




DATA NOTE

Box, stalked, and upside-down? Draft genomes from diverse jellyfish (Cnidaria, Acraspeda) lineages: *Alatina alata* (Cubozoa), *Calvadosia cruxmelitensis* (Staurozoa), and *Cassiopea xamachana* (Scyphozoa)

Aki Ohdera ¹, Cheryl L. Ames^{2,3}, Rebecca B. Dikow⁴, Ehsan Kayal^{2,5}, Marta Chiodin^{6,7}, Ben Busby³, Sean La^{3,8}, Stacy Pirro⁹, Allen G. Collins^{2,10}, Mónica Medina ^{1,*} and Joseph F. Ryan ^{6,7,*}


¹Department of Biology, Pennsylvania State University, 326 Mueller, University Park, PA, 16801, USA;

²Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, 10th Street & Constitution Avenue NW, Washington DC, 20560, USA; ³National Center for Biotechnology Information, 8600 Rockville Pike MSC 3830, Bethesda, MD, 20894, USA; ⁴Data Science Lab, Office of the Chief Information Officer, Smithsonian Institution, 10th Street & Constitution Avenue NW, Washington DC, 20560, USA; ⁵UPMC, CNRS, FR2424, ABiMS, Station Biologique, Place Georges Teissier, 29680 Roscoff, France;

⁶Whitney Laboratory for Marine Bioscience, University of Florida, 9505 Ocean Shore Boulevard, St. Augustine, FL, 32080, USA; ⁷Department of Biology, University of Florida, 220 Bartram Hall, Gainesville, FL, 32611, USA;

⁸Department of Mathematics, Simon Fraser University, 8888 University Drive, Barnaby, British Columbia, BC, V5A 1S6, Canada; ⁹Iridian Genomes, Inc., 6213 Swords Way, Bethesda, MD, 20817, USA and ¹⁰National Systematics Laboratory of NOAA's Fisheries Service, 1315 East-West Highway, Silver Spring, MD, 20910, USA

*Correspondence address. Joseph F. Ryan, Address: 9505 Ocean Shore Boulevard, St. Augustine, FL, 32080, USA; E-mail:

joseph.ryan@whitney.ufl.edu  <http://orcid.org/0000-0001-8367-0293>; Mónica Medina, Address: 326 Mueller, University Park, PA, 16801, USA;

E-mail: mum55@psu.edu;  <http://orcid.org/0000-0001-5478-0522>.

Abstract

Background: Anthozoa, Endocnidozoa, and Medusozoa are the 3 major clades of Cnidaria. Medusozoa is further divided into 4 clades, Hydrozoa, Staurozoa, Cubozoa, and Scyphozoa—the latter 3 lineages make up the clade Acraspeda. Acraspeda encompasses extraordinary diversity in terms of life history, numerous nuisance species, taxa with complex eyes rivaling other animals, and some of the most venomous organisms on the planet. Genomes have recently become available within Scyphozoa and Cubozoa, but there are currently no published genomes within Staurozoa and Cubozoa. **Findings:** Here we present 3 new draft genomes of *Calvadosia cruxmelitensis* (Staurozoa), *Alatina alata* (Cubozoa), and *Cassiopea xamachana* (Scyphozoa) for which we provide a preliminary orthology analysis that includes an inventory of their respective venom-related genes. Additionally, we identify synteny between POU and Hox genes that had previously been reported in a

Received: 3 April 2018; Revised: 27 March 2019; Accepted: 21 May 2019

© The Author(s) 2019. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

hydrozoan, suggesting this linkage is highly conserved, possibly dating back to at least the last common ancestor of Medusozoa, yet likely independent of vertebrate POU-Hox linkages. **Conclusions:** These draft genomes provide a valuable resource for studying the evolutionary history and biology of these extraordinary animals, and for identifying genomic features underlying venom, vision, and life history traits in Acraspeda.

Keywords: staurozoa; scyphozoa; cubozoa; acraspeda; cnidaria; medusozoa

Introduction

Some of the most fascinating and outstanding mysteries related to the genomic underpinnings of metazoan biology are centered around cnidarians [1]. Active areas of research include the basis of venom evolution and diversification [2–4], mechanisms of independent evolution of image-forming vision (lens eyes) [5–7], and the emergence of a pelagic adult stage within a biphasic (or multiphasic) life cycle [8]. Cnidaria encompasses 3 major clades: Anthozoa, Endocnidozoa, and Medusozoa [9–12]. Anthozoa comprises Hexacorallia and Octocorallia. Hexacorallia includes scleractinian corals, anemones, and zooanthids and is characterized by a 6-fold symmetry, with species exhibiting both colonial and solitary forms. Octocorallia includes sea fans, gorgonians, and soft corals; these animals are characterized by pinnate tentacles in 8-fold symmetry. Endocnidozoa, comprising the parasitic lineages Myxozoa and Polypodiozoa, was only recently properly classified as Cnidaria [13–15]. Medusozoans are characterized by the emergence of a medusa life history stage within some taxa of the clade, their high diversity in regards to life history and morphology, the presence of a linear mitochondrial genome (with a variable number of chromosomes), and by the presence of a hinged cap at the apex of the cnidocyst (cnidarian stinging organelles) [8, 16, 17].

There are ~3,900 described species within Medusozoa, classified into 4 diverse lineages: Hydrozoa (hydroids, hydromedusae, siphonophores), Staurozoa (stalked jellyfish), Cubozoa (box jellyfish), and Scyphozoa (true jellyfish) (Fig. 1A–C) [1, 11]. There exists much debate regarding the phylogenetic relationships among these lineages [10, 16, 18–20]. Recent phylogenomic analyses have placed Staurozoa as the sister to a clade that contains Cubozoa and Scyphozoa, reuniting these lineages into a group called Acraspeda (Fig. 1D) [9, 15, 21]. Given the extensive morphological diversity within Cnidaria, understanding the evolutionary relationships and mechanisms leading to lineage-specific innovations has been of considerable interest but has been fraught with challenges. In particular, the evolution and subsequent loss of the medusoid form in some lineages hints at a complex evolutionary history within Medusozoa [22].

The mechanisms of medusa formation are variable amongst medusozoans, often involving 2 phenotypically distinct life stages—polyp and medusa—that are genotypically identical (reviewed by Lewis Ames [1]). Cubozoan polyps undergo partial or complete metamorphosis and develop into the adult medusoid form capable of sexual reproduction, although in some cases a polyp rudiment remains [23]. Scyphozoan polyps (scyphistomae) undergo a transition known as strobilation, in which the upper calyx proceeds through metamorphosis and transverse fission to produce a medusa [24]. Unlike other medusozoans, staurozoans lack a free-swimming medusa stage but exhibit medusa-associated characters that are present in other medusozoans [25]. The basal portion of the adult forms a stalk, or peduncle, while coronal muscles and gastric filaments, among other features, characterize the apical portion (calyx) of the adult [20, 25–27]. Hydrozoans exhibit the greatest variation in life history strategies and often lack a medusa form [1]. Species that

give rise to the medusoid form do so via lateral buds generated by asexual polyps, while others possess sexual polyps without a free-swimming stage [28, 29]. Elsewhere within Cnidaria, Anthozoa and the parasitic Endocnidozoa lack the medusa stage or medusoid characters entirely. Research on medusa development has shown similar gene expression patterns between hydrozoans and scyphozoans, with developmental genes co-opted for patterning the medusa body plan [30, 31]. Interestingly, strobilation in scyphozoans was recently shown to be under the control of the retinoic acid pathway [32–34]; these same genes are involved in metamorphosis of insects and amphibians, hinting towards regulation of metamorphosis being a conserved function in metazoans [35]. The study also found that potential lineage-specific genes were involved in controlling strobilation, suggestive of genomic innovations within Medusozoa playing a role in medusa morphogenesis, or more specifically within Scyphozoa.

Hox genes, which control body formation during early embryonic development, predate the emergence of both Bilateria and Cnidaria, and the evolution of these genes played a crucial role in the diversification of these lineages [36–38]. In particular, clustering and synteny has been shown to be important in bilaterians [39], but also in some cnidarians, such as the anthozoan *Nematostella vectensis* [36, 40, 41], and in several hydrozoan species [42–46]. Other than an initial characterization of select Hox genes in *Cassiopea xamachana* [47], information about Hox genes and Hox gene clustering in Acraspeda species has been limited. In hydrozoans and vertebrates, Hox genes were shown to be linked to another class of homeobox genes, the POU genes [48], but this linkage has not been demonstrated in any other cnidarian lineages. These new Acraspeda genomes provide us with an opportunity to investigate the evolutionary history of the POU-Hox linkage in more detail.

Genomic resources necessary to understand medusozoan evolution have been lacking, with genomes predominantly available for anthozoans and hydrozoans [49–56]. However, 3 new scyphozoan genomes, 2 genomes of the moon jellyfish *Aurelia* spp. and the giant Nomura's jellyfish *Nemopilema nomurai* were recently sequenced [57–59]. In addition, the genome for the cubozoan *Morbakka virulenta* has also recently been released [59]. While the majority of Medusozoa species are represented by hydrozoans (>90%), both cubozoans and scyphozoans garner significant attention as a result of their impact on economy and tourism [1, 60]. Largely due to venom being used as a mechanism of defense and prey capture, the inherent risk of jellyfish sting has been exacerbated by uncertainty about how cnidarians will respond to modern-day anthropogenic disturbances [61, 62]. Despite these risks, relatively little is known about cnidarian venom, as compared to snakes, cone snails, and other venomous organisms. Given the great phylogenetic distance between cnidarians and these well-studied venomous organisms, a better understanding of the cnidarian venom repertoire can provide insight into the evolution of venom and venom-encoding genes.

