against a metazoan genome database for all myxozoan sequences present in both the AG_62tx and OF-PTP_62tx matrices. In addition, we noted how many myxozoan species were present in each partition. Myxozoans are internal parasites of teleost fishes and we noted significant contamination in transcriptome data from these host species. The Agalma pipeline produces a large, but sparse matrix as compared to OF-PTP. In cases where contamination is common, as with myxozoan data, sparse data matrices have high numbers of partitions with single species represented per clade, which in turn are enriched for contaminant sequences. Partitions with greater than one species of myxozoan present have a lower potential to include contamination. The OF-PTP pipeline produces a denser data matrix, which makes it inherently less prone to selecting contaminants.
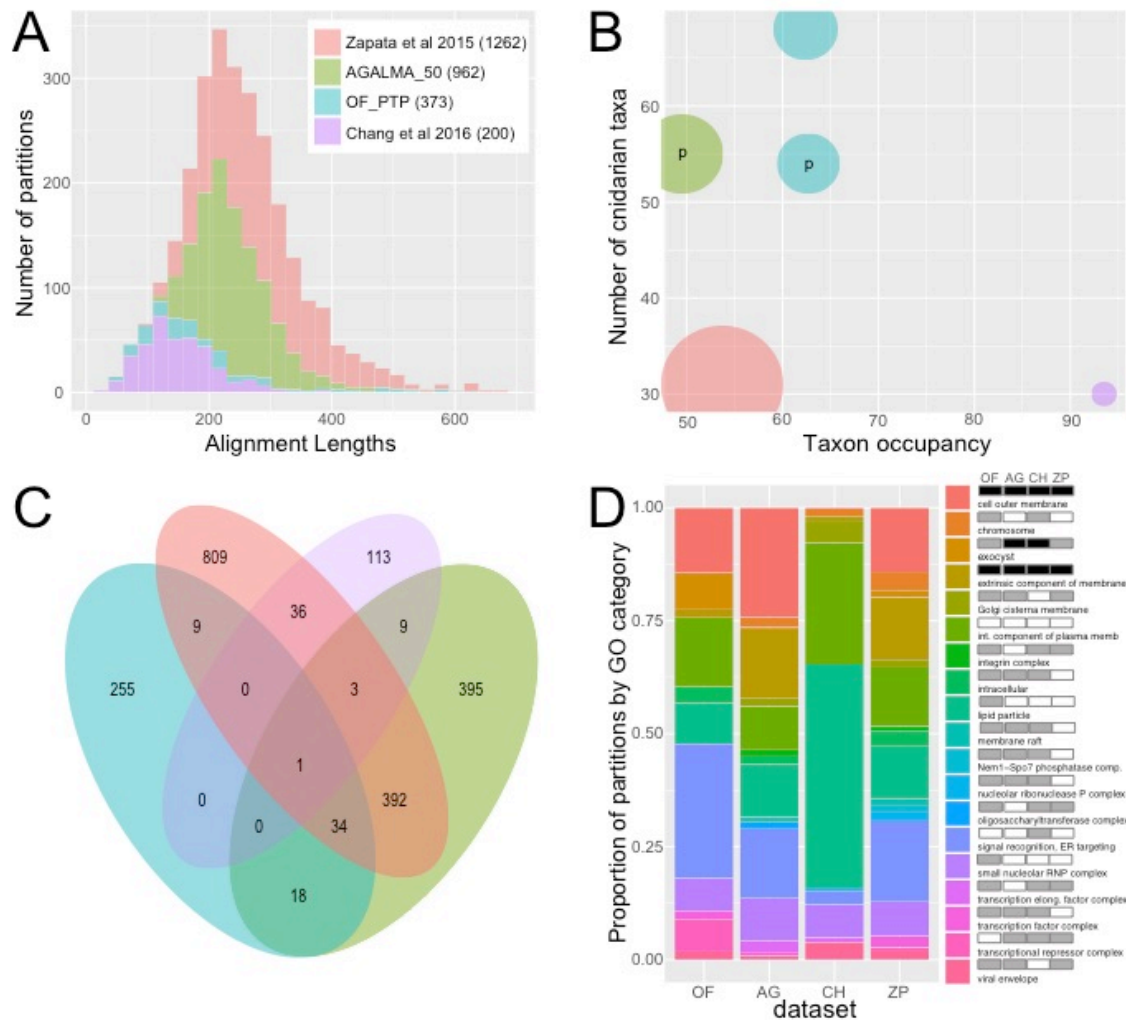
Figure 4. Comparisons of new and previously published phylogenomic datasets for Cnidaria reveal significant differences. **A.** Histograms showing the distributions of alignment lengths for partitions included in the OF-PTP_62tx, AG_62tx and two previously published phylogenomic matrices (Chang, et al. 2015; Zapata, et al. 2015). **B.** The number of taxa plotted against the average taxon occupancy for each dataset. Size of each circle is based on number of partitions (see inset in **A**). **C.** Venn diagram indicating the overlap in *N. vectensis* gene identities from partitions present in each dataset. The most similar datasets are AG_62tx and that from Zapata et al. (2015), which are both based on Agalma (Dunn, et al. 2013). **D.** Composition, enrichment and depletion of GO terms associated with the cellular component category from the *N. vectensis* sequences present across partitions. Left, the composition of unique cellular component terms are shown as bar plots for each dataset. Only terms that are significantly enriched or depleted relative to their frequencies in the *N. vectensis* protein set in at least one dataset are shown. Right, the identities of each cellular component term and their enrichment or depletion for each dataset. Black = depleted. White = enriched. Grey = not significant. For **A-C**, datasets are color coded as in **A**.

between the composition of datasets produced by OF-PTP, Agalma and those reported in recent phylogenomic analyses of Cnidaria (Chang, et al. 2015; Zapata, et al. 2015).

Next, we sought to understand how these datasets differed in terms of the functional classes of genes present. We first compared the number of *N. vectensis* genes shared in each dataset and found surprisingly low levels of overlap in each set (Figure 4C). This estimation could be confounded if different pipelines retained different, but closely related, *N. vectensis* paralogs during their distinct tree pruning procedures, potentially exaggerating differences between them. We therefore conducted gene ontology (GO) analyses of enrichment and depletion of GO terms present in each dataset by comparing the relative proportions of each GO term for the *N. vectensis* genes from each dataset to their relative proportion in the background *N. vectensis* v1.0 protein set (Putnam, et al. 2007). These analyses show that significant differences in GO term representation, relative to the background, pervade each dataset across each GO category examined (e.g. cellular component, molecular function and biological process) and are evident by comparing the composition and relative enrichment and depletion of GO terms between each of the data matrices (Figure 4D, Supplementary Figure S5). Our findings demonstrate that the two independent datasets produced here, together with the two most recent phylogenomic analyses of Cnidaria (Chang, et al. 2015; Zapata, et al. 2015), are comprised of data partitions that differ profoundly in terms of gene identity and functional class (Figure 4C-D).

Consistent phylogenomic results from different data matrices

Preliminary phylogenetic analyses of the AG_62tx and OF-PTP_62tx matrices were largely congruent in our ML analyses, with the exception that the Myxozoa was unexpectedly positioned within the vertebrates in our analyses of AG_62tx, likely due to contamination (Figure 3, Supplementary Figure S1). The unexpected position of Myxozoa in the AG_62tx analysis corresponds to data sparseness. On average, myxozoan species are represented in a significantly smaller proportion (Fisher's Exact Test) of partitions in the AG_62tx dataset as compared to the OF-PTP_62tx dataset: *Myxobolus pendula P* = 0.027; *Thelohanellus kitauei P* = 0.0001; *Myxobolus cerebralis P* = 0.0001 (Figure 5, Supplementary Figure S1). We hypothesized that if contaminants were still present in the myxozoan datasets despite our filtering efforts, they would be minimized in partitions that had more than one myxozoan species present, as the likelihood of selecting more than one orthologous contaminant sequences from multiple datasets for the same partition would be low. In fact, we recovered a monophyletic Cnidaria with the AG_62x matrix after pruning myxozoan partitions with fewer than three myxozoan species (Supplementary Figure S2). The sparse representation of myxozoan sequences across AG_62tx makes it more likely that a single myxozoan species is represented per partition, as compared to the less sparse OF-PTP matrix (Figure 3). This relationship, in turn, makes it more likely that contaminants are incorporated into the Agalma matrix. Given the sparse nature of the matrix produced by Agalma, we decided to focus our deeper analyses on the matrix derived from the OF-PTP approach.

