ABSTRACT

Title of Dissertation: TAXONOMY, MORPHOLOGY, AND RNA-SEQ TRANSCRIPTOMICS OF THE CUBOZOAN *ALATINA ALATA*, AN EMERGING MODEL CNIDARIAN

Cheryl L Ames, Doctor of Philosophy 2016

Dissertation directed by: Associate Professor Alexandra E. Bely, Biology Department
Adjunct Professor Allen G. Collins, Biological Sciences Graduate Program

Cnidarians are often considered simple animals, but the more than 13,000 estimated species (e.g., corals, hydroids and jellyfish) of the early diverging phylum exhibit a broad diversity of forms, functions and behaviors, some of which are demonstrably complex. In particular, cubozoans (box jellyfish) are cnidarians that have evolved a number of distinguishing features. Some cubozoan species possess complex mating behaviors or particularly potent stings, and all possess well-developed light sensation involving image-forming eyes. Like all cnidarians, cubozoans have specialized subcellular structures called nematocysts that are used in prey capture and defense.

The objective of this study is to contribute to the development of the box jellyfish *Alatina alata* as a model cnidarian. This cubozoan species offers numerous advantages for investigating morphological and molecular traits underlying complex processes and coordinated behavior in free-living medusozoans (i.e., jellyfish), and more broadly throughout Metazoa.

First, I provide an overview of Cnidaria with an emphasis on the current understanding of genes and proteins implicated in complex biological processes in a few select cnidarians. Second, to further develop resources for *A. alata*, I provide a formal redescription of this cubozoan and establish a neotype specimen voucher, which serve to stabilize the taxonomy of the species. Third, I generate the first functionally annotated transcriptome of adult and larval *A. alata* tissue and apply preliminary differential expression analyses to identify candidate genes implicated broadly in biological processes related to prey capture and defense, vision and the phototransduction pathway and sexual reproduction and gametogenesis. Fourth, to better understand venom diversity and mechanisms controlling venom synthesis in *A. alata*, I use bioinformatics to investigate gene candidates with dual roles in venom and digestion, and review the biology of prey capture and digestion in cubozoans.

The morphological and molecular resources presented herein contribute to understanding the evolution of cubozoan characteristics and serve to facilitate further research on this emerging cubozoan model.

TAXONOMY, MORPHOLOGY, AND RNA-SEQ TRANSCRIPTOMICS OF THE CUBOZOAN *ALATINA ALATA*, AN EMERGING MODEL CNIDARIAN

by

Cheryl L Ames

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2016

Advisory Committee:
Associate Professor Alexandra E. Bely, Chair
Adjunct Professor Allen G Collins
Associate Professor Eric Haag
Special Member Dr. Paulyn Cartwright
Professor Najib El-Sayed (Dean's Representative)

UNIVERSITY OF MARYLAND
BIOLOGICAL SCIENCES GRADUATE PROGRAM

2101 Bioscience Research Building
College Park, Maryland 20742-4415
301.405.6905/6991 TEL, 301.314.9921 FAX

Alexander Chen, Ph.D.
Associate Dean of the Graduate School
University of Maryland

Dear Dr. Chen,
This letter is written to signify that the dissertation committee, committee chair, and the graduate director have all approved the use of previously published co-authored work in the final dissertation of Cheryl Ames, Biological Sciences Graduate Program, UID 111757596. In accordance with the Graduate School's policy the dissertation committee has determined that they made substantial contributions to the included work.

The citations for the published works are:
Chapter 2 (Published)
Lewis, C., Bentlage, B., Yanagihara, A.A., Gillan, W., Blerk, J.V.A.N., Keil, D.P., Bely, A.E. & Collins, A.G. (2013) Redescription of *Alatina alata* (Reynaud, 1830) (Cnidaria: Cubozoa) from Bonaire, Dutch Caribbean. Zootaxa 3737, 473–487.

Chapter 3 (In Press with imminent release)
Lewis Ames, C., Collins, A.G., Ryan, J.F., Bely, A.E. & Cartwright, P. (In Press) 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.

Per Graduate School policy the dissertation forward will identify the scope and nature of the student's contributions to the jointly authored work included in the dissertation and a copy of this letter will be submitted with the dissertation.

Sincerely,

Dr. Alexa Bely, Dissertation Committee Chair, Professor, Department of Biology

Dr. Allen Collins, Dissertation Committee Co-Chair, Zoologist, NMFS, NOAA

Dr. Michelle Brooks, Associate Director, Biological Sciences Graduate Program

Cheryl Ames, Graduate Student, Biological Sciences

From: Zhi-Qiang Zhang
To: Cheryl Lynn Ames
Subject: RE: permission to reprint Zootaxa 3737 (4): 473–487 in dissertation
Date: Tuesday, July 5, 2016 7:20:43 PM

Permission to print the above referenced material granted by:
Name/Title: ___Zhi-Qiang Zhang_____
Conditions: __Attribution of copyrighted material by citing the original paper and also indicated "reproduced here with copyright owner with permission_____
Signature: ___Zhi-Qiang Zhang_____

From: Cheryl Lynn Ames [mailto:clames1@umd.edu]
Sent: Wednesday, 6 July 2016 6:17 a.m.
To: Zhi-Qiang Zhang
Subject: permission to reprint Zootaxa 3737 (4): 473–487 in dissertation

Attention: Magnolia Press

Dear Dr. Zhang,

I request permission to reprint the following material from your publication Lewis *et al.* 2013 in Zootaxa 3737 (4): 473–487:

Lewis, C., Bentlage, B., Yanagihara, A.A., Gillan, W., Blerk, J.V.A.N., Keil, D.P., Bely, A.E. & Collins, A.G. (2013) Redescription of Alatina alata (Reynaud, 1830) (Cnidaria: Cubozoa) from Bonaire, Dutch Caribbean. 3737, 473–487.

I have attached a copy of the article for your convenience.

The material will be printed in my dissertation thesis at the University of Maryland, College Park Biological Science Graduate Program entitled: "USING HIGH THROUGHPUT SEQUENCING TO ELUCIDATE MOLECULAR PROCESSES OF VENOM, VISION AND SEX IN THE AGGREGATING BOX JELLYFISH *ALATINA ALATA*"

I anticipate the dissertation will be published in August 2016.
To provide permission, please complete the bottom of this message and return it to me via email to clames1@umd.edu before July 20, 2016.

I look forward to hearing from you at your earliest convenience.

Sincerely,

Chery l Lewis Ames
PhD Candidate
Biological Sciences Graduate Program

University of Maryland, College Park, USA

_____

Permission to print the above referenced material granted by:
Name/Title: _____
Conditions: _____
Signature:

_____

_____

Cheryl Lewis Ames
Ann G. Wylie Predoctoral Fellow
Biological Sciences Graduate Program, BEES Concentration, University of Maryland, College Park, MD
& National Museum of Natural History, Smithsonian Institution, Washington DC
Twitter: @boxjellytalk
Department of Invertebrate Zoology No Bones Blog: http://nmnh.typepad.com/no_bones/
Email: amesc@ si.edu or clames1@ umd.edu
[* *]
/ / \ \

Dedication

To my guys Andy and Kaito. Thank you for your unfailing love and support.

Acknowledgements

I am grateful to Alexa Bely and Allen Collins for their steadfast support and invaluable advice throughout every part of this journey until I reached the goal of disseminating my findings, and to all members of my dissertation committee who have been essential in keeping me on track. Additionally, I thank the members of the Bely lab, both past and present, as well as the staff in the Department of Invertebrate Zoology at the Smithsonian National Museum of Natural History, as well as the staff of the Laboratories of Analytical Biology, Smithsonian High Performance Computing Cluster, and Museum Support Center. I am thankful for the moral and academic support provided by fellow BEES graduate students. I also thank Joseph Ryan, Jason Macrander, and Yan Wang for their valuable collaborations, as well as Stacy Pirro, Joseph Burnett and Ron Larson for their wisdom and advice on methodology and logistical support. A special thanks goes out to Rita Peachy and staff and interns of the Bonaire Research Station of The Council on International Educational Exchange (CIEE), and to Bud Gilan for introducing me to this fantastic Caribbean island paraside. This work was funded by a Peter Buck Predoctoral Fellowship at the Smithsonian National Museum of Natural History, an Ann G. Wylie Dissertation Improvement Fellowship and Graduate School Dean Summer Fellowships at the University of Maryland. Chapter 2 has been published in its entirety in Zootaxa (3737 (4): 473–487), and permission has been granted by Magnolia Press to include these materials herein. Chapter 3 is in press with BMC Genomics, and Chapter 4 is in revision for Integrative and Comparative Biology. Figure 1.1 was previously

published in PNAS (10(1): e0139068), and permission has been granted to reproduce it herein under an open access license called “CC-BY.”