Here we present 3 new genomes for species of the 3 major Acraspeda lineages: *Calvadosia cruxmelitensis* (formerly *Lucernar-*

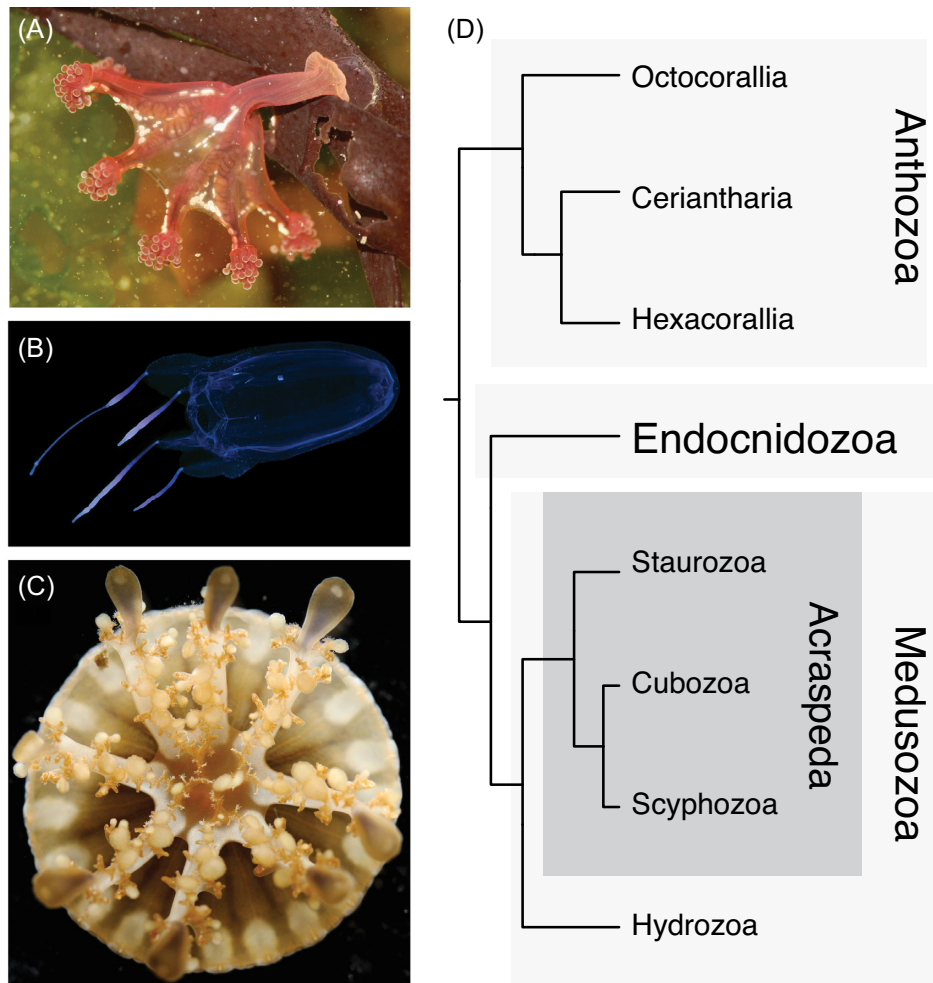


Figure 1: A, *Calvadosia cruxmelitensis* (Staurozoa); B, *Alatina alata* (Cubozoa); and C, *Cassiopea xamachana* (Scyphozoa). D, Phylogenetic relationship of major cnidarian lineages after Kayal et al. [11], revealing Cubozoa and Scyphozoa as sister groups, united with Staurozoa to form Acraspida.

iopsis cruxmelitensis) (Staurozoa, NCBI:txid1843192), *Alatina alata* (Cubozoa, NCBI:txid1193083), and *C. xamachana* (Scyphozoa, NCBI:txid12993). The winged box jellyfish *Alatina alata* (Cnidaria: Cubozoa: Carybdeida: Alatinidae) has been of interest due to its unusual circumtropical distribution [63], extraordinarily rapid gonad development [64], and its reputation as a potent stinger, earning it the honor of being the only jellyfish species to have its own category in US weather reports [65]. The stalked jellyfish *C. cruxmelitensis* has been the recent subject of detailed anatomic [20] and biodiversity studies [27]. The upside-down jellyfish *C. xamachana* is an established model for understanding cnidarian-dinoflagellate endosymbiosis [66] and, with its ease of culturing and tractability in the laboratory setting, is poised as a model system for evolutionary developmental biology research and other laboratory-based studies [67].

The genomes and corresponding gene annotations from these 3 lineages will serve as useful resources aimed at sparking investigative research into the evolution and diversification of life history strategies across cnidarians. Furthermore, future studies examining cnidarian venom evolution, and phylogeographic patterns of venomous jellyfish, may provide opportunities for development of jellyfish-derived therapeutic drugs,

and additional novel biopharmaceuticals (reviewed by Lewis Ames [1]).

Data Description

Genome sampling, sequencing, and assembly

These 3 acraspedan genomes were assembled at different times throughout a 5-year period as part of several independent projects overseen by the coauthors, using separate methods for collection, extraction, sequencing, and assembly (see below). This valuable resource to the scientific community is the culmination of an extensive collaborative effort to respond to the need for model medusozoan systems in a plethora of research fields.

Cassiopea xamachana sample collection and DNA extraction

We propagated *C. xamachana* polyps from a single polyp via asexual budding (Line T1-A). Polyps were maintained symbiont-free at 26°C and fed 3 times weekly with *Artemia* nauplii. To avoid the possibility of food source contaminants interfering

with downstream bioinformatic analysis, we starved the polyps for 7 days in antibiotic-treated seawater prior to preservation in 95% ethanol; any *Artemia* cysts retained within the gut were manually removed before preservation. We extracted genomic DNA from the apo-symbiotic (lacking endosymbionts) polyps using a CTAB (cetyl trimethylammonium bromide) phenol chloroform extraction, first performing an overnight digestion of whole polyp tissue with proteinase K (20 mg/L) in CTAB buffer before proceeding with the standard protocol. DNA extract was stored at -20°C until further processing.

Calvadosia cruxmelitensis sample collection and DNA extraction

We collected adult specimens of *C. cruxmelitensis* in January 2013 at Chimney Rock, off the coast of Penzance, Cornwall, England. Specimens were immediately preserved in ethanol and stored at -20°C until further processing. We extracted genomic DNA using a phenol-chloroform protocol in an Autogen mass extractor (Holliston, MA, USA), and stored the DNA extract at -20°C until further processing.

Alatina alata sample collection and DNA extraction

We collected *A. alata* material during a spermcasting aggregation in Bonaire, The Netherlands (April 2014, 22:00–01:00) according to the methods in Lewis Ames et al. [7]. We selected a single live spermcasting male medusa from the same cohort as the female medusa used in previously published RNA sequencing studies (Genbank Accession: GEUJ01000000) [7, 9]. The medusa was divided into 4 longitudinal sections, and one quarter was placed into a 15-mL tube with pure (99%) ethanol, flash-frozen at -180°C (using a dry shipper), and subsequently transported to the Smithsonian National Museum of Natural History and stored at -20°C . We extracted genomic DNA using a DNeasy Blood & Tissue Kit (Qiagen: Hilden, Germany), following the manufacturer's protocol. DNA extract was stored at -20°C until further processing.

Cassiopea xamachana sequencing and assembly

Library construction and sequencing was performed at Hudson-Alpha Institute for Biotechnology. Four 350-bp paired-end linear libraries with insert sizes of 500 bp were generated with Illumina TruSeq DNA PCR-Free LT Prep Kits (San Diego, CA, USA) and sequenced on the Illumina HiSeq2000. Approximately 634 million reads totaling 117.6 Gb of high-quality paired-end sequence data were generated. We performed adaptor trimming and quality filtering using Trimmomatic v0.36 [68] with default settings, followed by genome size estimation and error correction with Allpaths-LG version 52,488 [69]. We removed mitochondrial reads using FastqSifter v1.1.1 (RRID:SCR.017200) [70] using the *C. xamachana* mitochondrial genome as a reference (NCBI NC_016466.1). We performed *de novo* genome assemblies using ABySS 2.0.1 with default settings [71], SPAdes genome assembler v3.10.0 [72], and Platanus version 1.2.1 (with default parameters, $k = 89$) [73] (Table 1). We used a custom Perl script, plat.pl [74], to invoke the Platanus commands for assembly, scaffolding, and gap closing. Of the 3 assembly methods, Platanus produced the best draft assembly with 93,483 scaffolds measuring a total of 393.5 Mb with an N50 of 15,563 bp (Table 1) (European Nucleotide Archive [ENA] Accession OLM001000000). We recovered 82.66% (53.63% complete and 29.03% partial) of the core eukaryotic genes and 66.97% (58.59% complete and 8.38%

partial) of the core metazoan genes with CEGMA version 2.5 [75] and BUSCO version 2.01 [76], respectively, through the gVolante web server [77] (Table 1).

Calvadosia cruxmelitensis sequencing and assembly

Library construction and sequencing for *C. cruxmelitensis* were performed at the University of Florida Interdisciplinary Center for Biotechnology Research. Four 150-bp paired-end linear libraries and four 150-bp single-end linear libraries with insert size of 300 bp were generated and sequenced on the Illumina NextSeq 500, generating 291,944,064 paired-end reads and 141,911,072 single-end reads. We performed adaptor trimming and quality filtering using Trimmomatic-0.32 [68] with default settings, followed by genome size estimation error correction using Allpaths-LG v.44837 [69]. We removed mitochondrial sequences to improve the final assembly with FastqSifter v1.1.1 (RRID:SCR.017200) [70], using a *de novo* assembly of the *C. cruxmelitensis* mitochondrial genome. We assembled the *C. cruxmelitensis* mitochondrial genome by capturing contigs from an initial assembly using available staurozoan mitochondrial DNA sequences from NCBI as a reference, following the methods presented in Kayal et al. [78], using Geneious v9.0 to generate the final mitochondrial assembly. We checked completeness of the mitochondrial genome using NCBI BLAST against the non-redundant database in addition to a manually generated set of medusozoan genes, annotated transfer RNA (tRNA) genes separately by using tRNAscan-SE [79] and Arwen [80], and checked the integrity of the assembly by aligning the reads to the completed mitochondrial genome. With the mitochondrial sequences removed, we generated 2 “sub-optimal” assemblies using Platanus v1.2.1 with k -mer size of 32 and 45 bp and default settings. Subsequently, we used these sub-optimal assemblies to construct artificial mate-pair libraries for 9 insert sizes (1,000, 2,000, 3,000, 4,000, 5,000, 7,500, 10,000, 15,000, 20,000) with MateMaker v1.0 (RRID:SCR.017199) [81]. We used the artificial mate-pair libraries to scaffold the optimal assembly (generated using Platanus k -mer = 45) with SSPACE Standard v3.0 [82]. This process produced a draft assembly with 417,008 scaffolds measuring a total of 209.3 Mb with an N50 of 16,443 bp (Table 1) (ENA Accession OFHS01000000). We recovered 91.94% (61.29% complete and 30.65% partial) of the core eukaryotic genes and 85.07% (70.86% complete and 14.21% partial) of the core metazoan genes with CEGMA and BUSCO, respectively.