After our initial analyses, we added new data for thirteen additional taxa derived from our own sequencing efforts and from newly published studies, for a total of 67 cnidarian taxa (75 species in total, Supplementary Tables S1 and S2) and produced a new dataset containing 372 partitions (53,539 positions) which we refer to as OF-PTP_75tx. We performed both ML and Bayesian analyses on OF-PTP_75tx (Figure 5, Supplementary Figures S6 and S7). After pruning *T. adherens* from the trees sampled from both chains, we obtained a convergence with high confidence (maxdiff = 0.152284). In addition, *Renilla reniformis*, which had among the most limited data representation of any taxon, formed a monophyletic clade with *Corallium rubrum* in the Bayesian analyses, but was the sister to the remaining octocorals in ML analyses. We therefore collapsed the two conflicting nodes into polytomies (Figure 5, Supplementary Figures
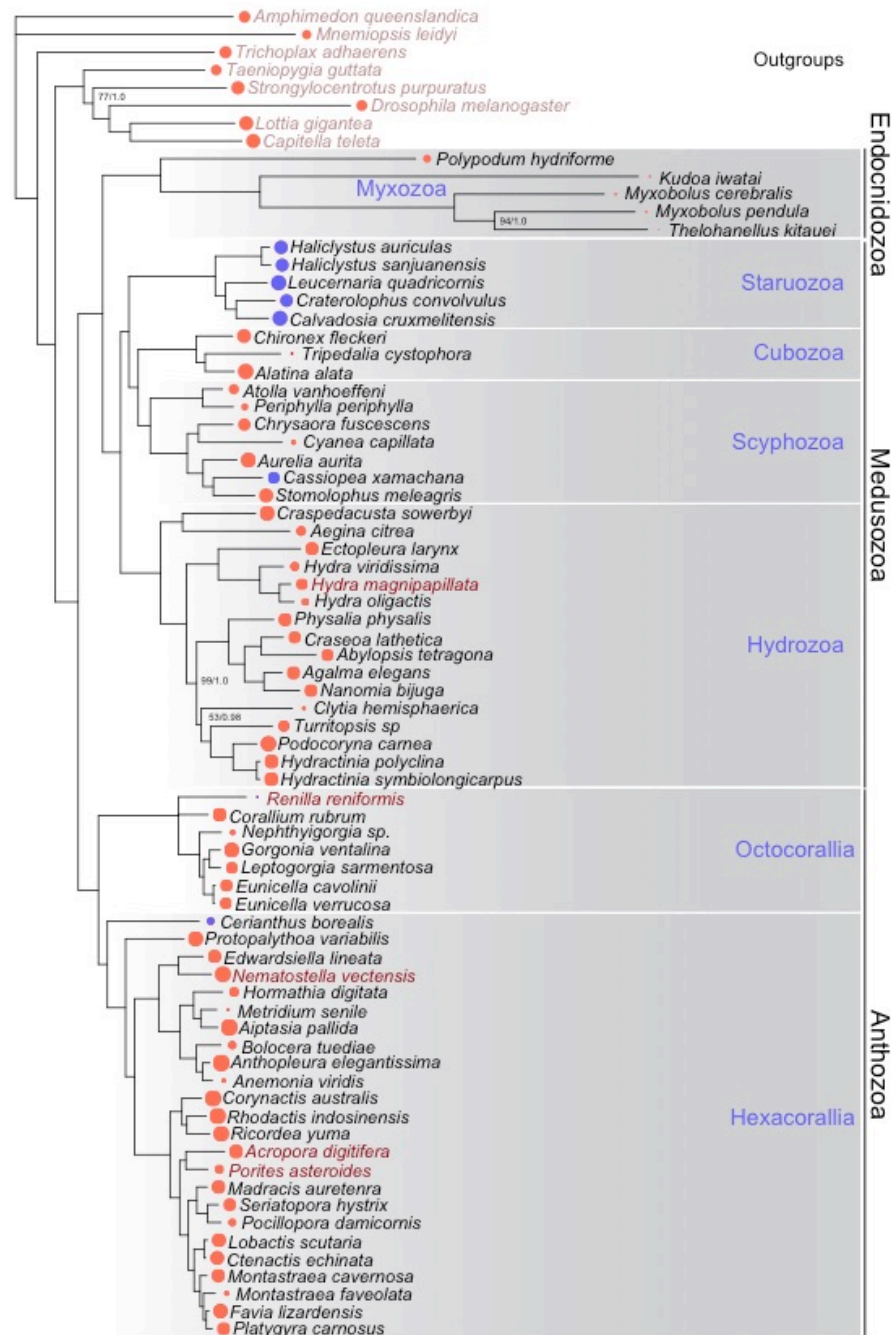
Figure 5. ML and Bayesian analyses of the OF-PTP_75tx dataset confidently resolve cnidarian phylogeny and depict the three major lineages. Shown is the topology from ML analyses (Supplementary Figure S5). Nodes not present in the Bayesian analysis (Supplementary Figure S6) are collapsed to polytomies. All nodes receive maximum support in both analyses except where indicated as bootstraps and posterior probabilities. Circles at terminal tips indicate the number of data partitions present per taxon. Blue circles indicate data produced here, while orange circles indicate publically available data. All datasets are derived from RNAseq data except when whole genome assemblies where used as indicated in Red.

S6 and S7). Overall, results from Bayesian and ML analyses were congruent, with all but three shallow nodes in the cnidarian ingroup receiving maximum support in both analyses.

The following phylogenetic findings related to our analyses of OF-PTP_75tx are recovered in both ML and Bayesian analyses and receive maximum support in each as detailed in Figure 5. We recovered a monophyletic Anthozoa as sister to a clade containing Medusozoa plus Endocnidozoa. In addition, Ceriantharia, represented by *Cerianthus borealis*, is sister to Hexacorallia. Consistent with Chang et al. (2015), our analyses recovered a monophyletic Endocnidozoa (Myxozoa + Polypodiozoa) as sister to Medusozoa. Our analyses split Medusozoa into two monophyletic clades consisting of Hydrozoa (Trachylinae + Hydroidolina) and Acraspeda, a lesser-known clade uniting Staurozoa, Cubozoa and Scyphozoa. Within Scyphozoa, we recovered a paraphyletic Semaeostomeae where *Aurelia aurita* grouped with Rhizostomeae. Within Hydrozoa, Trachylinae is the sister lineage to the remaining Hydroidolinia, which is further divided into Aplanulata and a clade comprised of siphonophores, the leptothecate *Clytia hemisphaerica* and species of the Filifera IV group.

The histories of key cnidarian traits

We applied explicit statistical approaches and stochastic mutational mapping (Bollback 2006; Nielsen 2002; Revell 2012) to reconstruct ancestral character states for selected traits on our topology (Figure 6). In addition, we conducted a Bayes Factor test comparing the prior and posterior probabilities of each trait evolving either once or multiple times (Minin 2014) using a range of gain and loss rate parameters including empirical estimates (Beaulieu, et al. 2013) (Table 1). Of the characters we examined, we recovered strong support for multiple origins of the intracellular, autotrophic, eukaryotic symbiont character (P = 0.96; Table 1), which occurred independently within all major classes of Cnidaria except the parasitic Endocnidozoa (Figure 6). In contrast, results from our analyses of coloniality were less clear. We find marginal support for a single origin of coloniality across the tree (Table 1; P = 0.83) but ancestral state reconstructions provided marginal support for the hypothesis that the last common ancestor of the included cnidarian taxa possessed the alternative, solitary, character state (PP=0.76).

Life history stages within Cnidaria are strikingly plastic, making universal definitions difficult (Cartwright and Nawrocki 2010; Leclere, et al. 2009; Marques and Collins 2004; Morandini AC 2016; Okamura 2015). Here we consider the medusa to be a sexually mature, solitary, free-swimming life history stage that spawns after separation or metamorphosis from a polyp. We consider the polyp stage to be a post-larval non-medusa stage. Our ancestral state reconstructions recovered only equivocal support (PP= 0.50) for the presence of medusa in the last common ancestor of Medusozoa with several inferred losses throughout the group, most likely in the lineages leading to Staurozoa and again within Aplanulata and Siphonophora. However, we recover strong support for a single origin of medusa (P = 0.98) on our tree. The polyp life history stage is common across cnidarian phylogeny with notable losses in Endocnidozoa and Trachylinae. A single origin of the polyp stage is well supported (P = 0.93) and the last common ancestor of our well-sampled cnidarian phylogeny is strongly inferred to have a polyp stage (PP=1.0). Our results from Bayes Factor tests for multiple origins are based on empirically derived gain and loss rate parameters, but hold true under a wide range of exaggerated rates of gain and loss (Table 1).