Table of Contents

List of Tables

List of Figures

List of Abbreviations

USNM =National Museum of Natural History, Smithsonian Institution Collections
BH=bell height, measured from tip of bell to velarial turnover of a box jellyfish
BW=bell width, the distance measured between two adjacent rhopalia of a box jellyfish
ICZN=International Code on Zoological Nomenclature
NGS=Next Generation Sequencing

Chapter 1: Introduction and Background

Cnidaria – an ancient and innovative phylum

The early diverging phylum Cnidaria consists of the Anthozoa (i.e., corals, anemones and sea pens), the Medusozoa (Bridge *et al.* 1992; Collins *et al.* 2006; Daly *et al.* 2007; Marques & Collins 2004b), consisting of the classes Cubozoa, Scyphozoa, Staurozoa, and Hydrozoa (i.e., sessile hydroids and motile jellyfish) and the group containing *Polypodium hydriforme* and Myxozoa (Chang *et al.* 2015) (Fig 1.1). The diploblastic cnidarian body has two layers of epithelial tissue (ectoderm and endoderm) separated by an extracellular collagen-based structural matrix known as mesoglea (Hausman & Burnett 1971).

Contrary to popular belief, only a subset of cnidarians exhibits exclusively radially symmetry, while the majority of cnidarian taxa exhibits bilateral symmetry at least in one life stage, and some taxa have directional symmetry (Manuel 2009). Despite the superficial simplicity of this group, the more than 13,000 estimated species of Cnidaria have evolved diverse forms and functions. Cnidarians exhibit a broad range in their sting potency, from mild to lethal in humans (Kitatani *et al.* 2015); they possess both sessile and motile life forms, including alternation of generations (Collins 2009); they display various modes of sexual and asexual reproduction via transverse fission (Fuchs *et al.* 2014; Hofmann *et al.* 1978; Lewis & Long 2005; Schiariti *et al.* 2012); and they also demonstarte varying degrees of light-sensitivity, from extraocular "dermal sense" to visually-guided behavior (Bielecki *et al.* 2014; Desmond Ramirez *et al.* 2011).

The emergence of evolutionary novelties within the Cnidaria may have permitted their successful invasion into all types of marine habitats (Appeltans *et al.* 2012), certain freshwater ecosystems, and in the case of endoparasitic myxozoans, certain terrestrial milieu (Gruhl &

Okamura 2012; Okamura *et al.* 2015). Recent analyses support Cnidaria as the sister group to Bilateria (Dunn *et al.* 2014; Zapata *et al.* 2015), the most intensively studied group of animals, providing an excellent outgroup for understanding cnidarian biology (Fig 1.1). This study proposes to establish a new model cnidarian species that would provide a number of advantages for investigating morphological and molecular traits underlying complex processes and coordinated behavior related to prey capture and defense, vision and the phototransduction pathway and sexual reproduction and gametogenesis.

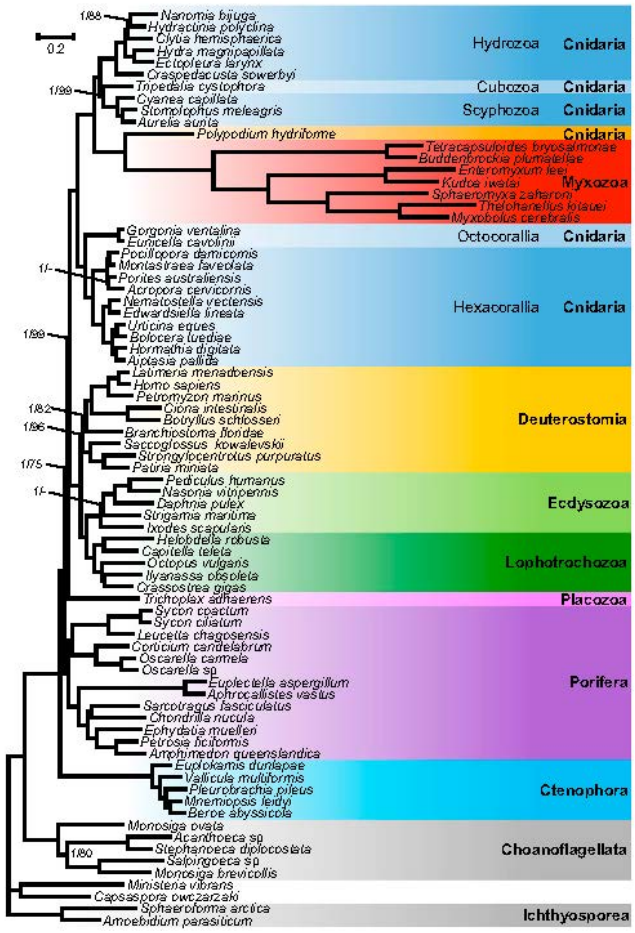

Figure 1.1 Metazoan phylogenetic tree
(Reproduced from Chang *et al.* 2015 Fig 2). Phylogenetic tree generated from a matrix of 51,940 amino acid positions and 77 taxa using Bayesian inference under the CAT model. Support values are indicated only for nodes that did not received maximal support. Bayesian posterior probabilities/ML bootstrap supports under the PROTGAMMAGTR model are given near the corresponding node.

Nematocysts - unifying the Cnidaria

The nematocyte is the unifying trait of the ancient phylum Cnidaria, which is estimated to have diverged over 650 million years ago (Cartwright & Collins 2007; Cartwright *et al.* 2007; Daly *et al.* 2007). Nematocytes are novel cell types found in all species of the early-branching phylum Cnidaria (Mariscal 1974). Comprising most of the nematocyte cytoplasm is the specialized subcellular stinging organelle – the nematocyst – which is armed with a venom-filled double-walled capsule and a harpoon-like tubule (Fig 1.2). Essential for prey capture and defense, the nematocyst is an excellent example of an evolutionary novelty that radiated throughout an entire phylum due to its adaptive benefits. However, the evolutionary origin of the nematocyst is not well understood.

Nematocysts are highly concentrated in the tentacles of most cnidarians (Gershwin 2006; Hessinger & Lenhoff 1988; Kass-Simon & Scappaticci, Jr. 2002), making the tentacles vital for prey capture and the first line of defense against potential predators. Typically, when nematocyst-laden tentacles come in contact with prey, numerous nematocysts deploy a cocktail of toxins (venom) at a rapid speed via a long spiny tubule immobilizing the target organism (Beckmann & Özbek 2012; David *et al.* 2008; Özbek *et al.* 2009). Approximately thirty different structural variants of the basic nematocyst design are known (Mariscal 1974; Östman 2000), and each cnidarian species possesses one to several types throughout its life history, suggesting that diversity of nematocyst types is an important adaptive evolutionary component in sessile and motile life styles. This inventory of nematocysts within a cnidarian species, called the "cnidome", has traditionally been characterized based on the morphology of the capsule and tubule (Weill 1934) and, more recently, the apical structure of the capsule (Reft & Daly 2012).