Alatina alata sequencing and assembly

Illumina library preparation and sequencing was conducted at the University of Kansas Genome Sequencing Core. Libraries were generated with the Illumina Nextera Library Preparation Kit (San Diego, CA, USA) and sequenced twice on the Illumina HiSeq 2500. The 2 different runs were performed on the same library: 1 with 100-bp paired-end, and 1 with 150-bp paired-end sequencing, resulting in 564 million reads totaling 148.6 Gb of paired-end sequence data. Pacific Biosciences (PacBio) library prep and sequencing were completed at the University of Washington Northwest Genomics Center. We constructed the libraries with unsharped DNA with end-cleanup only, with an average insert size of 6,000 bp. Sequencing was completed on the PacBio RS II platform, resulting in 486,000 long-reads totaling 990.2 Mb of data. We conducted hybrid assembly of Illumina short-reads and PacBio long-reads using MaSuRCA 3.2.2 [83] (which includes an error correction step for paired-end reads), which resulted in an assembly of 291,445 contigs and an N50 of 7,049

Table 1: Statistics of *Alatina alata*, *Calvadosia cruxmelitensis*, and *Cassiopea xamachana* genome assemblies

	<i>Alatina alata</i>	<i>Calvadosia cruxmelitensis</i>	<i>Cassiopea xamachana</i>
NCBI Taxa ID	1,193,083	1,843,192	12,993
No. of sequences	291,445	50,999	93,483
Estimated genome size	2,673,604,203	230,957,924	361,689,769
Total length (bp)	851,121,747	209,392,379	393,520,168
N50 (bp)	7,049	16,443	15,563
CEGMA (% complete)	8.06	61.29	53.63
CEGMA (% complete + partial)	29.84	91.94	82.66
BUSCO (% complete)	18.30	70.86	58.59
BUSCO (% complete + partial)	32.11	85.07	66.97
Guanine-cytosine content (%)	38.07	39.95	37.07
Assembly accession	PUGI00000000	OFHS01000000	OLMO01000000
NCBI raw read accession	PRJNA421156	PRJEB23739	PRJEB23739
Specimen Voucher ID	USNM 1,248,604	USNM 1,286,381	UF Cnidaria 12,979

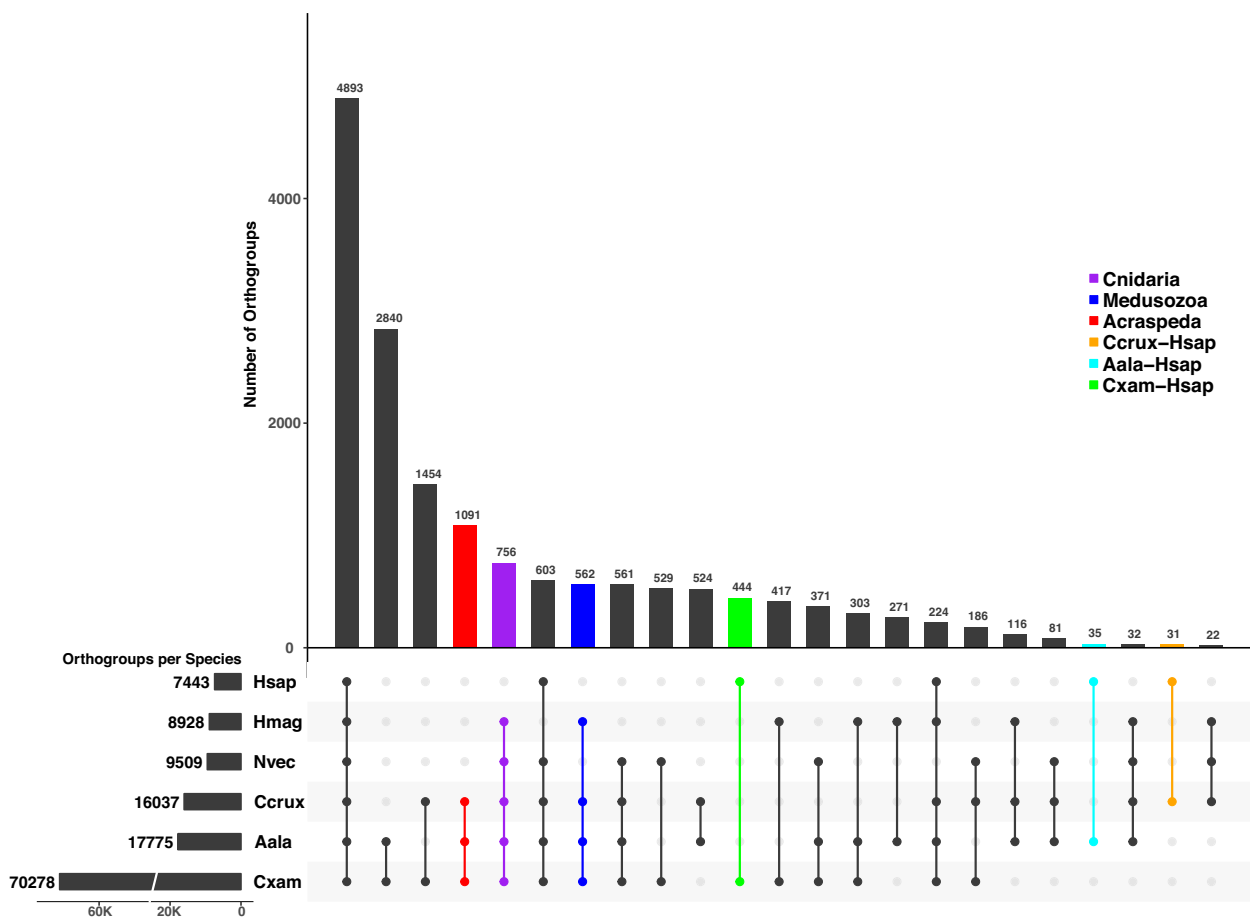


Figure 2: Gene content distribution in cnidarian lineages. Filled circles in the bottom panel indicate shared orthogroups in those lineages. Bar graphs indicate the number of orthogroups corresponding to each filled-circle pattern. Numbers next to each species abbreviation indicate the total number of orthogroups identified for that species. Hsap = *Homo sapiens*; Nvec = *Nematostella vectensis*; Hmag = *Hydra magnipapillata*; Ccrux = *Calvadosia cruxmelitensis*; Aala = *Alatina alata*; Cxam = *Cassiopea xamachana*.

bp (NCBI Accession PUGI00000000). We did not perform adapter trimming prior to assembly because the MaSuRCA manual advises against preprocessing of reads, including adapter removal. Nevertheless, we identified considerable adapter contamination in our final assembly. Subsequently, we used a custom script (`remove_adapters_and_200.pl` [74]) to remove adapters and se-

quences shorter than 200 nucleotides. The total length of the assembly was 851.1 Mb. We recovered 29.84% (8.06% complete and 21.78% partial) of the core eukaryotic genes and 32.11% (18.30% and 13.81% partial) of the core metazoan genes with CEGMA and BUSCO, respectively. The low recovery rates for conserved genes in the *A. alata* genome are likely due to the considerably larger

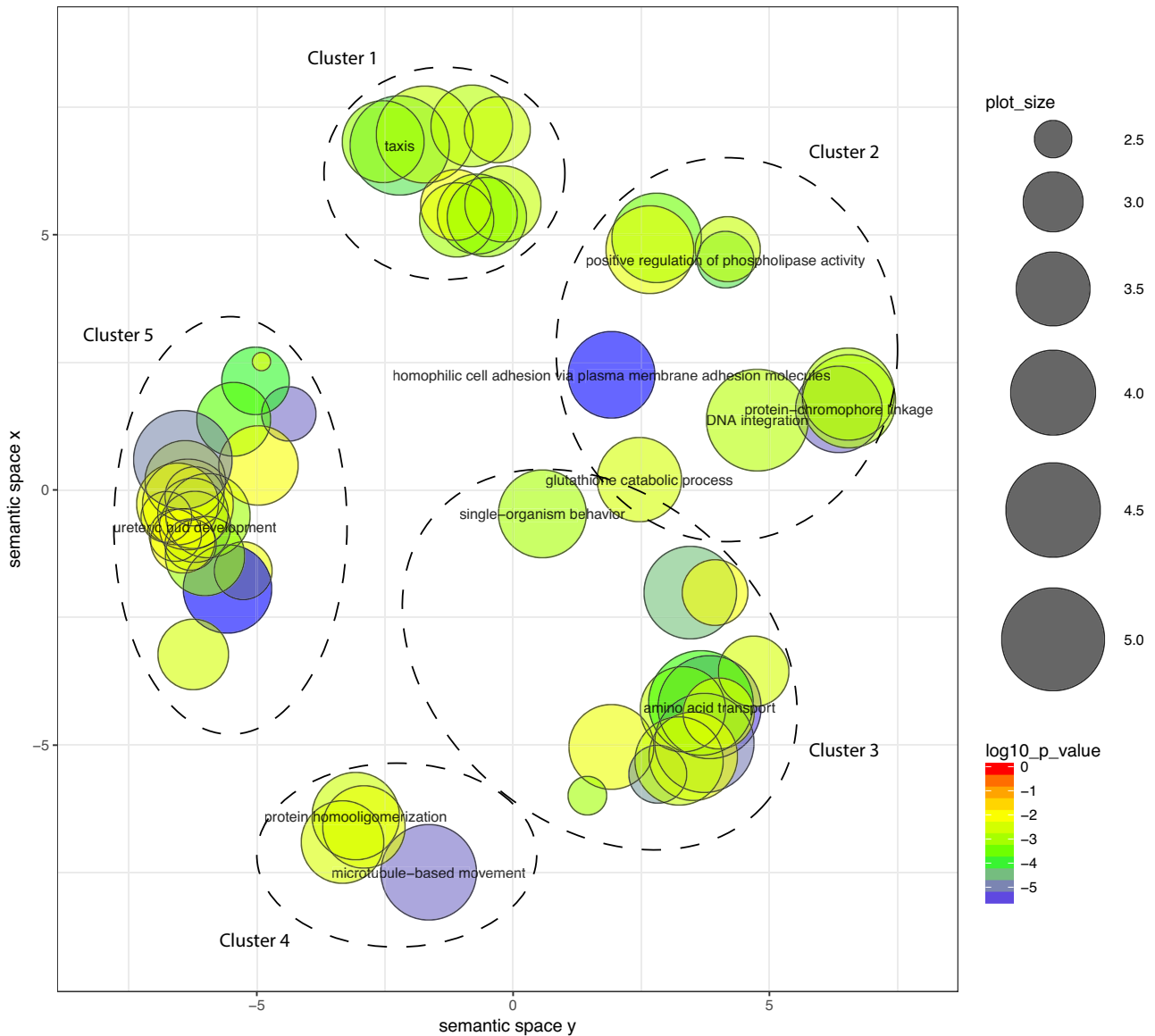


Figure 3: Gene ontology biological processes over-enriched within Cnidaria-specific orthogroups visualized using ReViGO. Over-representation analysis was performed with ClusterProfiler, with a P-adjusted cutoff of 0.01. Color indicates \log_{10} -transformed P-adjusted value. Terms are plotted within an x-y semantic space, in which similar terms are clustered within closer proximities. Color indicates P-value and circle size indicates frequency of GO term in the *Cassiopea* database.

size of the genome, which tends to be coupled with long introns, and therefore higher rates of gene fragmentation in a draft assembly [84–86].