Discussion

The Phylogenetic structure of Cnidaria

Cnidaria is a large and diverse clade that has produced numerous fascinating evolutionary novelties since at least the middle Cambrian (Cartwright, et al. 2007).
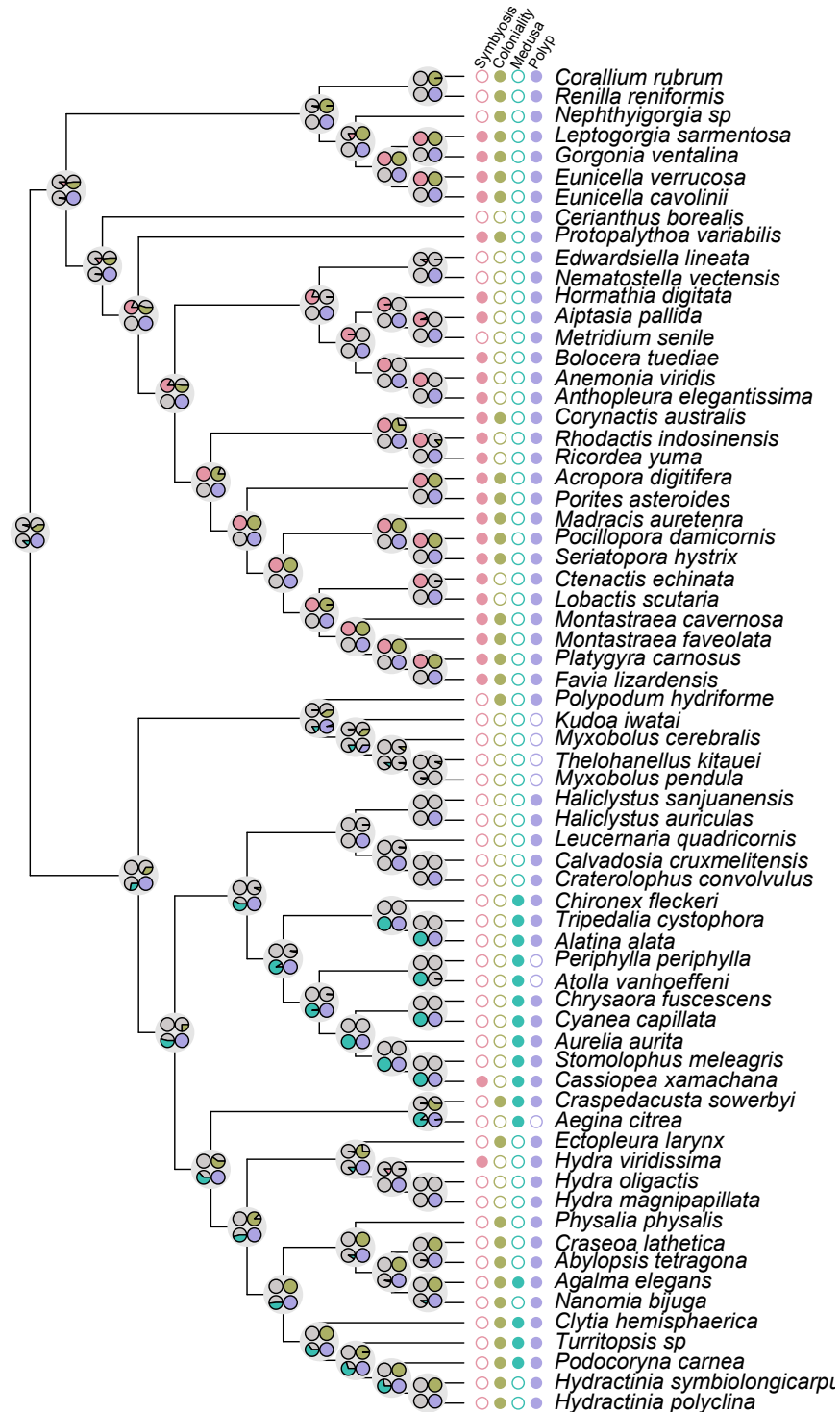
Figure 6. Studies of character evolution reveal a dynamic history of gain and loss for each trait examined. We conducted simulation mutational mapping (Bollback 2006; Revell 2012) on our cnidarian ingroup topology (branch lengths not shown) for each character state included. Ancestral state reconstructions of each character at each node are shown as pie charts representing posterior probabilities. The color-coded presence or absence of each character is shown at Right. Our results strongly suggest that the ancestral state of Cnidaria was a solitary, aposymbiotic polyp, however other characters are equivocal at this node.

Understanding the origin and evolution of these innovations requires a stable phylogenetic framework, but resolving the relationships of the major lineages has eluded evolutionary biologists. Here we present a well-resolved cnidarian phylogeny using genomic and transcriptomic data from representatives of all classes.

Our analyses support the monophyly of Anthozoa, with the enigmatic Ceriantharia placed as sister to Hexacorallia, and corroborate previous findings of a sister relationship between Medusozoa and Endocnidozoa (Myxozoa and *Polypodium hydriforme*) (Chang, et al. 2015; Jimenez-Guri, et al. 2007; Siddall, et al. 1995). Importantly, our findings strongly support the existence of a clade consisting of Staurozoa, Cubozoa and Scyphozoa, which drew only weak support in previous analyses (Zapata, et al. 2015). Our placement of Staurozoa revives evolutionary hypotheses put forward more than half a century ago by Hyman (1940) and Thiel (1966) and includes the union of the rhopalia-bearing Cubozoa and Scyphozoa in the clade Rhopaliophora (Ax 1996).

Our study provides the most comprehensive taxon sampling of Cnidaria for phylogenomic analysis to date (Supplementary Table 3). While our attempts to account for potential contamination resulted in the dramatic reduction of data for some taxa we managed to recover sufficient partitions with at least 50% taxon occupancy to estimate a robust phylogeny for Cnidaria (Figure 5). We show that given the same input data, the Agalma pipeline under default (Dunn, et al. 2013) tended to produce larger though sparser data matrices as compared to the OF-PTP procedure (Emms and Kelly 2015; Kocot, et al. 2013). In our analyses, the data matrix produced by the Agalma pipeline resulted in the misplacement of the data-poor Myxozoa within the vertebrates due to contamination (Figure 3, Supplementary Figures S1 and S2) despite our extensive data filtering steps. We note that our analyses are based on the earlier version of Agalama v0.5-devel and a subsequent update has been made available https://bitbucket.org/caseywdunn/agalma.

The minuscule amount of overlap between the Agalma and OF-PTP datasets (only 52 *N. vectensis* loci shared across partitions between AG_62tx and OF_62tx; Figure 4C) was surprising, given that both pipelines use similar approaches: TransDecoder (Haas 2016) to produce translated peptidomes, all-by-all BLAST to generate similarity graphs, and MCL clustering (Van Dongen 2000) to define orthologous gene clusters. Why would similar approaches produce largely different datasets using identical input data? Minor differences between the Agalma and OF-PTP pipelines include the default MCL inflation parameters (2.1 and 1.5 respectively) and that the Agalma BLAST searches are done using TBLASTX (translated BLAST) while OrthoFinder uses BLASTP searches on TransDecoder (Haas 2016) amino acid translations. However, it seems unlikely that different flavors of BLAST would drive extensive differences in similarity search outcomes, and it has been demonstrated that varying the inflation parameter in MCL clustering does not have a major effect in resulting orthogroups (Li, et al. 2003). One important difference that could explain the construction of largely different datasets by the two pipelines is the procedure used for pruning orthogroups when multiple representative sequences from each taxon are present. The treeprune procedure in Agalma may produce several partitions per orthogroup, while the OF-PTP pipeline uses PhyloTreePruner (Kocot, et al. 2013), which produces only a single partition (the largest monophyletic clade) from each orthogroup, regardless of the topology. Differences in orthogroup pruning are also likely to drive differences in matrix sparseness, which can exacerbate the influence of contamination in phylogenomic datasets (Figure 3). While there are advantages and limitations to each approach, the important consideration here is that phylogenetic analyses of very different matrices produced by both pipelines yielded identical topologies when free of contamination. (Supplementary Figure S2).
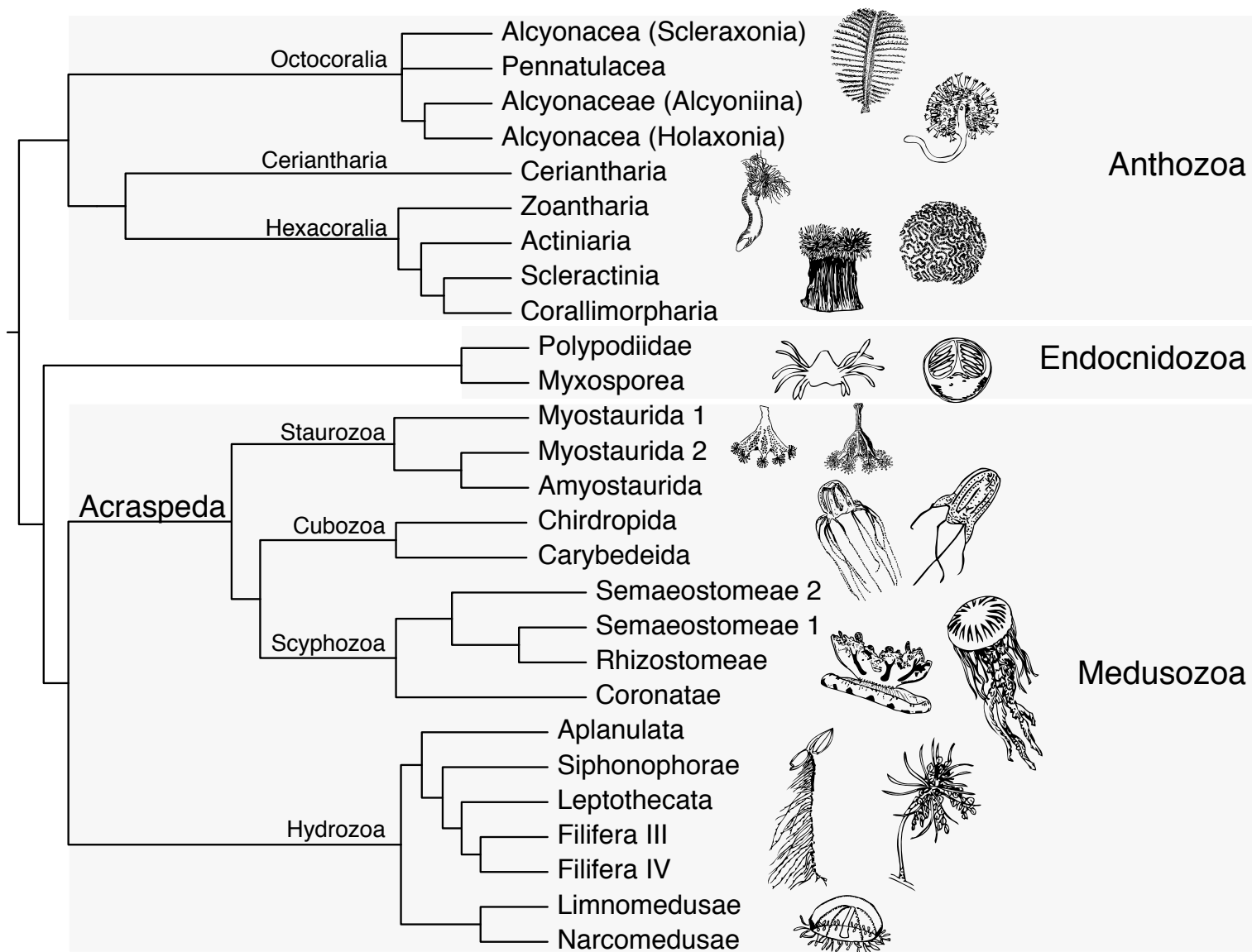
Figure 7. Summary of results. Our working hypothesis for the major lineages of Cnidaria based on the present study.