Nematocysts are secreted from a post-Golgi vesicle within nematocytes by proteins from the secretory pathway; hence the classification of nematocysts as "secretory organelles" (Ozbek 2011). The morphological processes involved in nematocyst development and regeneration, known as nematogenesis, have been characterized in some model cnidarians (e.g., *Hydra*) (Balasubramanian *et al.* 2012; Beckmann & Özbek 2012; David *et al.* 2008; Houliston *et al.* 2010; Ozbek 2011; Zenkert *et al.* 2011). Nematocysts are one of three major categories of such specialized organelles called cnidocysts (the other two being spirocysts and ptychocysts (Mariscal *et al.* 1977)), and the only type found in all cnidarian taxa (Daly et al. 2007). Nematocysts, commonly called "stinging cells" are thought to the ancestral cnidocyst type, and are the only kind found in the subphylum Medusozoa (Collins 2009; Marques & Collins 2004a) (see (Reft & Daly 2012) for discussion on all cnidocysts types). However, the evolutionary history of nematocysts remains elusive, and despite cnidarians being the earliest animals to evolve a specialized structure for injecting venom, the site of nematogenesis is poorly understood outside of model cnidarians (e.g., *Hydra* (Galliot 2012; Kurz *et al.* 1991)). Indeed, the prevailing hypothesis of nematocyst evolution involves a postulated scenario in which cells secreting soluble toxins acquired extracellular matrix components that likely permitted effective storage and delivery of toxins (Beckmann & Özbek 2012; Rachamim *et al.* 2014). Nematocysts are secreted by nematocytes, thus it has long been suggested that the toxin peptides and proteins making up the venom within the nematocyst capsule are also secreted directly by nematocytes during nematogenesis (Beckmann & Özbek 2012).

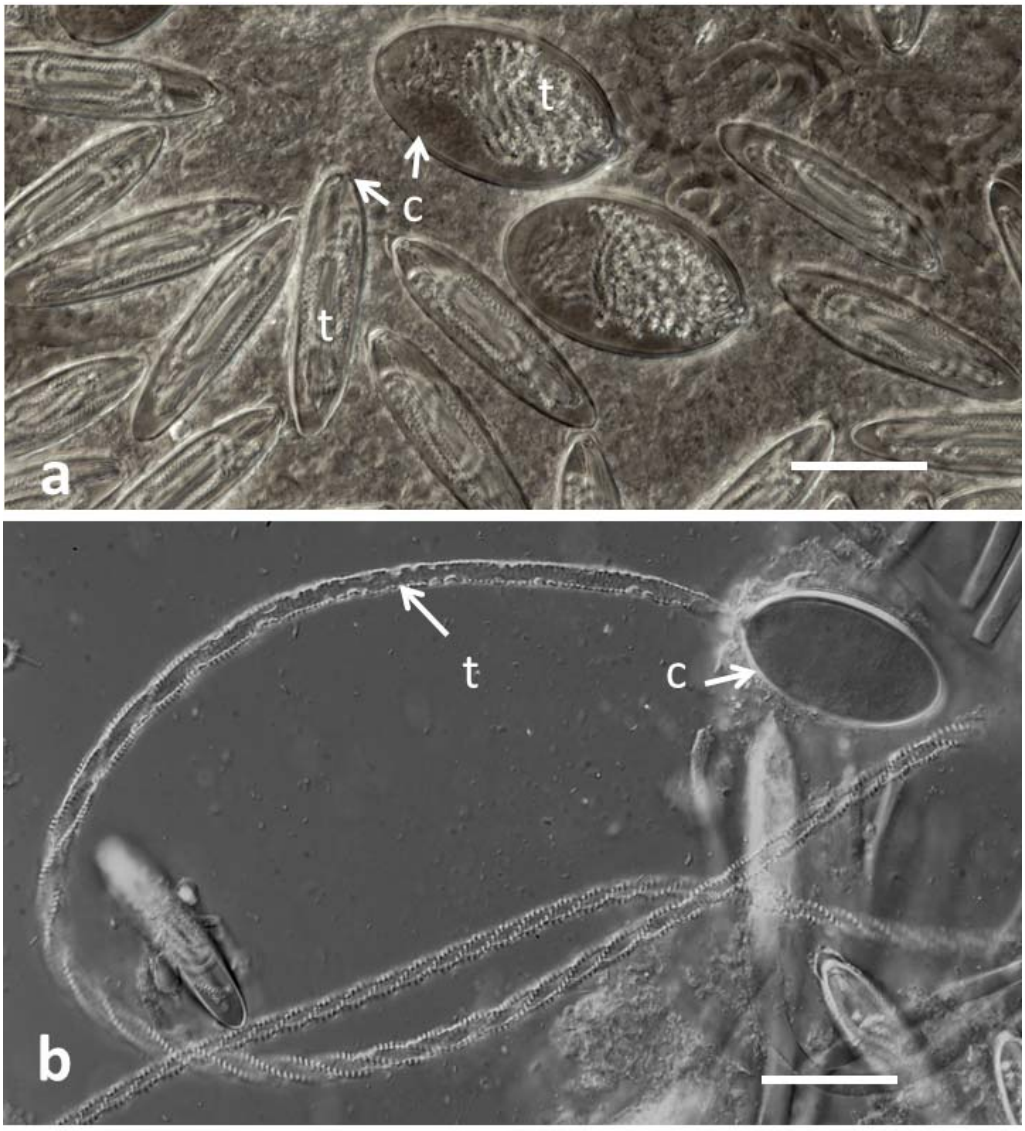

Figure 1.2 Nematocysts isolated from a cubozoan jellyfish (*Morbakka* sp.) tentacle tip. a. Two kinds of nematocysts (penetrant and adherent); b. A single discharged nematocyst. Photos courtesy of Tara Lynn and Allen G. Collins; edited by Cheryl Lewis Ames. Scale bars=50 microns. Abbreviations, c=nematocyst capsule, t=nematocyst tubule.

Bioactive proteins have been isolated from nematocyst venom, and the expression of genes encoding these toxins has been documented in tentacles and body parts where nematocysts are present (Balasubramanian *et al.* 2012; Brinkman *et al.* 2015; Moran *et al.* 2013; Ponce *et al.* 2016). However, the process by which these venom toxins are synthesized within the nematocyte is not known, nor whether cnidarian venom is synthesized exclusive within the nematocysts.

Recently a study on the model cnidarian *Nematostella vectensis* (Moran *et al.* 2012, 2013) localized neurotoxin proteins and their encoding genes in regions outside of the nematocysts, suggesting that certain soluble venom toxins are produced elsewhere in the animal and recruited by nematocytes during nematogenesis. However, it is not known how typical this is of all cnidarian taxa. Studies investigating the expression of venom candidate genes in multiple cnidarian taxa are needed to corroborate these preliminary findings of venom production outside of the nematocyte.

Next Generation Sequencing (NGS) – a useful tool for elucidating the basis of evolutionary novelties

Since the discovery of regeneration in *Hydra* by Trembley almost 275 years ago, much work has been done on this "fruitful" model cnidarian from an organismal biology perspective (e.g., taxonomy, systematics, ecology, neurogenesis, embryogenesis) and from a molecular stand point (e.g., cell fate, innate immunity, transgenics) (Galliot 2012). These prolific studies on *Hydra* have provided the basis of our understanding of the evolution of novel traits within the Cnidaria. However, given that so few studies have ben replicated on free-swimming cnidarians (i.e., medusozoan jellyfish), it is premature to speculate about the potential homology of these novelties across the phylum. Until recently, very little was known about the biological processes, or cellular and molecular components underlying complex behaviors (e.g., venom, vision, sex and development) in non-model medusozoan. Next Generation Sequencing (NGS) techniques and, RNA-Seq transcriptomics in particular, have proven to be an effective method of profiling putative candidate genes in various organisms, including cnidarians.

Transcriptomics has been used successfully in cnidarians to identify genes involved in different stages of a scyphozoan life cycle (Brekhman *et al.* 2015) and in different polyp types in

colonial hydrozoan (Sanders *et al.* 2014); to identify candidate venom genes and characterize bioactive proteins in nematocyst venom (Brinkman & Burnell 2009; Brinkman *et al.* 2012, 2015; Jouiaei *et al.* 2015a; Li *et al.* 2014; Macrander *et al.* 2015, 2016; Ponce *et al.* 2016; Weston *et al.* 2013; Yanagihara & Shohet 2012); to understand the evolution of nematocysts structural genes in a number of cnidarian species (Beckmann & Özbek 2012; Chang *et al.* 2015; David *et al.* 2008; Engel *et al.* 2002; Hwang *et al.* 2010; Özbek *et al.* 2009; Shpirer *et al.* 2014); to characterize molecular components of the opsin-mediated phototransduction pathway in cnidarians with and without lens eyes (Bielecki *et al.* 2014; Feuda *et al.* 2012; Liegertová *et al.* 2015; Mason *et al.* 2012; Porter *et al.* 2012; Schnitzler *et al.* 2012; Shinzato *et al.* 2011; Speiser *et al.* 2014; Suga *et al.* 2008); to localize egg-yok protein precursors in developing egg follicles in corals (Shikina *et al.* 2013); and to reconstruct cnidarian phylogenetic relationships within the Metazoa (Chang *et al.* 2015; Yang *et al.* 2014; Zapata *et al.* 2015).