Comparison of Assemblies

A comparison of the draft genomes assembled in this study reveals that the genome of *A. alata* is 4 times the size of that of *C. cruxmelitensis* and almost twice the size of the genome of *C. xamachana*. The apparent contiguity of the assemblies is reflected in this size difference, with the N50 of the *A. alata* genome (7,049 bp) being considerably smaller for both *C. cruxmelitensis* (16,443 bp) and *C. xamachana* (15,563 bp). The N50, the minimum length of at least half the contigs/scaffolds in an assembly, tends to scale with the level of completeness as measured

by CEGMA and BUSCO. For example, CEGMA recovered 91.94% of 248 conserved eukaryotic genes (complete + partial) in *C. cruxmelitensis* and 82.66% in *C. xamachana*, and 29.84% in *A. alata* (Table 1).

Gene Model Prediction

We predicted genes for all 3 genomes using Augustus v3.2.2 [87], with the *Homo sapiens* training set and hits generated with BLAT [88] alignments of published transcriptome data (*C. cruxmelitensis* ENA accession = HAHCO1000000; *C. xamachana* ENA accession = PRJEB21012; *A. alata* accession = PRJNA312373) to the genome assemblies of the respective taxa [11]. The *H. sapiens* training set was used because Augustus gene predictions using the *Nematostella vectensis* v1.0 training set failed to detect intronic re-

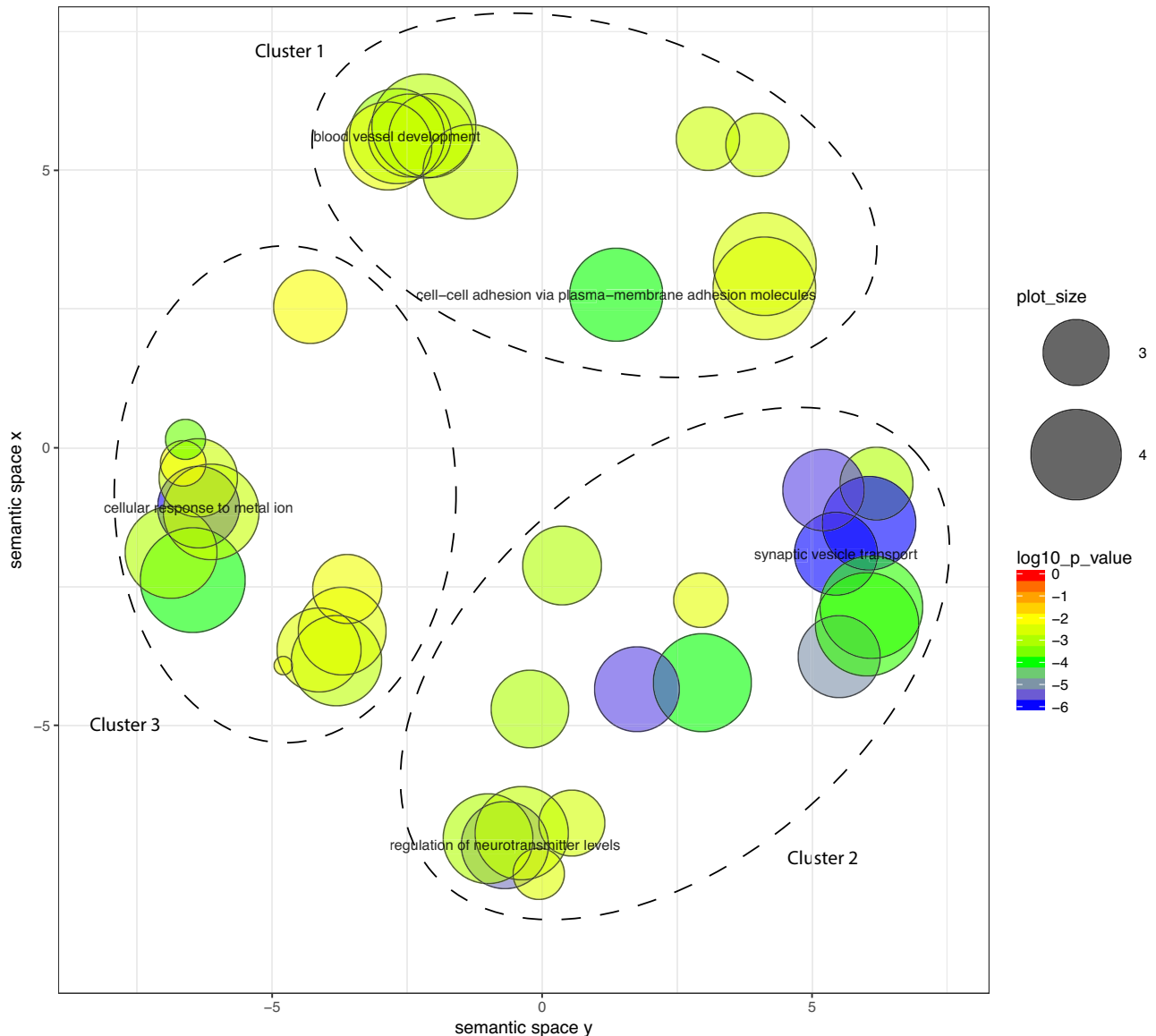


Figure 4: Gene ontology biological processes over-enriched within Medusozoa-specific orthogroups visualized using ReViGO. Over-representation analysis was performed with ClusterProfiler, with a P-adjusted cutoff of 0.01. Color indicates \log_{10} -transformed P-adjusted value. Terms are plotted within an x-y semantic space, in which similar terms are clustered within closer proximities. Color indicates P-value and circle size indicates frequency of GO term in the *Cassiopea* database.

gions within predicted genes, thereby resulting in predicted proteins consisting of single exons. We generated 66,156 gene models for *A. alatina*, 26,258 for *C. cruxmelitensis*, and 31,459 for *C. xamachana*.

Gene Orthology and Lineage-Specific Gene Ontology

We used OrthoFinder v1.1.4 [89] to construct orthologous groups between gene models of *A. alatina*, *C. cruxmelitensis*, *C. xamachana*, *N. vectensis*, *Hydra magnipapillata*, and *H. sapiens*. We also included translated transcriptome assemblies for *A. alatina*, *C. cruxmelitensis*, and *C. xamachana* in these ortholog analyses, as well as an additional transcriptome of the apo-symbiotic polyp stage of *C. xamachana*, which was assembled using Trinity v2.4.0 [90] with default settings (ENA Project Accession: PRJEB23739). All transcriptomes were translated using TransDecoder 3.0.0 [91]

with minimum protein length (-m) set to 50 and all other settings as default. We annotated orthogroups by BLASTing a representative species against the Uniprot/Swiss-Prot database [92]. Orthogroups with annotations were further mapped to Gene Ontology (GO) terms, and analysed for putative enrichment related to biological function using ClusterProfiler [93], against a *C. xamachana* annotation database generated using AnnotationForge [94].

Our OrthoFinder analysis generated a total of 80,482 orthogroups for the combined genomic datasets. Using a custom script, we identified 756 Cnidaria-specific orthogroups, another 562 medusozoan-specific orthogroups, and yet another 1,091 Acraspeda-specific orthogroups (Fig. 2); genes in each taxon-specific orthogroup were non-overlapping. Of these unique orthogroups, we were able to retrieve Swiss-Prot annotations for 57% of Cnidaria-specific orthogroups, 32% of Medusozoa-specific orthogroups, and 55% of Acraspeda-specific orthogroups

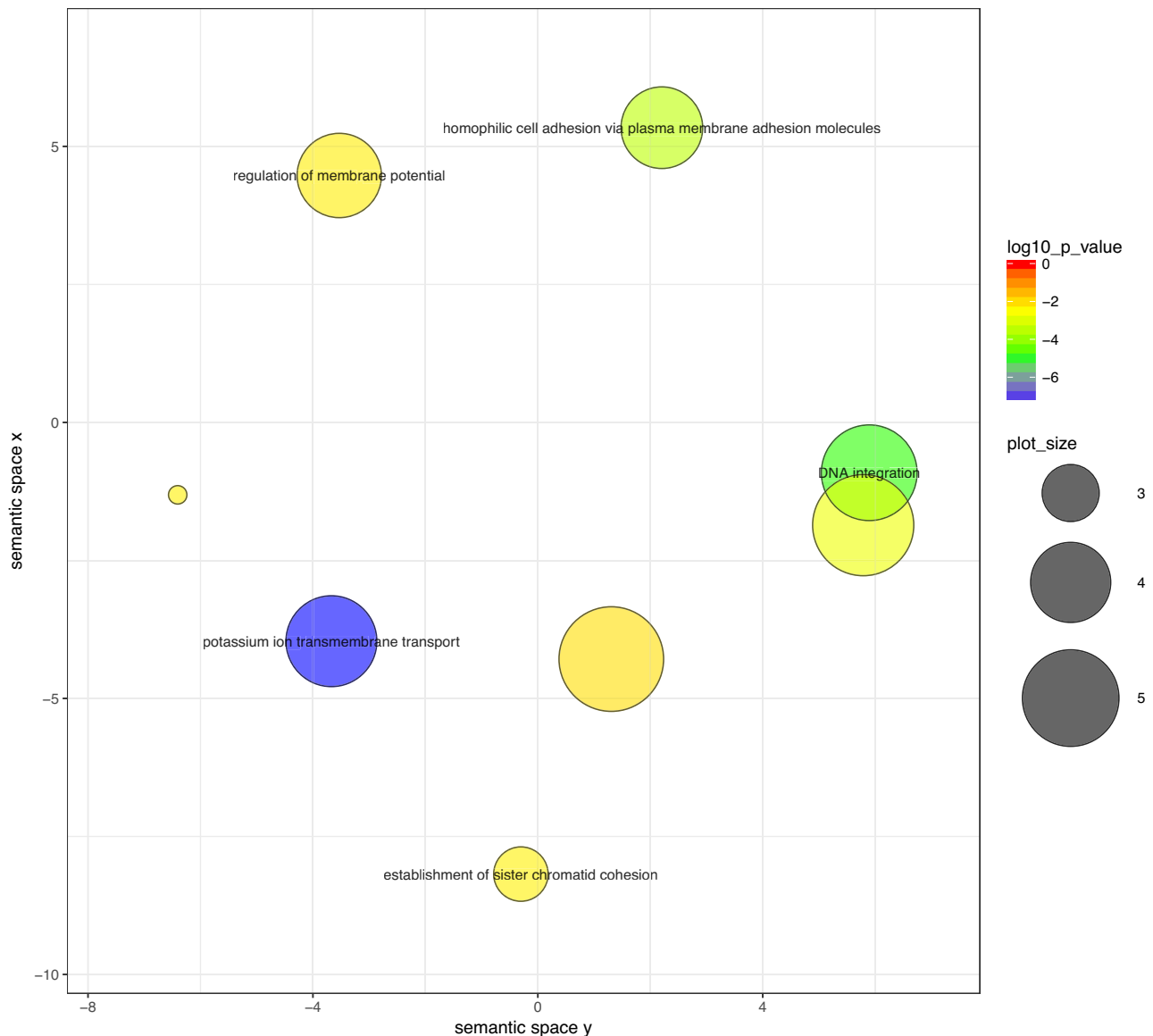


Figure 5: Gene ontology biological processes over-enriched within Acraspeda-specific orthogroups visualized using ReViGO. Over-representation analysis was performed with ClusterProfiler, with a P-adjusted cutoff of 0.01. Color indicates \log_{10} -transformed P-adjusted value. Terms are plotted within an x-y semantic space, in which similar terms are clustered within closer proximities. Color indicates P-value and circle size indicates frequency of GO term in the *Cassiopea* database.