Deciphering relationships within Anthozoa

This study is the first to confidently determine the position of the ceriantharian tube anemones (Figure 2) as the sister group to Hexacorallia within Anthozoa (Figures 5 and 7). Our result contradicts the favored hypothesis of Stampar et al. (2014) that Ceriantharia is the sister to the remaining Anthozoa, and corroborates earlier but weakly supported hypotheses based on morphology (Daly, et al. 2003), mitochondrial genomes (Kayal, et al. 2013), and phylogenomic datasets (Zapata, et al. 2015). Ceriantharians possess several unique characteristics (Daly, et al. 2003) and following the suggestion of Stampar et al. (2014) we treat it as a unique taxon among the other anthozoan clades (Figure 7). Although often treated as members of Hexacorallia (Daly, et al. 2003), ceriantharian mesenteries, which divide the gastric cavity, are not paired as they are in Hexacorallians. Both taxa do, however, possess a distinctive nematocyst type known as a spirocyst, which is likely a synapomorphy for the clade (Daly, et al. 2003).

In addition to other unusual characteristics, ceriantharians possess swimming larvae called cerinula that are somewhat similar to medusae (Figure 2B) (Molodtsova 2003). In some cases, these stages even develop gonads (Bourne 1919; Stampar 2015; Vanhoeffen 1895) and indeed such stages are responsible for one of the more interesting, if obscure, confusions in cnidarian biodiversity studies. Haeckel (1880) observed several such stages and erected a family, Tesseridae, that he concluded was part of Stauromedusaea, whose members are otherwise benthic (see below). For decades, the swimming pelagic species of Tesseridae went mostly unobserved and were neglected in compendia of known medusae (Kramp 1961) until Goy (1979) reported an observation and documented its veracity. It was not until a few years ago (Rodriguez, et al. 2011) that Goy's species and Haeckel's family were recognized for what they are, precocious larval tube anemones that had yet to settle and secrete their tubes into which they would project their soft adult bodies.

Our taxon sampling for Hexacorallia is extensive, with all orders present except Antipatharia (black corals), providing confidence in the sister-relationship of Ceriantharia with Hexacorallia (Figures 5 and 7). Our phylogeny, however, does not address many open questions regarding relationships within the second anthozoan subclass Octocorallia. Addressing these issues will require increasing the breadth of taxon sampling across Alcyonacea and the inclusion of Helioporacea (e.g., the reef-forming blue corals) in future phylogenomic studies.

Resolving key controversies within Medusozoa

The benthic Stauromedusae of the class Staurozoa have had a long and confused taxonomic history (Figure 1). The earliest studies classified them as anthozoans in the 19th century taxon Polypi of the Actiniae (Cuvier 1830, 1817; Gosse 1858), but Sars (1846) was the first to note that the finger-like gastric cirri and the four-part arrangement of gonads in Stauromedusae bore a striking resemblance to similar features of non-hydrozoan medusae. Indeed, our results strongly suggest that gastric cirri and a quadripartite body plan are synapomorphies of the clade uniting Staurozoa, Cubozoa and Scyphozoa. In the late 1800's, the so-called stalked jellyfishes were the subject of discourse in the nascent field of evolutionary biology where they were viewed as "degenerate scyphomedusae" (Hornell 1893), "arrested scyphistoma[e]" (Hurst 1893), or as "ancestral forms" representing an early diverging lineage "equivalent in value" to the scyphomedusae (Clark 1862). This former view is borne out by our results showing strong support for Staurozoa as the sister group to Cubozoa plus Scyphozoa (Figures 5, 7) and is also in agreement with the topology of Zapata et al. (2015), which had only weak support. Earlier phylogenetic analyses of rDNA and morphology also supported the view that Staurozoa is a distinct clade from Scyphozoa and Cubozoa, but instead suggested that

Staurozoa was the sister group to the remaining medusozoans (Collins, et al. 2006b; Van Iten, et al. 2006).

Our data strongly suggest that Staurozoa is a member of a monophyletic clade containing Cubozoa and Scyphozoa. The earliest taxon name that could apply to this clade is Acraspeda (Gegenbaur 1886), which was originally restricted to scyphozoan and cubozoan species, but later included Staurozoa in a discussion of an evolutionary series linking Stauromedusae to Coronatae and Discomedusae (Claus 1883; Haeckel 1880; Maas 1907, 1903). During this same period, Goette (1887) originated the name Scyphozoa and included Stauromedusae as one of its orders. Based on distinct life cycle and polyp traits, Werner (1973) extracted Cubomedusae from Scyphozoa as the Cubozoa, and by a similar analysis of life history and anatomical traits, Marques and Collins (2004) established Stauromedusae as the medusozoan class Staurozoa. In addition, the first explicit name for Cubozoa plus Scyphozoa, the Rhopaliophora, was introduced by Ax (1996) and we follow that here, while using Gegenbaur's Acraspeda as the clade uniting Staurozoa and Rhopaliophora (Figure 7). We note that Haeckel (1880) appears to be the first to use the clade name Acraspeda in its present sense.

Evolution of complex characters in Cnidaria

Our phylogenomic analyses of Cnidaria provide a framework for understanding the evolutionary histories of several important organismal traits that contributed to the success of the phylum. We scored each taxon in our dataset for the presence or absence of photosynthetic eukaryotic endosymbionts, colonial body plans, a medusa stage, and, a polyp stage, as discrete characters (Figure 6). We used explicit models for character evolution to reconstruct individual character histories, rates of gain and loss and numbers of origins (Bollback 2006; Revell 2012; Minn; Beaulieu). The efficacy of ancestral state reconstruction is dependent upon taxon sampling and, while our dataset is larger than previous phylogenomic analyses of cnidarians (Figure 4B), we emphasize that taxon sampling in certain octocoral and hydrozoan subclades remain sparse.