I propose the use of RNA-Seq transcriptomics in this study as a reasonably unbiased method to profile putative candidate genes of interest broadly implicated in venom, vision and sex in different body parts and life stages of the medusozoan box jellyfish *Alatina alata* (Cubozoa: Alatinidae).

The class Cubozoa

Cubozoans (box jellyfish) are the most venomous animals in the world (Barnes 1966; Bentlage *et al.* 2010; Gershwin *et al.* 2010), and depending on the species, cubozoan stings in humans can range from asymptomatic to potentially life-threatening (Barnes 1964, 1966; Bentlage & Lewis 2012; Bentlage *et al.* 2010; Brinkman *et al.* 2015; Conant 1898; Gershwin *et al.* 2013; Lewis & Bentlage 2009). Furthermore, despite being a small class made up of less than 50 described species, cubozoans have evolved a number of unique and distinguishing features.

Some cubozoans have evolved complex sexual behavior including synchronous spawning aggregations, mating and internal fertilization (Garm *et al.* 2015b; Lewis & Long 2005; Lewis *et al.* 2013). In species with internal fertilization, sperm are taken up by the female, and following fertilization blastulae or planulae are released into the water (Bentlage & Lewis 2012; Marques *et al.* 2015). Swimming planula larva ultimately settles onto a substrate to become an asexually reproducing polyp. Polyps then give rise to medusae (jellyfish), which have separate sexes and are the sexually reproductive stage (Fig 1.3). A particularly note-worthy character of cubozoan medusae is their image-forming lens eyes located on sensory structures known as rhopalia (singular rhopalium), which have been implicated in visually-guided behavior (Garm *et al.* 2007a; b; Nilsson *et al.* 2005; O'Connor *et al.* 2010). These rhopalia are also equipped with additional photoreceptors that lack lens, and are structurally similar to eyes spots on cubozoan planulae (Kozmik *et al.* 2008b).

Few accounts of cubozoan prey capture have been documented in natural settings, but from the limited available reports it appears that box jellyfish feed on a variety of prey items, from microscopic zooplankton (Arneson & Cutress 1976) to larval fish and eels (Júnior & Haddad 2008; Larson 1976). In cubozoans, digestion is presumed to be facilitated by gastric cirri, which line the edge of the quadrangular stomach and contain nematocysts (Larson 1976). However, some cubozoans lacks nematocysts in the gastric cirri (Gershwin 2005; Lewis *et al.* 2013), or completely lack gastric cirri (Bentlage & Lewis 2012), suggesting that extra-cellular digestion occurs with the stomach of cubozoans.

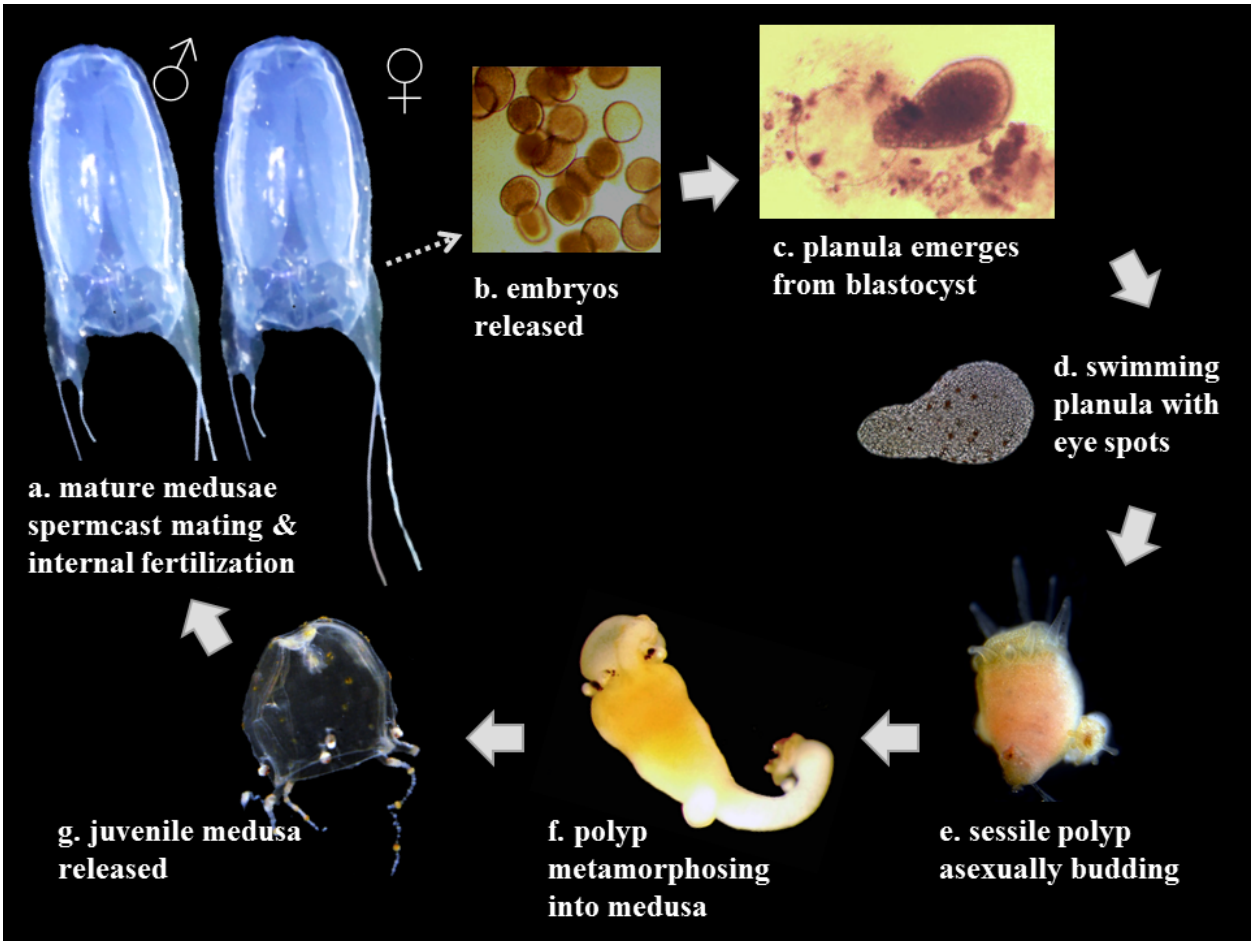

Figure 1.3 The Cubozoan life cycle.
a. Mature medusae mate via spermcasting aggregations and internal fertilization; b. Embryos released into the water; c. Planulae hatch from blastocysts with perisarc; d. Sessile polyps adhere to the substrate and bud additional polyps asexually; e. Polyp tentacles retract and the apical type metamorphosis into a medusa; f. Juvenile medusa is released. This figure is a composite of the life stages of two different cubozoan taxa: *Alatina alata* in Figs a-d, and cf. Alatinidae in Fig e-g. Photos by Cheryl Lewis Ames and Allen G Collins.

The target organism – Alatina alata

The target species of this study is the box jellyfish *Alatina alata* (Reynaud 1830). This species is the second oldest named box jellyfish (*Carybdea marsupialis* Linnaeus 1758 is the oldest), and has more than a century-old reputation as a noxious marine stinger (Fig 3a) (Conant 1898; Mayer 1915). Although it was previously called *Carybdea alata* (family Carybdeidae), it is now properly known as *Alatina alata* (family Alatinidae). *Alatina alata* has both complex and

simple eyes, its potent sting causes serious human envenomation, and mature medusae of this species form monthly nearshore spawning aggregations at predictable times (8–10 days after the full moon) in Atlantic and Indo-Pacific localities (Arneson & Cutress 1976; Carrette *et al.* 2014; Chiaverano *et al.* 2013; Crow 2015; Lawley *et al.* In Press; Lewis *et al.* 2013), making both adult and larval material reliably accessible.