(Tables S1–S3). Unannotated orthogroups may represent taxonomically restricted genes or genes no longer discernable as such due to possibly extensive genetic mutation experienced in evolutionary history. Using this framework, we identified enriched GO terms, uncovering 123 terms corresponding to biological processes that appear to be enriched within Cnidaria: 107 for Medusozoa, and 14 for Acraspeda (adjusted P-value < 0.01). We used ReViGO to remove redundant GO terms from these initial lists and grouped them further through k-means clustering by Euclidean distance. The optimal number of clusters was predicted using the R package NbClust v3.0.

Our ReViGO analysis reduced the 123 Cnidaria-specific GO terms for biological processes to 59 non-redundant ReViGO terms comprising 5 clusters (Fig. 3, Table S1, Fig. S1). Within the 5 clusters, many genes putatively encoding proteins for the cnidarian nerve net were represented (e.g., development and sensory perception), indicative of a system exhibiting

a complex response to physical and chemical stimuli. Additionally, terms related to transport (ion, amines, carbon compounds) and the extracellular matrix were also represented. Genes associated with the extracellular matrix are possibly linked to the cnidarian novelty, the mesoglea (the proteinaceous layer between the endoderm and ectoderm in these diploblastic animals). The 107 Medusozoa-specific GO terms were reduced to 41 non-redundant ReViGO terms, which when grouped by k-means formed 3 clusters (Fig. 4, Table S2, Fig. S2). Similar to terms represented within Cnidaria, Medusozoa terms were also associated with response to stimuli and the nervous system. Unique terms seemingly important to medusozoan biology were those related to wound healing and tissue migration, as well as apoptotic signaling regulation. These terms are possibly associated with unique asexual reductive traits (e.g., budding, fission, strobilation) seen within the medusozoan lineage. Despite initially identifying 1,091

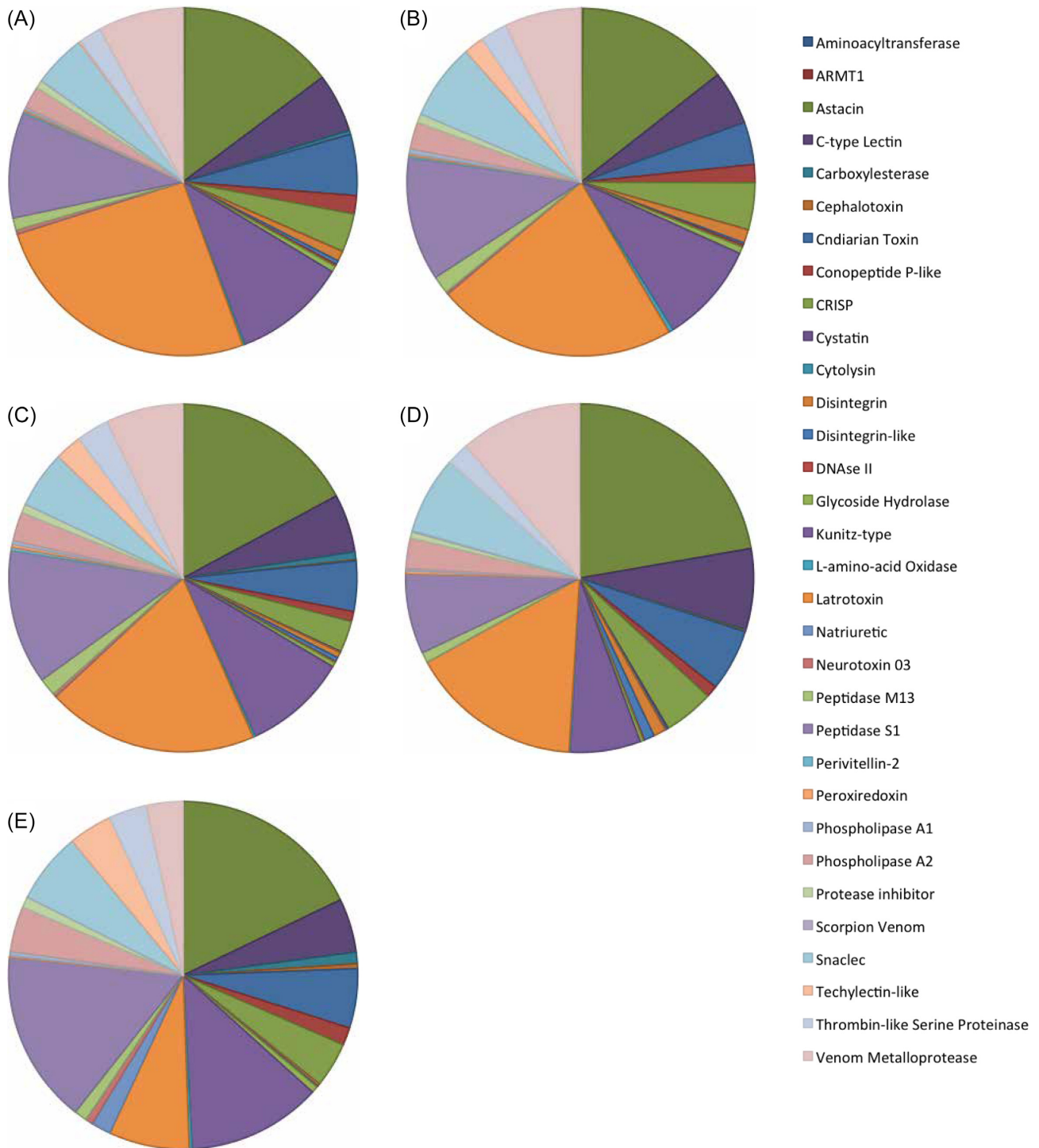


Figure 6: Venom-encoding gene repertoire of 5 cnidarian genomes. Venom-encoding genes were identified with the venomix database using OrthoFinder and BLAST. A, *Calvadosia cruxmelitensis*; B, *Alatina alata*; C, *Cassiopea xamachana*; D, *Hydra magnipapillata*; and E, *Nematostella vectensis*.

orthogroups unique to Acraspeda, only 14 GO terms were enriched (highly abundant); this number was further reduced to 8 non-redundant ReViGO terms (Fig. 5, Table S3). The apparently low enrichment may reflect an under-representation of Acraspeda genes within the reference database. Interestingly, half of the terms were associated with DNA recombination.

Venom Analysis

We identified potential venom-encoding genes within the cnidarian transcriptomes using the venomix database (a publicly available curated set of 6,622 venom-related proteins) and associated pipeline [95]. Additional venom-encoding genes were identified by BLASTing the >6,000 venom-related protein sequences of the venomix database to the Augustus protein pre-

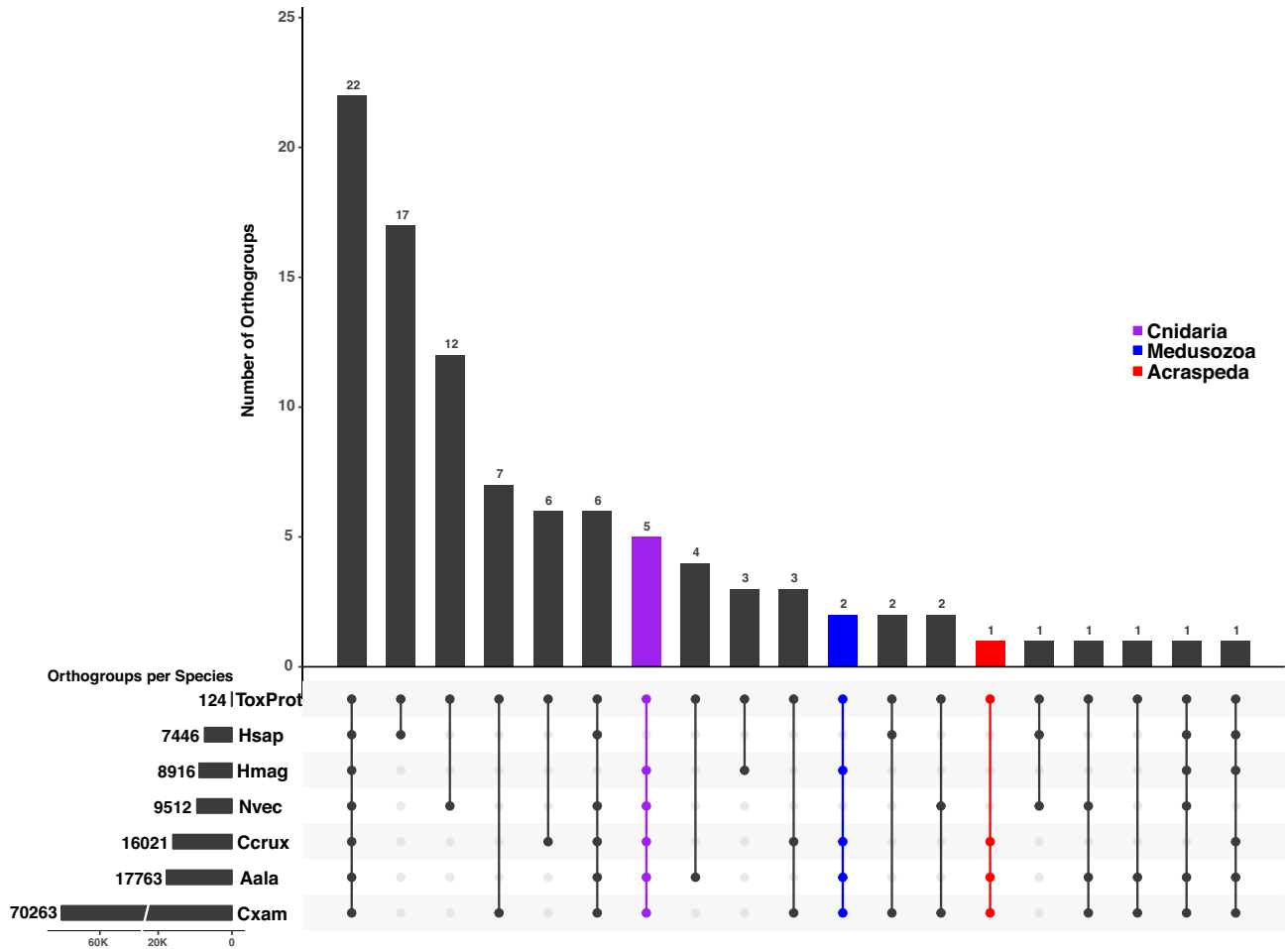


Figure 7: Distribution of venom-related genes in cnidarian lineages. Filled circles in the bottom panel indicate presence of shared venom-related genes in those lineages. Bar graphs indicate the number of venom-related orthogroups corresponding to each filled-circle pattern. Numbers next to each species abbreviation indicate the total number of venom-related orthogroups identified for that species. Hsap = *Homo sapiens*; Nvec = *Nematostella vectensis*; Hmag = *Hydra magnipapillata*; Ccrux = *Calvadosia cruxmelitensis*; Aala = *Alatina alata*; Cxam = *Cassiopea xamachana*.