The acquisition of phototrophic endosymbionts, including zooxanthellae and/or zoochlorellae, constitutes a major ecological innovation in the evolutionary history of Cnidaria allowing these organisms to thrive in oligotrophic waters (Roth 2014). Endosymbiosis in hexacorals, especially the scleractinian corals, is a major area of interest, but is also found in every other cnidarian class except Endocnidozoa. Our analyses support the hypothesis that endosymbiosis has evolved multiple times during the evolutionary history of Cnidaria, with independent origins likely occurring in hexacorals, octocorals, scyphozoans and hydrozoans (Figure 6). This finding may highlight both the adaptive utility of photosynthetic endosymbionts in oligotrophic environments and the possibility of a shared underlying mechanism for the establishment and maintenance of endosymbiotic autotrophs across disparate cnidarian clades.

Colonial organization among disparate cnidarian clades may facilitate adaptations related to maximizing nutrient and spatial resources, defense and reproductive success and is associated with enhanced modularity and the potential for division of labor among zooids (somatically integrated individuals that arise by budding or division) (Hughes 1989). In colonial anthozoans, colonies are monomorphic and exhibit only feeding gastrozooids, whereas in colonial hydrozoans, colonies may be polymorphic with multiple functional classes of zooids present. Division of labor of this type reaches its pinnacle in the siphonophores where it parallels the level of functional specialization exhibited by the cells of other multicellular organisms (Dunn 2009). Our analysis of the character history of coloniality across Cnidaria shows that this trait was likely present in the last common ancestor of octocorals, scleractinians and hydrozoans, but was absent or equivocal in all other deeper nodes including the last

common ancestor of Cnidaria. Strong evidence for loss of coloniality is found only in the Aplanulata hydrozoans, which include the species within the genus *Hydra*.

The polyp is generally regarded as the ancestral life history state in Cnidaria, to which the medusa was added in one or more lineages (Daly, et al. 2007; Salvini-Plawen 1978). Defining different cnidarian life history stages is often difficult because of the many variations and exceptions exhibited within the wide variety of taxa (Cartwright and Nawrocki 2010; Leclere, et al. 2009; Marques and Collins 2004). Our definition of the medusa as a liberated, propulsive life history stage bearing gonads (Cartwright and Nawrocki 2010) requires that the solitary, benthic members of Staurozoa are scored as possessing a derived polyp rather than a degenerated medusa as suggested earlier (Collins, et al. 2006b; Kayal, et al. 2013; Thiel 1966; Uchida 1972; Werner 1973) and the stolon stage of *Polypodium hydriforme* is scored as neither a medusa nor a polyp (Evans, et al. 2008).

Our analyses strongly support the polyp-first hypothesis (Hyman 1940; Werner 1973) with the prediction of a polypoid ancestor to Cnidaria, with at least two independent losses of the polyp stage in lineages leading to Endocnidozoa and within Trachylina (Collins, et al. 2008) (Figures 6, 7). Previous studies have suggested a single innovation of the medusa form within Medusozoa (Kayal, et al. 2013), with independent losses in several Hydrozoa clades (Cartwright and Nawrocki 2010). Our results also favor a single origin of medusa with independent losses of this stage in the lineages leading to Staurozoa and Aplanulata (Table 1, Figure 6). These analyses illustrate the remarkable variation of life history strategies within Cnidaria and set the stage for research into the genomic and developmental factors underlying these transitions.

Conclusions

Cnidaria have experienced more than 600 million years of independent evolution and in the process generated an array of biological innovations. Some of these innovations (e.g., cnidocytes) evolved in the stem of Cnidaria, but many of the most intriguing (e.g., endosymbiosis, coloniality, and the medusa life history stage) likely evolved after the last common cnidarian ancestor and were lost in some lineages. The well-resolved phylogenetic relationships put forth in this study, as well as the ancestral reconstruction of some of these traits marks a major step towards understanding the extraordinary evolutionary history of Cnidaria. While our analyses do not reveal the states of all of our selected characters with confidence in the ancestral cnidarian, a solitary non-symbiotic polyp that lacked a medusa stage remains the most likely prediction, with multiple independent origins of symbiosis occurring subsequently.

Material and methods

Taxon sampling and sequencing

We generated new transcriptome data from a range of cnidarian taxa including five staurozoans (*Calvadosia cruxmelitensis*, *Craterolophus convolvulus*, *Haliclystus auricula*, *Halyclystus "sanjuanensis" (nomen nudum)* and *Leucernaria quadricornis*), one cerianthid (*Cerianthus borealis*), one scyphozoan (*Cassiopea xamachana*) and gene model data from whole-genome sequencing of one octocoral (*Renilla reniformis*). To these we added the following previously published data: 13 cnidarian transcriptomes from Zapata et al. (2015), 30 RNA-seq datasets from the NCBI SRA Archive, and16 transcriptomes and gene models from whole-genome data. We included the same seven outgroups used by Zapata et al. (2015) to which we added *Lottia gigantea* (Simakov, et al. 2013). Details on taxon sampling and data sources are provided in Supplementary Tables S1-S3. Materials used for sequencing were either sampled from whole organisms, or from multiple tissue types per taxon as to broaden

transcript diversity. Further details, including extraction methods, DNA and RNA library preparation, and sequencing are provided in the Supplementary Methods.

Sequence assembly and translation

After adaptor filtering using Trimmomatic v0.33 (Bolger, et al. 2014) with default settings and retaining reads greater than 80 bp for 100 bp-length sequencing runs and 100 bp for 150 bp-length sequencing runs, we assembled all *de novo* transcriptomes using Trinity v2.0.6 (Haas, et al. 2013) with default parameters (Trinity v2.3 was used for *H. auricula*, *L. quadricornis*, and *C. borealis*). For each transcriptome, transcripts were translated into peptides using default settings in TransDecoder v2.0.1 (Haas 2016). We generated an assembly of the *Renilla reniformis* nuclear genome from Illumina paired-end reads as follows: we trimmed adapters with Trimmomatic v0.32 (Bolger, et al. 2014), performed error-correction with Allpaths-LG version 44837 (Gnerre, et al. 2011), and assembled the processed reads using Platanus version 1.2.1 (with default parameters except k=48) (Kajitani, et al. 2014). We created a coding-region training set using the JGI genome annotations of *Nematostella vectensis* v1.0 (Putnam, et al. 2007) and then used Augustus 3.0.3 (Stanke, et al. 2008) with default parameters to generate *Renilla reniformis* protein predictions.

To minimize the possibility of integrating contaminant or laterally transferred sequences, we removed all sequences that had better BLAST (Altschul, et al. 1990) hits to outgroups than to ingroups. We did this in two steps: first against a database that included a representative set of metazoan and non-metazoan sequences (http://ryanlab.whitney.ufl.edu/downloads/alien_index/), and then against a database that included a set of representative cnidarian sequences and a set of representative bilaterian sequences. We used alien_index version 3.0 (Ryan 2015) to identify sequences with better hits to each outgroup and the remove_aliens script from the alien_index package to build a new FASTA sequence file that excluded potential contaminants. While this process likely removed numerous non-contaminant/non-laterally transferred sequences, our conservative approach made it less likely that we included contaminant sequences and the loss of data was acceptable given the great number of sequences that passed our conservative filter. All commands and scripts used for sequence assembly and translation can be found at https://github.com/josephryan/2017-Kayal_et_al.

Construction of phylogenomic datasets

We built two preliminary datasets consisting of 54 cnidarian taxa and eight outgroups using *1)* Agalma v0.5-devel with nucleotide sequences as input (https://bitbucket.org/caseywdunn/agalma) as in Zapata et al. (2015) and, *2)* a custom phylogenomics pipeline consisting of OrthoFinder v0.4.0 (Emms and Kelly 2015) followed by PhyloTreePruner (Kocot, et al. 2013) and our associated wrapper scripts that we refer to as OF-PTP. OF-PTP takes the TransDecoder-translated peptide sequences from each transcriptome as input. The final supermatrices produced by both approaches were filtered to include partitions with greater than 50% taxon occupancy, which were then used for phylogenetic analyses. Following preliminary phylogenetic analyses we noticed that the myxozoan taxa showed evidence of contamination even after extensive filtering with alien index and this contamination resulted in Endocnidozoa being placed within Verterbrata in the Agalma, but not the OF-PTP dataset. To investigate this further, we obtained all data partitions from AG_62tx and OF_PTP_62tx that had myxozoan data and assessed how many putative contaminant sequences were present in each. We also conducted BLAST (Altschul, et al. 1990) similarity searches for each myxozoan sequence captured by both pipelines against a BLAST database

comprised of the protein models from ten high-quality, phylogenetically disparate metazoan genomes, including teleosts and cnidarians. The frequency that the top BLAST hit for each myxozoan sequence resided in a data partition with one, two, three or four other myxozoan species was determined for each data matrix using custom scripts. In addition, while analyzing our preliminary datasets, our efforts and additional publications made available transcriptomic data for several additional cnidarian taxa. We therefore created a final dataset using the OF-PTP pipeline (OF-PTP_75tx) that included 13 of these taxa.