The study species *A. alata* provides a number of advantages for investigating morphological and molecular characters underlying complex processes, such as visually-guided behavior, sexual reproduction, venom synthesis, gametogenesis and fertilization. Furthermore, a recent population genetics study, incorporating morphological and behavioral similarities among geographically disparate *Alatina* populations (*A. mordens* and *A. moseri*), revealed that *A. alata* is a single pantropically distributed species (Lawley *et al.* In Press). While repeated human-mediated introductions of *A. alata* is a possible explanation for this phenomenon, it seems more likely that genetic metapopulation cohesion is maintained in *A. alata* via dispersal through the motile medusa stage or via dispersal of encysted embryo stages. (Fig 1.3).

This study

The overall objective of this study is to position the cubozoan *Alatina alata* as an emerging model cnidarian to elucidate the molecular components underlying the evolution of biological processes mediating complex behaviors in free-living medusozoans (i.e., jellyfish). Above in chapter 1, I present the tool of RNA-Seq transcriptomics as a practical method of profiling putative candidate genes in different body parts and life stages of cnidarians, and highlight the key genes and proteins that have been identified in a select set of cnidarians that are of particular interest for this study on *A. alata.*

In chapter 2, I provide a thorough taxonomic redescription of the second oldest box jellyfish (Cubozoa) species *Alatina alata* and establish a neotype for the species within the Smithsonian National Museum of Natural History Invertebrate Zoology collections. In addition to characterizing the cnidome and ovoviviparous reproduction of *A. alata*, I describe a new cubozoan character, the velarial lappet.

In chapter 3, I generate a functionally annotated transcriptome of adult and larval *Alatina alata* tissue, compare these transcripts to known eukaryote gene and protein databases, and identify candidate genes based on homology. Next, I apply preliminary differential expression analyses to identify tissue-specific gene expression profiles of candidate genes involved in nematogenesis (e.g., minicollagens) and venom production (e.g., CaTX jellyfish toxin family), vision (e.g., J-crystallins) and extraocular sensory perception (e.g., opsins), and sexual reproduction (e.g., Vitellogenins), in order to develop molecular resources for this emerging model. Finally, I propose putative sites of venom production, nematogenesis, opsin-mediated photoreception, and fertilization within the female *A. alata* medusa.

In chapter 4, I provide a review of cubozoan prey capture and defense, and a comparison of digestion in *Alatina alata* and *Carybdea* box jellyfish species. Next, I present evidence for gland cells associated with the gastric cirri with a putative dual role in secreting toxic-like enzymes and toxin proteins. Furthermore, I employ bioinformatics mapping techniques to validate the high quality of the *A. alata* transcriptome assembly (generated in chapter 3), and conduct a rigorous comparative analysis of the gene ontology (GO) of venom implicated homologs expressed in the tentacles and gastric cirri, with a particular focus on identifying zinc metalloproteinases.

Significance

This redescription and neotype specimen voucher for *Alatina alata* serve to stabilize the taxonomy of the species, and position *A. alata* as an emerging model box jellyfish. *Alatina alata* causes debilitating stings in humans along beaches in Hawaii, Australia and Puerto Rico (Bentlage & Lewis 2012; Bentlage *et al.* 2010; Carrette *et al.* 2014; Cutress 1970, 1971, 1972; Gershwin 2005; Gershwin *et al.* 2013; Lewis *et al.* 2013) resulting in serious local and systemic symptoms, including sharp pain, nausea, backache, aching joints, headache, hypertension, difficulty breathing and unconsciousness (Gershwin *et al.* 2013; Tamanaha & Izumi 1996; Yoshimoto & Yanagihara 2002). This study provides the first description of the repertoire of genes (and their gene ontologies) encoding putative bioactive proteins in the venom of *A. alata*. These findings might have future applications for developing effective methods of combating the debilitating effects of envenomation.

Presented here is the first functionally annotated transcriptome from multiple tissues of a cubozoan focusing on both the adult and larvae of *Alatina alata*, which serves as a valuable resource for understanding the molecular underpinnings of cubozoan biological processes and their mediation of complex behaviors. The approach of using multiple body parts and life stages to generate this transcriptome permits the identification of a broad range of candidate genes for the further study of biological processes associated with venom, vision and sex. The study also provides the first expression data for these candidate genes in box jellyfish planulae. This new genomic resource and the candidate gene dataset are valuable for further investigating the evolution of distinctive features of cubozoans, and of cnidarians more broadly. The *A. alata* transcriptome has been made publically available, and significantly adds to the genomic

resources for this emerging cubozoan model. The high quality of the transcriptome serves as an excellent tool for an array of additional downstream transcriptome profiling studies.

In recent years much work has been done on the taxonomy and sexual behavior (Lewis *et al.* 2013), life cycle (Carrette *et al.* 2014), mitochondrial genome (Kayal *et al.* 2012; Smith *et al.* 2012), and geographical distribution (Chiaverano *et al.* 2013; Crow 2015; Lawley *et al.* In Press), yet no studies exist describing the feeding behavior or digestion in this cubozoan species. The findings presented herein represent the first description of the ultrastructure of the gastric phacellae of in *A. alata*. Although understanding of the mechanics of digestion in cubozoans is taxonomically restricted, this analysis provides support for an alternative mechanism of venom production via putative gland cells exhibiting secretory properties in the cubozoan stomach associated with the gastric cirri.

Chapter 2: Redescription of *Alatina alata* (Reynaud 1830) (Cnidaria: Cubozoa) from Bonaire, Dutch Caribbean

Abstract

Here we establish a neotype for *Alatina alata* (Reynaud, 1830) from the Dutch Caribbean island of Bonaire. The species was originally described one hundred and eighty three years ago as *Carybdea alata* in La Centurie Zoologique—a mono-graph published by René Primevère Lesson during the age of worldwide scientific exploration. While monitoring monthly reproductive swarms of *A. alata* medusae in Bonaire, we documented the ecology and sexual reproduction of this cubozoan species. Examination of forty six *A. alata* specimens and additional archived multimedia material in the collections of the National Museum of Natural History, Washington, DC revealed that *A. alata* is found at depths ranging from surface waters to 675 m. Additional studies have reported it at depths of up to 1607 m in the tropical and subtropical Atlantic Ocean. Herein, we resolve the taxonomic confusion long associated with *A. alata* due to a lack of detail in the original description and conflicting statements in the scientific literature. A new cubozoan character, the velarial lappet, is de-scribed for this taxon. The complete description provided here serves to stabilize the taxonomy of the second oldest box jellyfish species, and provide a thorough redescription of the species.

Keywords: Neotype, *Carybdea alata*, Carybdeida, cubomedusae, box jellyfish, Atlantic Ocean, aggregation, sexual re-production, deep-sea, taxonomy

Introduction

The winged box jellyfish *Alatina alata* (Cnidaria: Cubozoa: Carybdeida: Alatinidae) has appeared in the scientific literature (as *Carybdea alata*) many times (see (Gershwin 2005) for list of citations) since its original description (by Reynaud) and illustration (by Prêtre; reproduced

herein as Fig 2.1a) appeared in the extensive monograph by French naturalist René Primevère Lesson in 1830. Reynaud's brief description gave no details about the collection events or the whereabouts of the specimen (or possibly multiple specimens), stating only that this box jellyfish "lives in the Atlantic Ocean". Almost a century later, (Bigelow 1938) redescribed *A. alata* (as *C. alata*) (original drawings reproduced herein as Fig 2.1b–g) from Jamaica and Cuba, effectively clarifying its identity in the Atlantic Ocean and the associated Gulf of Mexico and Caribbean Sea, where the name has since been in common usage (see additional reports by (Arneson 1976) from Puerto Rico, (Morandini 2003) from Brazil, (Graham 1998), as *C. alata* var. *grandis* and (Larson *et al.* 1991) from the Gulf of Mexico. In the last decade, nine nominal species from various disparate localities around the world formerly recognized under the name *C. alata* (see (Bigelow 1918, 1938; Mayer 1910)) were revived within the newly established genus *Alatina* Gershwin 2005. As a result, *Carybdea alata*, under the new combination *Alatina alata* (Gershwin 2005), became the oldest available name within the genus. With this work, we aim to stabilize the species name *A. alata* because it is the oldest species within its genus, the original description lacks detail, and no type material exists. Following a thorough examination of live and preserved material from several Atlantic localities, we provide a detailed redescription of *A. alata* and establish a neotype from a population in Bonaire, Dutch Caribbean. *A. alata* forms monthly aggregations at this locality, making regular collection and in situ observations feasible. We also report on its ovoviviparous mode of sexual reproduction, and bathymetric and geographic distribution in the Atlantic Ocean. This study sets the groundwork for future studies of *Alatina* species, which are notorious for their painful stings (see (Yoshimoto & Yanagihara 2002) for detailed list of symptoms related to the debilitating Irukandji syndrome). Furthermore, *Alatina* is poised to emerge as a key cnidarian model organism, with the mitochondrial genome

recently characterized (Kayal *et al.* 2012; Smith *et al.* 2012), and a nuclear genome assembly currently underway (Genbank accession PRJNA167165 and PRJNA41627).