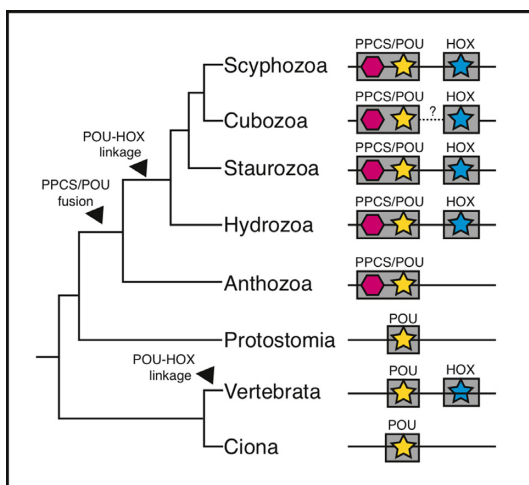


Figure 8: Evolution of PPCS/POU gene fusion and POU-Hox linkage. A fusion event involving a POU and PPCS domain occurred in the stem of Cnidaria. The syntenic linkage of POU and Hox genes occurred at least twice in animal evolution: once in the stem of Medusozoa and once in the vertebrate lineage.

dictions (BLAST v2.2.31+ e -value = 10^{-6}) [96]. By combining the results of both approaches, we identified 93 types of venom-encoding genes in *C. cruxmelitensis*, 93 in *A. alata*, 97 in *C. xamachana*, 96 in *H. magnipapillata*, and 91 in *N. vectensis*. In total, we identified 117 types of putative venom proteins, organized into 32 families, that were present in ≥ 1 of the 5 cnidarian taxa (Fig. 6, Table S4). To attempt to reconstruct evolutionary relationships among venom proteins in cnidarians, we added the venom database to our initial set of input protein sequences and reran our OrthoFinder pipeline. Using this process, we identified 124 orthogroups encoding venom genes in the 5 cnidarian genomes and the human genome (Fig. 7). Of the 124 venom orthogroups, few were found to be specific to any 1 cnidarian lineage, with 5 orthogroups present across all cnidarians, 2 spanning medusozoans, and 1 shared between Acraspeda. Most of the proteins in the venom database were identified first in bilaterian animals, and properly curated based on extensive supporting data, whereas putative toxins identified in non-model cnidarians often lack robust evidence to support annotations, precluding their entry into curated databases; hence the limited number of proteins returned in our homology search. However, we were successful in identifying 9 cnidarian-specific toxin proteins [4, 97–100]. Four of these proteins (potassium channel toxin

BcsTx, peptide toxin Am-1, AvTx, MsepPTx) were found exclusively in *N. vectensis*, while CqTx was exclusive to the genome of *A. alatina*. Interestingly, CrTx, a toxin previously identified in Cubozoa, including *A. alatina* [99, 101], was also found in *C. xamachana* and *H. magnipapillata*. While the genome of *A. alatina* appears to possess 5 copies of the CrTx gene, we identified 3 putative copies in *C. xamachana* and 1 copy in *H. magnipapillata* genomes, respectively. However, CrTx was absent from both the *C. cruxmelitensis* genome and transcriptome, suggesting that this gene may have been lost from the staurozoan lineage. Additionally, the pore-forming toxins sticholysin and hydralysin [97, 102] were found only in the *H. magnipapillata* genome in our analysis. Sticholysin was originally identified in the sea anemone *Stichodactyla helianthus*, but its absence from the *N. vectensis* genome may indicate that it is not an Anthozoa-specific protein, but rather variably distributed in Cnidaria.

Hox-POU Synteny Analysis

In the hydrozoan *Eleuthera dichotoma*, a POU6-class homeobox gene is fused with a phosphopantothenoylcysteine-synthetase (PPCS), and this PPCS/POU6 fusion is linked to a Hox class homeobox gene, *Cnox5* [48]. The PPCS/POU6 fusion is known only in Cnidaria, and its presence in the anthozoan *N. vectensis* suggests that it was likely present in the last common cnidarian ancestor. On the other hand, PPCS/POU6 fusion is not linked to a Hox-class gene in *N. vectensis* (cf. Putnam et al. 2007 assembly [55]), suggesting that POU-Hox linkage might be a more recent event. Of the Acraspeda genomes we find the PPCS/POU6 fusion gene linked to a *Cnox5* ortholog in *C. cruxmelitensis* and *C. xamachana* (Fig. 8, Fig. S3). Considering that the last common medusozoan ancestor likely >500 million years ago [103], it is reasonable to conclude that a functional constraint has led to conserved synteny for PPCS/POU6 fusion. However, the linkage to *Cnox5* was not recovered in the *A. alata* genome, preventing further speculation about whether POU-Hox linkage was present in the last common acraspedan ancestor.

Based on the well-established linkage of a POU-class homeobox gene to Hox clusters in vertebrates [48], it had been suggested that a POU-Hox linkage may have been present in the last common ancestor of cnidarians and bilaterians. To check this, we searched several additional anthozoan genomes: *Stylophora pistillata* [54] and *Acropora digitifera* [104], as well as several invertebrate bilaterian genomes: *Capitella teleta* (Polychaeta) [105], *Strigamia maritima* (Chilopoda) [106], *Octopus bimaculoides* (Cephalopoda) [107], *Mizuhopecten yessoensis* (Bivalvia) [108], and *Ciona intestinalis* (Ascidiacea) [109] that were not available at the time of the original study. We found no evidence for ancient POU-Hox synteny in these anthozoans nor in the invertebrate bilaterian genomes, suggesting that the POU-Hox linkage in medusozoans was achieved independently from the vertebrate POU-Hox linkage (Fig. 8, Fig. S3). These findings demonstrate how the 3 new medusozoan genomes allow us to address questions pertaining to molecular evolution, as well as the synergistic benefit of increased genomic-level taxon sampling when testing hypotheses about ancestral states.

Conclusions

In this note we describe draft genomes for 3 species of the medusozoan sub-group Acraspeda (Cnidaria)—*C. cruxmelitensis* (Staurozoa), *A. alata* (Cubozoa), and *C. xamachana* (Scyphozoa)—and our corresponding bioinformatics workflows for their assem-

blies and partial annotations. The findings of our preliminary orthology analyses and annotation of Hox-linked and venom-related genes provide a glimpse into genetic components underlying the evolution of certain traits in these early metazoans. Coupled with appropriate bioinformatics tools and data management pipelines, researchers across a broad range of scientific fields can utilize these resources to investigate the genetic basis of defense, reproduction, and communication in this ancient and species-rich group that encompasses a diversity of life histories, some of which exhibit pelagic life stages. Furthermore, cnidarian genomes offer strategic opportunities to investigate possible genetic links to any number of ecological issues related to jellyfish that are frequently reported in the scientific literature, or in news media.

These medusozoan genomes will be useful resources in developing functional constructs (e.g., CRISPR/Cas9 guide RNAs) that can be used to understand the genomic basis for some of the captivating biological innovations of these animals, and eventually for the design of probes for target-capture DNA sequencing. Last, the availability of these genomic-level sequence data is an important step forward in the pursuit to elucidate evolutionary events that may have shaped Medusozoa, and in reconstructing the last common ancestor of Cnidaria and Bilateria. Therefore, we are confident that these new genomes will prove valuable for understanding the biology of these fascinating creatures, and for exploring key genomic events that were formative in the early evolution of animals.

Availability of supporting data and materials

Accession numbers for raw sequencing reads and assemblies are available in Table 1. Custom scripts and parameters used for the analyses are available in a github repository [74]. Other data supporting this work are available in the GigaScience repository, GigaDB [110].

Additional files

Supplementary Figure S1. ReViGO output for Cnidaria genes clustered through *k*-means clustering by Euclidean distance. Number of optimal clusters predicted prior to clustering using NbClust.

Supplementary Figure S2. ReViGO output for Medusozoa genes clustered through *k*-means clustering by Euclidean distance. Number of optimal clusters predicted prior to clustering using NbClust.

Supplementary Figure S3. Linkage of PPCS-POU genes with Hox genes in cnidarian genomes. Genomic scaffolds for 3 Medusozoa lineages (*Eleuthera dichotoma*, *Calvadosia cruxmelitensis*, and *Cassiopea xamachana*) show linkage of the PPCS-POU gene linked to a Hox gene (dark green). This linkage is not seen in Anthozoa (*Nematostella vectensis*). The light green region indicates the transcribed portion of the scaffold, and exons are represented within by curved rectangles (PPCS exons = purple, POU exons = yellow). Scaffold length shown to the right of each bar. Edic: *E. dichotoma*; Ccrux: *C. cruxmelitensis*; Cxam: *C. xamachana*; Nvec: *N. vectensis*.

Table S1. Orthogroups specific to Cnidaria identified using OrthoFinder and annotated by a representative gene from the *Hydra magnipapillata* genome. Protein annotations were retrieved from Swiss-Prot.

Table S2. Orthogroups specific to Medusozoa identified using OrthoFinder and annotated by a representative gene from the

Hydra magnipapillata genome. Protein annotations were retrieved from Swiss-Prot.

Table S3. Orthogroups specific to Acraspeda identified using OrthoFinder and annotated by a representative gene from the *Hydra magnipapillata* genome. Protein annotations were retrieved from Swiss-Prot.

Table S4: Venom-encoding gene repertoire of 5 cnidarian genomes identified via the venomix database and pipeline. Venom genes are categories by families (column 1). Both genomic and transcriptomic data were used, with transcriptomic isoforms counted as a single venom-encoding gene.

Abbreviations

ABYSS: Assembly By Short Sequencing; BLAST: Basic Local Alignment Search Tool; BLAT: BLASTlike Alignment Tool; bp: base pairs; BUSCO: Benchmarking Universal Single-Copy Orthologs; CEGMA: Core Eukaryotic Genes Mapping Approach; CRISPR: clustered regularly interspaced short palindromic repeats; CTAB: cetyl trimethylammonium bromide; ENA: European Nucleotide Archive; Gb: gigabase pairs; GO: Gene Ontology; MaSuRCA: Maryland Super-Read Celera Assembler; Mb: megabases pairs; NCBI: National Center for Biotechnology Information; PacBio: Pacific Biosciences; Platanus: PLAT-form for Assembling Nucleotide Sequences; PPCS: phosphopantothienoylcysteine synthetase; ReViGO: Reduce + Visualize Gene Ontology; SPAdes: St. Petersburg genome assembler; tRNA: transfer RNA.