We estimated overlap in data composition between the OF-PTP and Agalma matrices by directly comparing the complement of *N. vectensis* sequences present in each data partition. To do this, we first cross referenced the *N. vectensis* sequences from each partition for each dataset to their full length protein model in the *N. vectensis* v1.0 genome release (Putnam, et al. 2007) using BLAST (Altschul, et al. 1990). We then compared the single best hits for each partition recovered from each dataset. We also explored possible differences in the functional classes represented by sequences present in each data matrix by assigning GO terms to the *N. vectensis* sequences recovered by each pipeline, and to the global *N. vectensis* protein models, using Interproscan v5 (Jones, et al. 2014). GO terms were summarized using REVIGO (Supek, et al. 2011) and their relative enrichment and/or depletion compared to the *N. vectensis* v1.0 protein models (Putnam, et al. 2007) was assessed using Fisher's Exact Tests in custom R scripts. In addition, we determined the number of partitions that included data-per-taxon (taxon occupancy) in our final supermatrices using custom R scripts. Commands and scripts used in the construction of phylogenomic datasets can be found at https://github.com/josephryan/2017-Kayal_et_al.

Phylogenetic analyses and character mapping

For all datasets, preliminary phylogenetic analyses were conducted under the Maximum Likelihood (ML) framework with the best-fit model (PROTGAMMAAUTO) on a single partition using RAxML v8 (Stamatakis 2014). In addition, we estimated an ML tree for our final dataset (OF-PTP_75tx) using the partitioning scheme predicted by PartitionFinder2 (Lanfear, et al. 2017). For all ML analyses, we first performed 20 independent runs using random starting trees under the best-fit model (preliminary analyses) or the modeling scheme predicted by PartitionFinder2 and from these, chose the best scoring tree. In addition, for each analysis, we generated 500 bootstrap replicates under the cognate model as a measure of nodal support. We also conducted Bayesian analyses of OF-PTP_75tx by running two independent chains with PhyloBayes MPI v.1.6 (Lartillot, et al. 2013) under the CAT-GTR model. Each chain was run for more than 4,000 cycles and the resulting topologies were summarized using bpcomp with a burn-in of 0.25 and sampling every 10 trees using PhyloBayes v.4.1. The independent chains did not converge due to the position of the outgroup *Trichoplax adhaerens*. To estimate convergence for the cnidarian ingroup, we removed *T. adhaerens* from all sampled trees using the prune function in Phyutility (Smith and Dunn 2008) and reran bpcomp on both chains.

We conducted character-mapping analyses under the explicit statistical models for character evolution described in SIMMAP and implemented in phytools (Neilsen 2002; Bollback 2006; Revell 2012). SIMMAP uses stochastic mutational mapping to simulate the evolution of characters on a posterior distribution of trees, resulting in estimates of Posterior Probability (PP) for the presence or absence of each trait at each node. We scored each taxon for presence or absence of photosynthetic endosymbionts (including zoozanthellae and zoochlorellae), colonial body plan, a medusa stage, and a polyp stage as discrete characters (Figure 6). In addition, we estimated the rate of gain or loss of each character under a two-rate Markov process using corHMM (Beaulieu, et al. 2013) and, using these estimated rates, we then estimated the marginal likelihoods of single vs. multiple origins for each trait using indorigins (Minin 2014; Pankey, et al. 2014). To test the robustness of analyses of independent origins to differences in

rate parameters we also conducted these analyses with arbitrary, exaggerated rate parameters (Table 1). Commands and custom scripts used for phylogenetic analyses and character mapping can be found at https://github.com/josephryan/2017-Kayal_et_al.

Table 1. Bayes factor analyses of single vs. multiple origins of selected traits

Trait	Priors on rates of gain:loss	BF of single (HO) vs. multiple origins (HA)	Log10(BF)	2xlog_e(BF)	Posterior Probability on HO	Posterior Probability on HA
Symbiosis	0.47:0.47*	0.03683213	-1.433773	-6.60277	0.035523716	**0.964476284**
	10:1	0.008559815	-2.067536	-9.521353	0.008487166	**0.991512834**
	1:10	0.000400292	-3.397623	-15.64663	0.000400132	**0.999599868**
Coloniality	0.7:0.7*	5.003769	0.6992972	3.220383	0.833437962	0.166562038
	10:1	1.528053	0.1841384	0.8479886	0.604438673	0.395561327
	1:10	1.560139	0.1931633	0.8895497	0.609396209	0.390603791
Medusa	0.27:0.27*	63.42498	1.80226	8.299716	**0.98447807**	0.01552193
	10:1	4.380722	0.6415457	2.954427	0.814151335	0.185848665
	1:10	3.150691	0.4984058	2.295244	0.75907626	0.24092374
Polyp	0.2:0.2*	3.540258	0.5490349	2.528399	0.779748199	0.220251801
	10:1	12.90867	1.110882	5.115799	0.928102399	0.071897601
	1:10	8.429244	0.9257886	4.263414	0.893946959	0.106053041

*gain:loss rates estimated from observed data using *corHMM* (Beaulieu et al, 2013)
HO = null hypothesis that trait originated once
HA = alternative hypothesis that triat evolved more than once

List of Abbreviations
GO: Gene ontology
SRA: Sequence read archive

Declarations
Ethics approval
This article does not utilize regulated vertebrate or invertebrates; Not applicable.

Availability of data and material
All data and bioinformatics code utilized in this study is freely available at https://github.com/josephryan/2017-Kayal_et_al.

Competing interests
The authors declare no competing interests.

Authors Contributions
EK, BB, AGC, DCP and JFR conceived the paper, conducted the analyses and wrote the paper. MSP provided additional analyses. AHO and MM provided data sources.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ 1990. Basic local alignment search tool. J Mol Biol 215: 403-410. doi: 10.1016/S0022-2836(05)80360-2

Ax P. 1996. Multicellular animals. Berlin ; New York: Springer.

Beaulieu JM, O'Meara BC, Donoghue MJ 2013. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. Systematic Biology 62: 725-737. doi: 10.1093/sysbio/syt034

Berntson EA, France SC, Mullineaux LS 1999. Phylogenetic relationships within the Class Anthozoa (Phylum Cnidaria) based on nuclear 18S rDNA sequences. Molecular Phylogenetics and Evolution 13: 417-433. doi: Doi 10.1006/Mpev.1999.0649

Bolger AM, Lohse M, Usadel B 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114-2120. doi: 10.1093/bioinformatics/btu170

Bollback JP 2006. SIMMAP: Stochastic character mapping of discrete traits on phylogenies. BMC Bioinformatics 7. doi: Artn 88 10.1186/1471-2105-7-88

Borowiec ML, Lee EK, Chiu JC, Plachetzki DC 2015. Extracting phylogenetic signal and accounting for bias in whole-genome data sets supports the Ctenophora as sister to remaining Metazoa. BMC Genomics 16: 987. doi: 10.1186/s12864-015-2146-4

Bourne GC 1919. Memoirs: Observations on Arachnactis Albida, M. Sars. J. Cell Sci. s2-64: 27-62.

Cartwright P, Halgedahl SL, Hendricks JR, Jarrard RD, Marques AC, Collins AG, Lieberman BS 2007. Exceptionally preserved jellyfishes from the Middle Cambrian. Plos One 2: e1121. doi: 10.1371/journal.pone.0001121

Cartwright P, Nawrocki AM 2010. Character Evolution in Hydrozoa (phylum Cnidaria). Integrative and Comparative Biology 50: 456-472. doi: 10.1093/icb/icq089

Chang ES, Neuhof M, Rubinstein ND, Diamant A, Philippe H, Huchon D, Cartwright P 2015. Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. Proc Natl Acad Sci U S A 112: 14912-14917. doi: 10.1073/pnas.1511468112

Chen CA, Odorico DM, ten Lohuis M, Veron JE, Miller DJ 1995. Systematic relationships within the Anthozoa (Cnidaria: Anthozoa) using the 5'-end of the 28S rDNA. Molecular Phylogenetics and Evolution 4: 175-183.

Clark HJ 1862. Lucernaria, the Coenotype of Acalephae. Proc. Boston Nat. Hist. Soc. ix: 47-54.

Claus C. 1883. Untersuchungen über die Organisation und Entwicklung der Medusen. In. Frytag.

Collins AG 2002. Phylogeny of Medusozoa and the evolution of cnidarian life cycles. Journal of Evolutionary Biology 15: 418-432. doi: Doi 10.1046/J.1420-9101.2002.00403.X

Collins AG 2010. Phylogeny, evolution, and systematics of the stalked jellyfishes (Cnidaria, Staurozoa). Integrative and Comparative Biology 50: E32-E32.

Collins AG 2009. Recent insights into cnidarian phylogeny. Smithsonian Contributions to Marine Sciences: 139-149.