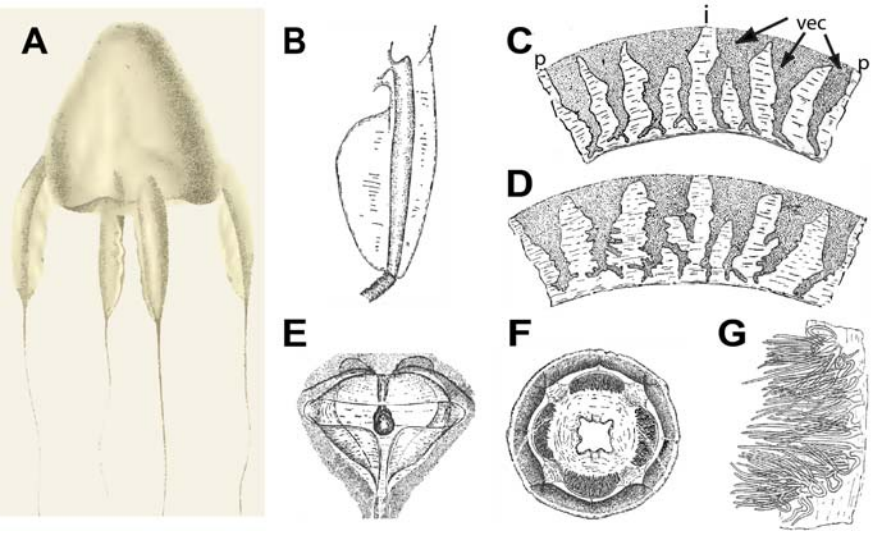

Figure 2.1 Line drawings of *A. alata* reproduced from Reynaud (1830) and Bigelow (1938) (as *Carybdea alata*).
A. Whole body showing bell, tentacles and pedalia (reproduced from Reynaud (1830). B–G, Reproduced from Bigelow (1938). B. Pedalium, C. Velarial canals (BH=70 mm), D. Velarial canals (BH=90 mm), E. Rhopalial niche with T–shaped opening, F. Bell apex showing crescentic gastric phacellae, G. Dissected gastric phacella (comprising gastric cirri). Abbreviations: i=interradius (location of pedalium), p=perradius (location of frenulum), vec=velarial canal (three per quadrant, simple to branching).

Materials and Methods

Sampling location

Bonaire is a small crescent–shaped island (approximately 40 km by 11 km, oriented NW to SE) surrounded by fringing coral reefs. It lies outside of the hurricane belt, but the exposed eastern side (windward) of the island experiences rough water conditions, in contrast to the protected western side (leeward) where only moderate swells occur. On the leeward side, the reef starts in the upper littoral zone and continues to a depth of 12 m where a steep slope of 20–50° leads to a flat at 25–55 m; and a second drop off descends to 250 m. Currents are slight with the

predominate current moving northward on the leeward shore. Water temperatures range from 26–28 °C, salinity from 34–36 ppt. Maximum annual tide range is about 1 m, and the average range is 0.30 m during a lunar cycle (adapted from De Meyer 1997).

Collection and lab culture

In 2008 the authors began documenting monthly swarms of *A. alata* medusae on the leeward side of the island of Bonaire (The Netherlands) during an initiative to describe the box jellyfish taxa of Bonaire (see also (Collins *et al.* 2011)). On June 24–25, 2011 (20:00–23:00), eight and nine days after the full moon respectively, live *A. alata* medusae were collected close to the surface about 25 m from shore off Karel's Pier (Kralendijk, Bonaire) using a hand–held net (medusae were attracted to the lights around the perimeter of the pier). Males and females were put together in a bucket of seawater (~5 individuals per bucket). Subsequently (23:00– 01:00), an additional 100 individuals that were stranded along the shore of Playa Lechi Beach were collected using gallon–sized Ziploc bags. Stranded medusae had truncated tentacles, diminished swimming ability, and disassociated pieces of opaque gonad material circulating in the gastro–vascular cavity. All medusae were transported to the Bonaire Research Station of The Council on International Educational Exchange (CIEE) several hours after collection. Blastulae were exuded from the manubria of the females, and became planulae within several hours. Planulae were reared for several days in glass bowls filled with filtered seawater in a climate– controlled room at approximately 27˚C in the lab at CIEE. Live medusae were placed in an acrylic and silicone aquarium tank filled with filtered seawater, and photographed. Eleven adult medusae (collected off Karel's Pier) were preserved in 8% Formalin for deposit at the National Museum of Natural History (NMNH) in Washington, D.C. To supplement our observations on live and preserved material from Bonaire, we examined a total of forty six museum specimens in the collection at

the NMNH, as well as a video and photo material that we identified as *A. alata*. Using oil immersion light microscopy at 600x and 1000x magnification on live specimens (for methods see (Yanagihara *et al.* 2002)) and preserved specimens (for methods see (Bentlage & Lewis 2012)), the cnidome of *A. alata* was characterized.

Abbreviations

National Museum of Natural History, Smithsonian Institution catalogue number=USNM; BH=bell height, measured from tip of bell to velarial turnover; BW=bell width, the distance measured between two adjacent rhopalia.

Results

Systematics
Phylum Cnidaria Verrill 1865

Subphylum Medusozoa Petersen 1979

Class Cubozoa Werner 1973

Order Carybdeida Gegenbaur 1857

Family Alatinidae Gershwin 2005

Genus *Alatina* Gershwin 2005

Species *Alatina alata* (Reynaud 1830)

Table 2.1, Figs 2.1–2.6

Carybdea (medusa) *alata* Reynaud, 1830 (in Lesson 1830, pl. 33; Fig 1 reproduced herein as Fig 2.1a)

Marsupialis alata Lesson 1837, p. 9, n. 26; Lesson 1843, p. 278

Tamoya alata Agassiz 1862, p. 174

Charybdea alata Haeckel 1880, p. 441; p. 42; 1940a, p. 5

Carybdea alata Mayer 1910, p. 508–510; Bigelow, 1918, p. 400; 1938, pp. 144–151, Text–Figs 11–16; Kramp 1961, p. 304; Arneson 1976, pp. 36, Figs 1,2, Table 1,2, pl. I–V; Arneson & Cutress 1976, pp. 227–236, Table 1, pl. I A–G; Cutress 1971, p. 19, pl. 1; Larson, 1976, pp. 242; Larson *et al.* 1991, p. 313, Table 2; Humann & Deloach 2002; Morandini, 2003, p. 15–17, Fig. 2; Gershwin 2005 pp. 501–523; Calder 2009, pp. 12, 13, Fig. 1; Bentlage 2010, p. 52; Bentlage *et al.* 2010, p. 498; Bentlage and Lewis, 2012, p. 2602

Carybdea alata var. *grandis* Graham 1998, pp 28–30;

Material examined

Neotype: USNM 1195802, 1 ind, female, BW 40 mm, BH 70 mm (live), BW 30 mm, BH 69 mm (preserved), 24 June 2011, Karel's Pier, Kralendijk, Bonaire, The Netherlands, 12 09' 06.37 N 6816' 40.84 W, depth=surface.