Competing interests

The authors declare that they have no competing interests.

Funding

The sequencing of *Alatina alata* was funded by Iridian Genomes. A.O. acknowledges funding from an NSF Dimensions Grant (OCE 1442206); C.L.A. acknowledges funding from The University of Maryland Eugenie Clark Scholarship and a fellowship from Oak Ridge Institute for Science and Education (ORISE); startup funds from the University of Florida DSP Research Strategic Initiatives #00114464 and University of Florida Office of the Provost Programs to J.F.R.

Author contributions

Alatina alata samples were collected by C.L.A., A.G.C., and S.P. *Alatina alata* assembly: R.B.D. and C.L.A. with B.B. and S.L.; *C. xamachana* assembly: A.O. and J.F.R.; *Calvadosia cruxmelitensis* assembly: M.C. and J.R. *Calvadosia cruxmelitensis* mitochondrial genome assembly: E.K. Orthology analyses: A.O. and J.R. The manuscript was drafted by A.O. with substantial contributions from C.L.A., A.G.C., and J.F.R. All coauthors read and provided feedback on the final manuscript. S.P., A.G.C., M.M., and J.R. oversaw the project from start to finish.

References

- Lewis Ames C. Medusa: a review of an ancient cnidarian body form. In: Kloc M, Kubiak JZ, eds. *Marine Organisms as Model Systems in Biology and Medicine Results and Problems in Cell Differentiation*. Cham, Switzerland: Springer; 2018:105–36.
- Brinkman DL, Burnell JN. Biochemical and molecular characterisation of cubozoan protein toxins. *Toxicon* 2009;54(8):1162–73.
- Jouiaei M, Sunagar K, Federman Gross A, et al. Evolution of an Ancient venom: recognition of a novel family of cnidarian toxins and the common evolutionary origin of sodium and potassium neurotoxins in sea anemone. *Mol Biol Evol* 2015;32(6):1598–610.
- Jouiaei M, Yanagihara AA, Madio B, et al. Ancient venom systems: a review on cnidaria toxins. *Toxins (Basel)* 2015;7(6):2251–71.
- Coates MM. Visual ecology and functional morphology of cubozoa (cnidaria). *Integr Comp Biol* 2003;43(4):542–8.
- Liebertova M, Pergner J, Kozmikova I, et al. Cubozoan genome illuminates functional diversification of opsins and photoreceptor evolution. *Sci Rep* 2015;5:11885.
- Lewis Ames C, Ryan JF, Bely AE, et al. A new transcriptome and transcriptome profiling of adult and larval tissue in the box jellyfish *Alatina alata*: an emerging model for studying venom, vision and sex. *BMC Genomics* 2016;17:650.
- Collins AG. Phylogeny of Medusozoa and the evolution of cnidarian life cycles. *J Evol Biol* 2002;15:418–32.
- Zapata F, Goetz FE, Smith SA, et al. Phylogenomic analyses support traditional relationships within cnidaria. *PLoS One* 2015;10(10):e0139068.
- Marques AC, Collins AG. Cladistic analysis of Medusozoa and cnidarian evolution. *Invertebr Biol* 2004;123(1):23–42.
- Kayal E, Bentlage B, Sabrina Pankey M, et al. Phylogenomics provides a robust topology of the major cnidarian lineages and insights on the origins of key organismal traits. *BMC Evol Biol* 2018;18(1):68.
- Holzer AS, Bartosova-Sojkova P, Born-Torrijos A, et al. The joint evolution of the Myxozoa and their alternate hosts: a cnidarian recipe for success and vast biodiversity. *Mol Ecol* 2018;27(7):1651–66.
- Siddall ME, Martin DS, Bridge D, et al. The demise of a phylum of protists: phylogeny of Myxozoa and other parasitic cnidaria. *J Parasitol* 1995;81(6):961–7.
- Holland JW, Okamura B, Hartikainen H, et al. A novel mini-collagen gene links cnidarians and myxozoans. *Proc Biol Sci* 2011;278(1705):546–53.
- Chang ES, Neuhof M, Rubinstein ND, et al. Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. *Proc Natl Acad Sci U S A* 2015;112(48):14912–7.
- Bridge D, Cunningham CW, Schierwater B, et al. Class-level relationships in the phylum Cnidaria: evidence from mitochondrial genome structure. *Proc Natl Acad Sci U S A* 1992;89:8750–3.
- Reft AJ, Daly M. Morphology, Distribution, and evolution of apical structure of nematocysts in hexacorallia. *J Morphol* 2012;273:121–36.
- von Salvini-Plawen L. On the origin and evolution of the lower Metazoa. *Zool Syst Evol* 1978;16:40–88.
- Ortman BD, Bucklin A, Pagès F, et al. DNA barcoding the Medusozoa using mtCOI. *Deep Sea Res Part II* 2010;57(24–26):2148–56.
- Miranda LS, Hirano YM, Mills CE, et al. Systematics of stalked jellyfishes (Cnidaria: Staurozoa). *PeerJ* 2016;4:e1951.
- Kayal E, Roure B, Philippe H, et al. Cnidarian phylogenetic relationships as revealed by mitogenomics. *BMC Evol Biol* 2013;13(5):1.
- Cartwright P, Nawrocki AM. Character evolution in Hydrozoa (phylum Cnidaria). *Integr Comp Biol* 2010;50(3):456–72.

23. Toshino S, Miyake H, Ohtsuka S, et al. Monodisc strobilation in Japanese giant box jellyfish *Morbakka virulenta* (Kishinouye, 1910): a strong implication of phylogenetic similarity between Cubozoa and Scyphozoa. *Evol Dev* 2015;17(4):231–9.
24. Helm RR. Evolution and development of scyphozoan jellyfish. *Biol Rev* 2018;93(2):1228–50.
25. Miranda LS, Collins AG, Hirano YM, et al. Comparative internal anatomy of Staurozoa (Cnidaria), with functional and evolutionary inferences. *PeerJ* 2016;4:e2594.
26. Kikinger R, von Salvini-Plawen L. Development from polyp to stauromedusa in *stylocoronella* (Cnidaria: Scyphozoa). *J Mar Biol Assoc U.K.* 2009;75(04):899.
27. Miranda LS, Mills CE, Hirano YM, et al. A review of the global diversity and natural history of stalked jellyfishes (Cnidaria, Staurozoa). *Mar Biodivers* 2018;48(4):1695–714.
28. Boero F, Bouillon J, Piraino S, et al. Diversity of hydroidomedusan life cycles: ecological implications and evolutionary patterns. In: den Hartog JC, van Ofwegen LP, van der Spoel S, eds. *Proceedings of the 6th International Conference on Coelenterate Biology*, Noordwijkerhout, the Netherlands, 1995. Leiden, The Netherlands: Nationaal Natuurhistorisch Museum; 1997:56–62.
29. Bentlage B, Osborn KJ, Lindsay DJ, et al. Loss of metagenesis and evolution of a parasitic life style in a group of open ocean jellyfish. *Mol Phylogenet Evol* 2018;123:50–9.
30. Reber-Muller S, Streitwolf-Engel R, Yanze N, et al. BMP2/4 and BMP5-8 in jellyfish development and transdifferentiation. *Int J Dev Biol* 2006;50(4):377–84.
31. Kraus JE, Fredman D, Wang W, et al. Adoption of conserved developmental genes in development and origin of the medusa body plan. *EvoDevo* 2015;6:23.
32. Fuchs B, Wang W, Graspeuntner S, et al. Regulation of polyp-to-jellyfish transition in *Aurelia aurita*. *Curr Biol* 2014;24(3):263–73.
33. Brekhman V, Malik A, Haas B, et al. Transcriptome profiling of the dynamic life cycle of the scyphozoan jellyfish *Aurelia aurita*. *BMC Genomics* 2015;16:74. doi:10.1186/s12864-015-1320-z.
34. Ge J, Liu C, Tan J, et al. Transcriptome analysis of scyphozoan jellyfish *Rhopilema esculentum* from polyp to medusa identifies potential genes regulating strobilation. *Dev Genes Evol* 2018;228(6):243–54.
35. Yamakawa S, Morino Y, Honda M, et al. The role of retinoic acid signaling in starfish metamorphosis. *EvoDevo* 2018;9(1), doi:10.1186/s13227-018-0098-x.
36. Ryan JF, Mazza ME, Pang K, et al. Pre-bilaterian origins of the Hox cluster and the Hox code: evidence from the sea anemone, *Nematostella vectensis*. *PLoS One* 2007;2(1):e153.
37. Ryan JF, Burton PM, Mazza ME, et al. The cnidarian-bilaterian ancestor possessed at least 56 homeoboxes: evidence from the starlet sea anemone, *Nematostella vectensis*. *Genome Biol* 2006;7(7):R64.
38. Chourrout D, Delsuc F, Chourrout P, et al. Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. *Nature* 2006;442(7103):684–7.
39. Slack JM, Holland PW and Graham CF. The zootype and the phylotypic stage. *Nature* 1993;361(6412):490–2.
40. Finnerty JR, Pang K, Burton P, et al. Origins of bilateral symmetry: Hox and dpp expression in a sea anemone. *Science* 2004;304(5675):1335–7.
41. He S, Del Viso F, Chen CY, et al. An axial Hox code controls tissue segmentation and body patterning in *Nematostella vectensis*. *Science* 2018;361(6409):1377–80.
42. Schummer M, Scheurlen I, Schaller C, et al. HOM/HOX homeobox genes are present in hydra (*Chlorohydra viridissima*) and are differentially expressed during regeneration. *EMBO J* 1992;11(5):1815–23.
43. Gauchat D, Mazet F, Berney C, et al. Evolution of Antp-class genes and differential expression of Hydra Hox/paraHox genes in anterior patterning. *Proc Natl Acad Sci U S A* 2000;97(9):4493–8.
44. Chiori R, Jager M, Denker E, et al. Are Hox genes ancestrally involved in axial patterning? Evidence from the hydrozoan *Clytia hemisphaerica* (Cnidaria). *PLoS One* 2009;4(1):e4231.
45. Shenk MA, Bode HB, Steele RE. Expression of Cnox-2, a HOM:HOX homeobox gene in hydra, is correlated with axial pattern formation. *Development* 1993;117:657–67.
46. Reddy PC, Unni MK, Gungi A, et al. Evolution of Hox-like genes in Cnidaria: study of Hydra Hox repertoire reveals tailor-made Hox-code for Cnidarians. *Mech Dev* 2015;138(Pt 2):87–96.
47. Kuhn K, Streit B, Schierwater B. Isolation of Hox genes from the Scyphozoan *Cassiopeia xamachana*: implications for the early evolution of Hox genes. *J Exp Zool* 1999;285:63–75.
48. Kamm K, Schierwater B. Ancient linkage of a POU class 6 and an anterior Hox-like gene in cnidaria: implications for the evolution of homeobox genes. *J Exp Zool* 2007;308(6):777–84.
49. Chapman JA, Kirkness EF, Simakov O, et al. The dynamic genome of Hydra. *Nature* 2010;464(7288):592–6.
50. Shinzato C, Mungpakdee S, Satoh N, et al. A genomic approach to coral-dinoflagellate symbiosis: studies of *Acropora digitifera* and *Symbiodinium minutum*. *Front Microbiol* 2014;5:336.
51. Cunning R, Bay RA, Gillette P, et al. Comparative analysis of the *Pocillopora damicornis* genome highlights role of immune system in coral evolution. *Sci Rep* 2018;8(1):16134.
52. Jiang J, Quattrini AM, Francis WR, et al. A hybrid *de novo* assembly of the sea pansy (*Renilla muelleri*) genome. *GigaScience* 2019;8(4):giz026, doi.org/10.1093/gigascience/giz026.
53. Ying H, Cooke I, Sprungala S, et al. Comparative genomics reveals the distinct evolutionary trajectories of the robust and complex coral lineages. *Genome Biol* 2018;19(1):175.
54. Voolstra CR, Li Y, Liew YJ, et al. Comparative analysis of the genomes of *Stylophora pistillata* and *Acropora digitifera* provides evidence for extensive differences between species of corals. *Sci Rep* 2017;7(1):17583.
55. Putnam NH, Srivastava M, Hellsten U, et al. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 2007;317(5834):86–94.
56. Leclère L. The genome of the jellyfish *Clytia hemisphaerica* and the evolution of the cnidarian life-cycle. *Nat Ecol Evol* 2019;3:801–10.
57. Gold DA, Katsuki T, Li Y, et al. The genome of the jellyfish *Aurelia* and the evolution of animal complexity. *Nat Ecol Evol* 2019;3:96–104.
58. Kim HM, Weber JA, Lee N, et al. The genome of the giant *Nomura's* jellyfish sheds light on the early evolution of active predation. *BMC Biol* 2019;17(1):28.
59. Khalturin K, Shinzato C, Khalturina M, et al. Medusozoan genomes inform the evolution of the jellyfish body plan. *Nat Ecol Evol* 2019;3(5):1.
60. Nastav B, Malej M, Malej A, Jr, et al. Is it possible to determine the economic impact of jellyfish outbreaks on fisheries? A case study – Slovenia. *Mediterr Mar Sci* 2013;14(1):214.