Collins AG, Bentlage B, Lindner A, Lindsay D, Haddock SHD, Jarms G, Norenburg JL, Jankowski T, Cartwright P 2008. Phylogenetics of Trachylina (Cnidaria: Hydrozoa) with new insights on the evolution of some problematical taxa. Journal of the Marine Biological Association of the United Kingdom 88: 1673-1685. doi: 10.1017/S0025315408001732

Collins AG, Schuchert P, Marques AC, Jankowski T, Medina M, Schierwater B 2006a. Medusozoan phylogeny and character evolution clarified by new large and small subunit rDNA data and an assessment of the utility of phylogenetic mixture models. Systematic Biology 55: 97-115. doi: 10.1080/10635150500433615

Collins AG, Schuchert P, Marques AC, Jankowski T, Medina M, Schierwater B 2006b. Medusozoan phylogeny and character evolution clarified by new large and small subunit rDNA data and an assessment of the utility of phylogenetic mixture models. Systematic Biology 55: 97-115. doi: 10.1080/10635150500433615

Cuvier G. 1830. Le Re`gne animal distribue´ d'apre`s son organisation, pour servir de base a` l'histoire naturelle des animaux et d'introduction a` l'anatomie compare´e. . In. Tome III Paris:: Chez De´terville Libraire.

Cuvier G. 1817. Le Re`gne animal distribue´ d'apre`s son organisation, pour servir de base a` l'histoire naturelle des animaux et d'introduction a` l'anatomie compare´e. . In. Tome IV Paris:: Chez De´terville Libraire.

Daly M, Brugler MR, Cartwright P, Collins AG, Dawson MN, Fautin DG, France SC, Mcfadden CS, Opresko DM, Rodriguez E, Romano SL, Stake JL 2007. The phylum Cnidaria: A review of phylogenetic patterns and diversity 300 years after Linnaeus. Zootaxa: 127-182.

Daly M, Fautin DG, Cappola VA 2003. Systematics of the Hexacorallia (Cnidaria : Anthozoa). Zoological Journal of the Linnean Society 139: 419-437. doi: Doi 10.1046/J.1096-3642.2003.00084.X

Dunn C 2009. Siphonophores. Curr Biol 19: R233-234. doi: 10.1016/j.cub.2009.02.009

Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, Smith SA, Seaver E, Rouse GW, Obst M, Edgecombe GD, Sorensen MV, Haddock SH, Schmidt-Rhaesa A, Okusu A, Kristensen RM, Wheeler WC, Martindale MQ, Giribet G 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 452: 745-749. doi: 10.1038/nature06614

Dunn CW, Howison M, Zapata F 2013. Agalma: an automated phylogenomics workflow. BMC Bioinformatics 14: 330. doi: 10.1186/1471-2105-14-330

Emms DM, Kelly S 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol 16: 157. doi: 10.1186/s13059-015-0721-2

Evans NM, Lindner A, Raikova EV, Collins AG, Cartwright P 2008. Phylogenetic placement of the enigmatic parasite, Polypodium hydriforme, within the Phylum Cnidaria. Bmc Evolutionary Biology 8: 139. doi: 10.1186/1471-2148-8-139

Foox J, Siddall ME 2015. The Road To Cnidaria: History of Phylogeny of the Myxozoa. J Parasitol 101: 269-274. doi: 10.1645/14-671.1

France SC, Rosel PE, Agenbroad JE, Mullineaux LS, Kocher TD 1996. DNA sequence variation of mitochondrial large-subunit rRNA provides support for a two-subclass organization of the Anthozoa (Cnidaria). Molecular Marine Biology and Biotechnology 5: 15-28.

Gegenbaur C 1886. Versuch eines Systemes der Medusen, mit Beschreibung neuer oder wenig gekannter Formen; zugleich ein Beitrag zur Kenntnis der Fauna des Mittelmeeres. Z. wiss. Zool. Leipzig 8: 202-273.

Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C,

Goette A. 1887. Entwickelungsgeschichte der Aurelia aurita und Cotylorhiza tuberculata. Hamburg und Lipzig: L. Voss.

Gosse PH 1858. Synopsis of the families, genera, and species of the British Actiniae. The Annals and Magazine of Natural History 3: 414-419.

Goy J. 1979. Campagne de la Calypso au large des côtes atlantiques de l'Amérique du Sud Sud (1961–1962) - 35. In. Méduses. Resultats scientifiques des campagnes de la Calypso. p. 263-296.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, Macmanes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8: 1494-1512. doi: 10.1038/nprot.2013.084

Haas BJaP, A. 2016. TransDecoder (find coding regions within transcripts).

Haeckel E. 1880. Das System der Medusen. I, 2: System der Acraspeden. . In. Gustav Fischer. Jenna. p. 672.

Hampl V, Hug L, Leigh JW, Dacks JB, Lang BF, Simpson AG, Roger AJ 2009. Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". Proc Natl Acad Sci U S A 106: 3859-3864. doi: 10.1073/pnas.0807880106

Hornell J 1893. The Lucernarians as degenerate Scyphomedusae. A note upon the phylogeny of the Order. . Nat . Sci. 3: 204-208.

Hughes RN. 1989. Functional biology of clonal animals.: Springer Science & Business Media.

Hurst CH. 1893. Biological theories. VI. The phylogeny of Lucernarians. In. Nat . Sci. London, New York. p. 208-217.

Hyman. 1940. The Invertebrates: Protozoa through Ctenophora. New York: McGraw-Hill.

Jimenez-Guri E, Okamura B, Holland PW 2007. Origin and evolution of a myxozoan worm. Integrative and Comparative Biology 47: 752-758. doi: 10.1093/icb/icm026

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A,

Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30: 1236-1240. doi: 10.1093/bioinformatics/btu031

Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, Yabana M, Harada M, Nagayasu E, Maruyama H, Kohara Y, Fujiyama A, Hayashi T, Itoh T 2014. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. Genome Res 24: 1384-1395. doi: 10.1101/gr.170720.113

Kayal E, Roure B, Philippe H, Collins AG, Lavrov DV 2013. Cnidarian phylogenetic relationships as revealed by mitogenomics. Bmc Evolutionary Biology 13. doi: Artn 5
10.1186/1471-2148-13-5

Kiderlen H 1937. Die Conularien. Über Bau and Leben der ersten Scyphozoa. Neues Jahrbuch für Mineralogie, Beilige-Band: 113-169.

Kikinger R, Vonsalviniplawen L 1995. Development from Polyp to Stauromedusa in Stylocoronella (Cnidaria, Scyphozoa). Journal of the Marine Biological Association of the United Kingdom 75: 899-912.

Kocot KM, Citarella MR, Moroz LL, Halanych KM 2013. PhyloTreePruner: A Phylogenetic Tree-Based Approach for Selection of Orthologous Sequences for Phylogenomics. Evol Bioinform Online 9: 429-435. doi: 10.4137/EBO.S12813

Kramp PL 1961. Synopsis of the medusae of the world. Journal of the Marine Biological Association of the United Kingdom 40: 7-382.

Lander ES, Jaffe DB 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci U S A 108: 1513-1518. doi: 10.1073/pnas.1017351108

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B 2017. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Mol Biol Evol 34: 772-773. doi: 10.1093/molbev/msw260

Lartillot N, Rodrigue N, Stubbs D, Richer J 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Systematic Biology 62: 611-615. doi: 10.1093/sysbio/syt022

Leclere L, Schuchert P, Cruaud C, Couloux A, Manuel M 2009. Molecular phylogenetics of Thecata (Hydrozoa, Cnidaria) reveals long-term maintenance of life history traits despite high frequency of recent character changes. Systematic Biology 58: 509-526. doi: 10.1093/sysbio/syp044

Li L, Stoeckert CJ, Jr., Roos DS 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13: 2178-2189. doi: 10.1101/gr.1224503

Maas O 1907. Die scyphomedusen. Ergebnisse und Fortschritte der Zoologie 1: 189-238.

Maas O 1903. Die Scyphomedusen der Siboga-Expedition. Siboga-Expeditie 6: 91.

Mariscal RN, Conklin EJ, Bigger CH 1977. Ptychocyst, a Major New Category of Cnida Used in Tube Construction by a Cerianthid Anemone. Biological Bulletin 152: 392-405. doi: Doi 10.2307/1540427

Marques AC, Collins AG 2004. Cladistic analysis of Medusozoa and cnidarian evolution. Invertebrate Biology 123: 23-42.

Minin VN SM, Imholte GC. 2014. indorigin: Testing how many times a trait of interest was regained during evolution (R package).