Other material. Collected and identified by Lewis *et al.*, depth=surface: USNM 1205450, 1 ind, female, BW 26 mm, BH 78 mm; USNM 1205449, 1 ind, female, BW 38 mm, BH 77 mm; USNM 1205448, 1 ind, male, BW 24 mm, BH 64 mm; USNM 1205447, 1 ind, female, BW 31 mm, BH 83 mm; USNM 1195807, 1 ind, female, BW 27 mm, BH 80 mm; USNM 1195806, 1 ind, male, BW 30 mm, BH 75 mm; USNM 1195805, 1 ind, male, BW 49 mm, BH 77 mm; USNM 1195804, 1 ind, male, BW 29 mm, BH 84 mm; USNM 1195803, 1 ind, female, BW 32 mm, BH 50 mm; USNM 1195801, 1 ind, female, BW 30 mm, BH 83 mm, 25 June 2011, Karel's Pier, Kralendijke, Bonaire, The Netherlands, 12 09' 06.37" N 68 16' 40.84" W. Collected by Ross *et al.* and identified by Bentlage, B.: USNM 1131246, 1 ind, gonads absent, BW 22 mm, BH 47 mm, 28 Aug 2007, depth=98–133 m (bottom depth=468–595 m), Lease Block VK826, Gulf of Mexico, 29 09' 34.99" N 88 01' 19.99" W. USNM 1131245, 1 ind, gonads absent, BW 19

mm, BH 25 mm, 25 Aug 2007, depth=surface (bottom depth=2206–2282 m), Lease Block AT340, Gulf of Mexico, 27 38' 38.00" N 88 20' 59.99" W. Identified by Burnett, J.W.: USNM 94780, 1 ind, female, BW 24 mm, BH 55 mm, Feb 1992, Guantanamo Bay, Cuba, Caribbean Sea. Collected by U S Navy, identified by Larson, R. J.: USNM 58692, 1 ind, gonads absent, BW 17 mm, BH 30 mm, 3 June 1970, depth=55m, Ocean Acre Area, Off Bermuda, North Atlantic, 31 55' 59.99 N 64 25' 00.00 W. USNM 58691, 1 ind, gonads absent, BW 7 mm, BH 17 mm, 7 Sept 1968, depth=327–335 m, Ocean Acre Area, Off Bermuda, North Atlantic, 31 52' 59.99 N 64 25' 00.00 W. USNM 58655, 1 ind, male, BW 20 mm, BH 49 mm, 28 Oct 1967, depth=550–675 m, Ocean Acre Area, Off Bermuda, North Atlantic, 32 34' 59.99 N 63 58' 00.00 W. USNM 58316, 1 ind, gonads absent, BW 8 mm, BH 16 mm, 28 Oct 1967, Depth=55m, Ocean Acre Area, Off Bermuda, North Atlantic, 31 55' 59.99" N 64 25' 00.00 W. USNM 54367, 1 ind, gonads absent, BW 8 mm, BH 22 mm, Apr 27 1969, depth=0–300 m, Open Ocean Area, Off Bermuda, North Atlantic, 31 55' 00.00 N 67 57' 00 W. USNM 54366, 1 ind, gonads absent, BW 16 mm, BH 55 mm, 25 Apr 1968, depth=350 m, Open Ocean Area, Off Bermuda, North Atlantic, 31 55' 59.99 N 63 46' 00.00 W. USNM 53694, Open Ocean Area, 1 ind, gonads absent, BW 5 mm, BH 9 mm, 6 Apr 1967, depth=224–298 m, Off Cape Hatteras, North Atlantic, 35 02' 59.99 N 74 40' 59.99 W. USNM 53659, 4 ind(s) gonads absent, BW 26 mm, BH 51 mm, BW 10 mm, BH 22 mm, BW 10 mm, BH 24 mm, BW 10 mm, BH 22.5 mm, 1 ind sex undetermined, BW 19 mm, BH 47 mm, 28 Sept 1965, depth=surface, Caracas, Caribbean Sea , 10 54' 00 N 67 58' 00.00 W. Identified by Larson: USNM 58211, 1 ind, female, BW 22 mm, BH 75 mm, 3 Apr 1978, Carrie Bow Cay, Lagoon, Dock, Belize, Caribbean Sea. USNM 58210, 1 ind, female, BW 25 mm, BH 52 mm, 25 Mar 1978, Carrie Bow Cay, Lagoon, Dock. USNM 54472, 4 ind(s), female 1, BW 16 mm, BH 58 mm, female 2, BW 18 mm, BH 64 mm, gonads absent, BW 17

mm, BH 32 mm, sex undetermined, BW 26 mm, BH 43 mm, 13 Oct 1974, Mona Island, Caribbean Sea, 18 04' 0.00 N 67 52' 59.99 W. USNM 54398, 5 ind(s), female, BW 8 mm, BH 17 mm, gonads absent, BW 7 mm, BH 19 mm, sex undetermined 1, BW 6 mm, BH 20 mm, sex undetermined 2, BW 8 mm, BH 16 mm, sex undetermined 3 , BW 7 mm, BH 18 mm, 13 Oct 1974, Mona Island, Caribbean Sea, 18 04' 00.00 N 67 52' 59.99 W. Collected by Lea and identified by Larson, R.J.: USNM 56737, 1 ind, gonads absent, BW 10 mm, BH 20 mm, 11 Sept 1977, depth=0–90 m, Open Ocean Area, Off Delaware, North Atlantic, 37 20' 41.99 N 69 10' 23.99 W. USNM 56736, 1 ind, gonads absent, BW 10 mm, BH 17 mm, 17 Sept 1977, depth=0–50 m, Open Ocean Area, Off Delaware, North Atlantic, 37 18' 24.00 N 66 51' 24.00 W. USNM 56735, 1 ind, gonads absent, BW 6 mm, BH 8 mm, 10 Sept 1977, depth=0–150 m, Open Ocean Area, Off Delaware, North Atlantic, 37 49' 59.99 N 67 25' 23.99 W. Collected by Chase and Nicholson, identified by Larson, R.J.: USNM 54385, 2 ind(s), female 1, BW 30 mm, BH 80 mm, female 2, BW 34 mm, BH 100 mm, 4 Apr 1956, Freemans Bay, English Harbor, Antigua Island, Caribbean Sea. Collected and identified by Bigelow, H.B: USNM 42017, 1 ind, female, BW 29 mm, BH 78 mm, 30 Jan 1914, depth=0–100 m, Open Ocean Area, E of Cape Romain, North Atlantic , 32 32' 59.99 N 72 13' 59.99 W. USNM 41921, 1 ind, gonads absent, BW 17 mm, BH 37 mm, collected 21 Mar 1914, depth=surface, N of Little Bahama Bank, Bahamas, North Atlantic, 27 46' 00.00 N 78 46' 00.00 W. USNM 41920, 1 ind, gonads absent, BW 18 mm, BH 40 mm, 18 Mar 1914, depth=surface, N of Havana, Cuba, Caribbean Sea, 22 31' 59.99 N 81 47' 59.99 W. USNM 41919, 1 ind, gonads absent, BW 19 mm, BH 36 mm, 3 Mar 1914, depth=surface, W of Eleuthera Island, Bahamas, North Atlantic, 25 26' 59.99 N 77 16' 00.00 W.

Catalogued multimedia material (no specimen collected). USNM 1195809, filmed by Harbison, R & Widde, E *et al.*, identified by Larson, R (Larson *et al.* 1991), still frame from a video voucher taken from the JSL manned submersible, 11 Nov 1989, depth=540 m (range=457–610 m), 26 04' 00.00 N 77 32' 59.99 W, Off Gordon Cay, 96 NM Off Rock Point. USNM 1005621, photographed by Continental Shelf Associates for BLM/ MMS and Texas A & M University taken from ROV SeaROVER, identified by Lewis, C., photo voucher, Aug 1998, depth=96.5–108.7 m, 29 19' 39 N 87 46' 00.00 W, Mississippi, MMS Lease Block Destin Dome 661, MMS–MAPTEM/M3–4, Gulf of Mexico.