61. Purcell JE, Arai MN. Interactions of pelagic cnidarians and ctenophores with fish: a review. *Hydrobiologia* 2001;451(1–3):27–44.
62. Purcell JE, Uye S, Lo W. Anthropogenic causes of jellyfish blooms and their direct consequences for humans: a review. *Mar Ecol Prog Ser* 2007;350:153–74.
63. Lawley JW, Ames CL, Bentlage B, et al. Box jellyfish *Alatina alata* has a circumtropical distribution. *Biol Bull* 2016;231(2):152–69.
64. Garcia-Rodriguez J, Lewis Ames C, Marian J, et al. Gonadal histology of box jellyfish (Cnidaria: Cubozoa) reveals variation between internal fertilizing species *Alatina alata* (Alatinidae) and *Copula sivickisi* (Tripedaliidae). *J Morphol* 2018;279(6):841–56.
65. Corw GL, Chiaverano LM, Crites J, et al. Box jellyfish (Cubozoa: Carybdeida) in Hawaiian waters, and the first record of *Tripedalia cystophora* in Hawai'i. *Bishop Mus Bull Zool* 2015;9:93–108.
66. Lampert KP. Cassiopea and its zooxanthellae. In: Goffredo S, Dubinsky Z, eds. *The Cnidaria, Past, Present and Future: The world of Medusa and her sisters*. Switzerland: Springer; 2016:415–23.
67. Ohdera AH, Abrams MJ, Ames CL, et al. Upside-down but headed in the right direction: review of the highly versatile *Cassiopea xamachana* system. *Front Ecol Evol* 2018;6, doi:10.3389/fevo.2018.00035.
68. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30(15):2114–20.
69. Gnerre S, MacCallum I, Przybylski D, et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A* 2011;108(4):1513–8.
70. Ryan JF. FastqSifter. 2015. <https://github.com/josephryan/FastqSifter>. Accessed on 10 June 2019.
71. Simpson JT, Wong K, Jackman SD, et al. ABySS: a parallel assembler for short read sequence data. *Genome Res* 2009;19(6):1117–23.
72. Bankevich A, Nurk S, Antipov D, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19(5):455–77.
73. Kajitani R, Toshimoto K, Noguchi H, et al. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res* 2014;24(8):1384–95.
74. Ohdera A, Ryan JF. Ohdera.et.al.2018. 2010. <https://github.com/josephryan/Ohdera.et.al.2018>. Accessed on 10 June 2019.
75. Parra G, Bradnam K, Korf I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* 2007;23(9):1061–7.
76. Simao FA, Waterhouse RM, Ioannidis P, et al. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 2015;31(19):3210–2.
77. Nishimura O, Hara Y, Kuraku S. gVolante for standardizing completeness assessment of genome and transcriptome assemblies. *Bioinformatics* 2017;33(22):3635–7.
78. Kayal E, Bentlage B, Cartwright P, et al. Phylogenetic analysis of higher-level relationships within Hydroidolina (Cnidaria: Hydrozoa) using mitochondrial genome data and insight into their mitochondrial transcription. *PeerJ* 2015;3:e1403.
79. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 1997;25(5):955–64.
80. Laslett D, Canback B. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* 2008;24(2):172–5.
81. Ryan JF. matemaker. 2015. <https://github.com/josephryan/matemaker>. Accessed on 10 June 2019.
82. Boetzer M, Henkel CV, Jansen HJ, et al. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 2011;27(4):578–9.
83. Zimin AV, Marcais G, Puiu D, et al. The MaSuRCA genome assembler. *Bioinformatics* 2013;29(21):2669–77.
84. Vinogradov AE. Intron-genome size relationship on a large evolutionary scale. *J Mol Evol* 1999;49(3):376–84.
85. Grau JH, Hackl T, Koepfli KP, et al. Improving draft genome contiguity with reference-derived in silico mate-pair libraries. *Gigascience* 2018;7(5):giy029, doi:10.1093/gigascience/giy029.
86. Miller DE, Staber C, Zeitlinger J, et al. Highly contiguous genome assemblies of 15 *Drosophila* species generated using Nanopore sequencing. *G3 (Bethesda)* 2018;8(10):3131–41.
87. Stanke M, Waack S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* 2003;19(Suppl 2):ii215–ii25.
88. Kent WJ. BLAT – the BLASTlike Alignment Tool. *Genome Res* 2002;12:656–64.
89. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 2015;16:157.
90. Grabherr MG, Haas BJ, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 2011;29(7):644–52.
91. Haas BJ, Papanicolaou A, Yassour M, et al. De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. *Nat Protoc* 2013;8(8):1494–512.
92. Boutet E, Lieberherr D, Tognolli M, et al. UniProtKB/Swiss-Prot. In: Edwards D, ed. *Plant Bioinformatics: Methods and Protocols*. Totowa, NJ: Humana Press; 2007:89–112.
93. Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012;16(5):284–7.
94. Carlson M, Pagès H. AnnotationForge: code for building annotation database packages. R package version 1240. 2018, doi:10.18129/B9.bioc.AnnotationForge.
95. Macrander J, Panda J, Janies D, et al. Venomix: a simple bioinformatic pipeline for identifying and characterizing toxin gene candidates from transcriptomic data. *PeerJ* 2018;6:e5361.
96. Camacho C, Coulouris G, Avagyan V, et al. BLAST+: architecture and applications. *BMC Bioinformatics* 2009;10:421.
97. Sher D, Fishman Y, Zhang M, et al. Hydralysins, a new category of beta-pore-forming toxins in cnidaria. *J Biol Chem* 2005;280(24):22847–55.
98. Oliveira JS, Fuentes-Silva D, King GF. Development of a rational nomenclature for naming peptide and protein toxins from sea anemones. *Toxicol* 2012;60(4):539–50.
99. Nagai H, Takuwa-Kuroda K, Nakao M, et al. A novel protein toxin from the deadly box jellyfish (sea wasp, habu-kurage) *Chiropsalmus quadrigatus*. *Biosci Biotechnol Biochem* 2002;66(1):97–102.
100. Nagai H, Takuwa K, Nakao M, et al. Novel proteinaeous toxins from the box jellyfish (sea wasp) *Caryb-*

- dea rastonii*. *Biochem Biophys Res Commun* 2000;275(2):582–8.
101. Nagai H, Takuwa K, Nakao M, et al. Isolation and characterization of a novel protein toxin from the Hawaiian box jellyfish (sea wasp) *Carybdea alata*. *Biochem Biophys Res Commun* 2000;275(2):589–94.
 102. Pedrera L, Fanani ML, Ros U, et al. Sticholysin I-membrane interaction: an interplay between the presence of sphingomyelin and membrane fluidity. *Biochim Biophys Acta* 2014;1838(7):1752–9.
 103. Rogers AD. Cnidarians (Cnidaria). In: Hedges SB, Kumar S, eds. *The Timetree of Life*. New York, NY, USA: Oxford University Press; 2009:233–8.
 104. Shinzato C, Shoguchi E, Kawashima T, et al. Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature* 2011;476(7360):320–3.
 105. Simakov O, Marletaz F, Cho SJ, et al. Insights into bilaterian evolution from three spiralian genomes. *Nature* 2013;493(7433):526–31.
 106. Chipman AD, Ferrier DE, Brena C, et al. The first myriapod genome sequence reveals conservative arthropod gene content and genome organisation in the centipede *Strigamia maritima*. *PLoS Biol* 2014;12(11):e1002005.
 107. Albertin CB, Simakov O, Mitros T, et al. The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature* 2015;524(7564):220–4.
 108. Wang S, Zhang J, Jiao W, et al. Scallop genome provides insights into evolution of bilaterian karyotype and development. *Nat Ecol Evol* 2017;1(5):120.
 109. Dehal P, Satou Y, Campbell RK, et al. The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* 2002;298(5601):2157.
 110. Ohdera AH, Ames CL, Dikow RB, et al. Supporting data for “Boxed, stalked and upside-down? Draft genomes from diverse jellyfish (Cnidaria, Acraspeda) lineages: *Alatina alata* (Cubozoa), *Calvadosia cruxmelitensis* (Staurozoa), and *Cassiopea xamachana* (Scyphozoa).” *GigaScience Database* 2019. <http://dx.doi.org/10.5524/100604>.