Miranda LS, Collins AG, Hirano YM, Mills CE, Marques AC 2016a. Comparative internal anatomy of Staurozoa (Cnidaria), with functional and evolutionary inferences. Peerj 4. doi: ARTN e259410.7717/peerj.2594

Miranda LS, Collins AG, Marques AC 2013. Internal Anatomy of Haliclystus antarcticus (Cnidaria, Staurozoa) with a Discussion on Histological Features Used in Staurozoan Taxonomy. Journal of Morphology 274: 1365-1383. doi: 10.1002/jmor.20185

Miranda LS, Hirano YM, Mills CE, Falconer A, Fenwick D, Marques AC, Collins AG 2016b. Systematics of stalked jellyfishes (Cnidaria: Staurozoa). Peerj 4. doi: ARTN e1951 10.7717/peerj.1951

Molodtsova. 2003. On the taxonomy and presumable evolutionary pathways of planktonic larvae of Ceriantharia (Anthozoa, Cnidaria). In: Fautin D.G. WJA, Cartwrigh P., Daly M., Wyttenbach C.R., editor. Coelenterate Biology. Dordrecht: Dordrecht.

Morandini AC SA, Stampar SN, Maronna MM, Straehler-Pohl I, Marques AC. 2016. Succession of generations is still the general paradigm for scyphozoan life cycles. Bull Mar Sci 92: 343-351.

Nielsen R 2002. Mapping mutations on phylogenies. Systematic Biology 51: 729-739. doi: 10.1080/10635150290102393

Okamura B, Gruhl, A. and Reft, A.J. 2015. Cnidarian origins of the Myxozoa. In. In Myxozoan evolution, ecology and development Springer International Publishing. p. 45-68.

Pankey MS, Minin VN, Imholte GC, Suchard MA, Oakley TH 2014. Predictable transcriptome evolution in the convergent and complex bioluminescent organs of squid. Proc Natl Acad Sci U S A 111: E4736-4742. doi: 10.1073/pnas.1416574111

Park E, Hwang DS, Lee JS, Song JI, Seo TK, Won YJ 2012. Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. Molecular Phylogenetics and Evolution 62: 329-345. doi: 10.1016/j.ympev.2011.10.008

Pratlong M, Rancurel C, Pontarotti P, Aurelle D 2017. Monophyly of Anthozoa (Cnidaria): why do nuclear and mitochondrial phylogenies disagree? Zoologica Scripta 46: 363-371. doi: 10.1111/zsc.12208

Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, Terry A, Shapiro H, Lindquist E, Kapitonov VV, Jurka J, Genikhovich G, Grigoriev IV, Lucas SM, Steele RE, Finnerty JR, Technau U, Martindale MQ, Rokhsar DS 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science 317: 86-94. doi: 10.1126/science.1139158

Qiu YL, Li L, Wang B, Chen Z, Knoop V, Groth-Malonek M, Dombrovska O, Lee J, Kent L, Rest J, Estabrook GF, Hendry TA, Taylor DW, Testa CM, Ambros M, Crandall-Stotler B, Duff RJ, Stech M, Frey W, Quandt D, Davis CC 2006. The deepest divergences in land plants inferred from phylogenomic evidence. Proc Natl Acad Sci U S A 103: 15511-15516. doi: 10.1073/pnas.0603335103

Revell LJ 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3: 217-223. doi: 10.1111/j.2041-210X.2011.00169.x

Rodriguez CS, Marques AC, Stampar SN, Morandini AC, Christiansen E, Genzano GN, Mianzan HW 2011. The taxonomic position of the pelagic 'staurozoan' Tessera gemmaria as a ceriantharian larva. Zootaxa: 49-58.

Roth MS 2014. The engine of the reef: photobiology of the coral-algal symbiosis. Frontiers in Microbiology 5. doi: Artn 42210.3389/Fmicb.2014.00422

Ryan JF. 2015. alien_index: alien_index first release: zenodo.

Salvini-Plawen Lv 1978. On the origin and evolution of the lower Metazoa. Journal of Zoological Systematics and Evolutionary Research 16: 40-87.

Sars M. 1846. Fauna littoralis Norvegiae oder Beschreibung und Abbildungen neuer oder wenig bekannter Seethiere, nebst Beobachtungen u¨ber die Organisation, Lebensweise u. Entwickelung derselben. In: Christiania: Druck und Verlag von Johann Dahk.

Shen X-X, Hittinger CT, Rokas A 2017. Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nature Ecology & Evolution 1: 0126. doi: 10.1038/s41559-017-0126

Siddall ME, Martin DS, Bridge D, Desser SS, Cone DK 1995. The demise of a phylum of protists: phylogeny of Myxozoa and other parasitic cnidaria. J Parasitol 81: 961-967.

Simakov O, Marletaz F, Cho SJ, Edsinger-Gonzales E, Havlak P, Hellsten U, Kuo DH, Larsson T, Lv J, Arendt D, Savage R, Osoegawa K, de Jong P, Grimwood J, Chapman JA, Shapiro H, Aerts A, Otillar RP, Terry AY, Boore JL, Grigoriev IV, Lindberg DR, Seaver EC, Weisblat DA, Putnam NH, Rokhsar DS 2013. Insights into bilaterian evolution from three spiralian genomes. Nature 493: 526-531. doi: 10.1038/nature11696

Simion P, Philippe H, Baurain D, Jager M, Richter DJ, Di Franco A, Roure B, Satoh N, Queinnec E, Ereskovsky A, Lapebie P, Corre E, Delsuc F, King N, Worheide G, Manuel M

2017. A Large and Consistent Phylogenomic Dataset Supports Sponges as the Sister Group to All Other Animals. Curr Biol 27: 958-967. doi: 10.1016/j.cub.2017.02.031

Smith SA, Dunn CW 2008. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. Bioinformatics 24: 715-716. doi: 10.1093/bioinformatics/btm619

Song JI, Won, J. H. 1997. Systematics Relationship of the Anthozoan Orders Based on the Partial Nuclear 18S rDNA Sequences. Korean Journal Biological Science: 43-52.

Srivastava M, Begovic E, Chapman J, Putnam NH, Hellsten U, Kawashima T, Kuo A, Mitros T, Salamov A, Carpenter ML, Signorovitch AY, Moreno MA, Kamm K, Grimwood J, Schmutz J, Shapiro H, Grigoriev IV, Buss LW, Schierwater B, Dellaporta SL, Rokhsar DS 2008. The Trichoplax genome and the nature of placozoans. Nature 454: 955-960. doi:10.1038/nature07191

Stamatakis A 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312-1313. doi: 10.1093/bioinformatics/btu033

Stampar SN, Morandini, A. C., Branco, L. C., Da Silveira, F. L., & Migotto, A. E. 2015. Drifting in the oceans: Isarachnanthus nocturnus (Cnidaria, Ceriantharia, Arachnactidae), an anthozoan with an extended planktonic stage. Marine biology 162: 2161-2169.

Stampar SN, Maronna MM, Kitahara MV, Reimer JD, Morandini AC 2014. Fast-Evolving Mitochondrial DNA in Ceriantharia: A Reflection of Hexacorallia Paraphyly? Plos One 9. doi: ARTN e8661210.1371/journal.pone.0086612

Stanke M, Diekhans M, Baertsch R, Haussler D 2008. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics 24: 637-644. doi: 10.1093/bioinformatics/btn013

Supek F, Bosnjak M, Skunca N, Smuc T 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. Plos One 6: e21800. doi: 10.1371/journal.pone.0021800
Tekaia F 2016. Inferring Orthologs: Open Questions and Perspectives. Genomics Insights 9: 17-28. doi: 10.4137/GEI.S37925

Thiel H. 1966. The evolution of Scyphozoa: A review. . London: Academic Press.

Uchida T 1972. The systematic position of the Stauromedusae. Publ Seto Mar Biol Lab: 133-139.

Van Dongen SM 2000. Graph clustering by flow simmulation. [University of Utrecht.

Van Iten H 1992. Morphology and phylogenetic significance of the corners and midlines of the conulariid test. Paleontology: 335-358.

Van Iten H, de Moraes Leme J, Simões MG, Marques AC, Collins AG 2006. Reassessment of the phylogenetic position of conulariids (?Ediacaran-Triassic) within the subphylum medusozoa (phylum cnidaria). Journal of Systematic Palaeontology 4: 109-118. doi: 10.1017/S1477201905001793

Vanhoeffen E 1895. Untersuchungen ueber Anatomie undEntwicklung von Arachanctis albida, M. Sars. Bibliotheca. Zoologica 20: 1-14.

Werner B 1973. New investigations on systematics and evolution of the class Scyphozoa and the phylum Cnidaria. Publ Seto Mar Biol Lab: 35-61.

Wietrzykowski W 1912. Recherches sur le de´veloppement des lucernaires. Arch. Zool. Exp. Gen. 5e Serie: 1-95.

Zapata F, Goetz FE, Smith SA, Howison M, Siebert S, Church SH, Sanders SM, Ames CL, McFadden CS, France SC, Daly M, Collins AG, Haddock SH, Dunn CW, Cartwright P 2015. Phylogenomic Analyses Support Traditional Relationships within Cnidaria. Plos One 10: e0139068. doi: 10.1371/journal.pone.0139068