Neotype
Neotype locality. Bonaire, Dutch Caribbean (Atlantic Ocean)

Diagnosis. *Alatina* with tall narrow bell, flared at base, tapering into truncated pyramid at apex; 4 crescentic gastric phacellae at interradial corners of stomach; 3 simple to palmate branching velarial canals per octant, each with a velarial lappet bearing a row of 3 to 4 nematocyst warts; 4 long wing–like (sensu Reynaud 1830) pedalia, each with a pink tentacle.
Cnidome consisting of heterotrichous microbasic p–euryteles and small birhopaloids in tentacles, and large isorhizas in nematocysts warts.

Description (Figs 2.1–2.6, Table 2.1).
Neotype (Fig 2.2a–f)

Mature female specimen (BW 40 mm, BH 70 mm live; BW 30 mm, BH 69 mm preserved), with tall narrow bell flared at base (Fig 2.2a), tapering into truncated pyramid at apex. Each interradial corner bearing a pedalium: 3 of the 4 pedalia long and broad (approximately 15 mm wide) and wing–like, each bearing a pink tentacle (2a, c) about 2 mm in diameter, with bands of nematocysts along the entire length (Fig 2.2c, e). Fourth pedalium bearing a remnant of a tentacle (only several millimeters in height and width), much smaller and

thinner than the other three (possibly damaged while being photographed together with other live specimens in an aquarium). Bell transparent and colorless in life (Fig 2.2a), translucent in fixed specimens (Fig 2.2e); exumbrella speckled with nematocyst warts. Stomach shallow (Fig 2.2b), lacking mesenteries. Manubrium short (2–3 mm long), wide and flat, with 4 mouth lips curled at the tips. Each of the four corners of the stomach housing a crescentic gastric phacella bearing about 20 gastric cirri (Fig 2.2a, b). A pair of leaf–like gonads, flanking each interradial septa and extending into the gastro– vascular cavity, filled with developing oocytes (Fig 2.2a, b). Adjacent gonads (i.e., ovaries) overlapping in the gastrovascular cavity, disassociated into large pieces following rupture due to internal fertilization event prior to preservation in formalin. Clumps of eggs found in various parts of the gastrovascular system (e.g., the velarial and pedalial canals, and the gastric pockets). Velarium wide, suspended by 4 perradial muscular brackets (frenulae) bracing the subumbrellar wall. Each octant bearing 3 simple to palmate branching velarial canals (variable with each octant), with the pair flanking the perradial frenulum simple to bifurcating, those flanking the pedalia at the interadius bifurcating as two main branches each with 3 to 4 distal branches (up to 7 branches in total), with the velarial canal in between bearing three distal branches (Fig 2.2d). Velarial canals in the two octants flanking the diminutive pedalium are not organized in the regular orderly fashion seen in the 6 other octants. Additionally, the velarium is torn near the base of the smaller pedalium, making it difficult to count the number of velarial canal branches there (N.B. the damaged pedalium is not readily apparent in the live photographs of the neotype specimen , and its small size may be exaggerated due to shrinkage following fixation in formalin). A new cubozoan character was discovered that appears to have been overlooked by previous workers. The structure which we call velarial lappets is found in sets of three in each octant overlaying the proximal two thirds of each velarial canal, each bearing a row

of 3 to 4 nematocyst warts (Fig 2.2d). Four club–shaped rhopalia (Fig 2.2a), each situated just above the point where the frenulum connects to the subumbrella bearing 2 median lensed eyes, 2 lateral slit eyes, 2 lateral pit eyes, and 1 statocyst (Fig 2.2f). Rhopaliar niche opening T–shaped (sensu Gershwin 2005), with a single upper scale and 2 lower scales (Fig 2.2f). Nematocyst warts are scattered over the entire exumbrella and bell apex; occurring in rows varying in number along the pedalial keel and the rhopaliar niche scales (Fig 2.2f). Cnidome characterized below.

Other material

Medusae measurements are given as min–mean–max values, and were made on preserved specimens unless otherwise stated.

In addition to the neotype, forty five *A. alata* specimens were examined (BH=8–48–100 mm, BW=5–20–49 mm). Bells tall, narrow truncated pyramids like the neotype (Fig 2.1a; 4g, 6a–c). Four long wing–like pedalia (Fig 2.1b; 3c; 6a–c), each with a single highly contractile tentacle, round in cross–section, 'fluorescent' pink in both live and preserved specimens, with numerous bands of nematocysts along the length (Fig 2.5a). Wide velarium (Fig 2.1c, d; 3a–d), suspended by 4 perradial muscular brackets (frenulae) bracing the subumbrellar wall (Fig 2.3a–d). Twenty four simple to branching velarial canals, 3 per octant, extending from the gastro–vascular space of the bell into the velarium (Fig 2.1c, d; 3a–c & e). Velarial canals flanking perradial frenulae simple, bifurcating distally, or giving rise to side branches; those flanking interradial pedalia splitting into two distinct sub–branches in first half to two– thirds of main branch; subsequent branching increasing in complexity (on average 4 sub–branches per primary branch; maximum 10), making it difficult to trace velarial canal tips back to the original branch, and often giving the appearance of greater than 3 primary velarial canals per octant (Fig 2.1c, d; 3a, b). Bifurcating velarial canals in individuals as small as BH=8 mm; some canals with three

branches in individuals of BH=16 mm. Larger medusae generally displaying more highly branched velarial canals; wide variation exists between each octant in a single individual (1 to 5 branches seen per canal). Velarial lappets are thick gelatinous pouch–like structures comprising the proximal 50–75% of each velarial canal (24 total; 3 per octant), each bearing a row of 3–4 nematocyst warts (Fig 2.3a, b, e), function unknown. Bell transparent and colorless in life (Fig 2.4g, 6a–c), translucent in fixed specimens, speckled with nematocyst warts (nematocyst warts in older material were often lacking, and may have rubbed off in collection gear or after many years in preservation fluids, despite velarial lappet warts being still recognizable in most specimens).

Rhopaliar niche opening T–shaped (sensu Gershwin 2005), with a single upper scale and 2 lower scales enclosing the rhopaliar niche (Fig 2.1e). Four club–shaped rhopalia, each with 2 median lensed eyes, 2 lateral slit eyes, 2 lateral pit eyes, and 1 statocyst (Fig 2.3f, g). All eyes are conspicuous in newly collected material from Bonaire, but in older preserved museum specimens the lens eyes are discolored (having a brown tinge), and bilateral paired pigmented pit eyes and slit eyes are faded to absent, leaving only the complex lens eyes visible. This may be an artifact of prolonged exposure to fixatives.

Stomach shallow, lacking mesenteries; manubrium short (2–3 mm long), wide and flat, 4 mouth lips curled at the tips (Fig 2.4a, f). Four crescentic gastric phacellae, 1 in each corner of the shallow stomach, each with 6–24 basal trunks branching several times at the base, giving rise to up to 100 terminal filaments (gastric cirri) in larger specimens (Fig 2.1f, g; 4a, b, d–g). Smaller individuals (BH<20 mm) possessing long filaments (6–8 basal trunks) extending from each corner into the stomach, giving appearance of a central mass (Fig 2.4c). Eight narrow leaf–like gonads extend in pairs within the gastric pockets, along either side of the interradial septa; filled with developing sperm or eggs in mature males and females (Fig 2.4d, g, h). We examined

preserved individuals with gonads (BH=16–61–100, BW=6–23–34, n=23), as well as preserved medusae without gonads (BH=8–32–88, BW=5–15–32, n=22). All medusae < BH=16 mm lacked gonads, as did many larger medusae (n=18). Gonads presumably are shed in mature medusae following a spawning event (see Arneson 1976). Gonads, translucent in live medusae (Fig 2.4g; 6a, b), turning cream to pale amber in spawning medusae; and opaque in fixed material. Ripe gonads overlap along the perradial plane, becoming increasingly pleated (Fig 2.4g).

Bands of nematocysts found along the entire length of the tentacles primarily contain oval, heterotrichus microbasic p–euryteles (Table 2.1, Fig 2.5a–e) both at tentacle tips, and tentacle base, but more abundant in tentacle tips. Discharged tubules sometimes possessed intact characteristic lancet (Fig 2.5d) contiguous to the shaft (see Yanagihara *et al.* 2002 for description of lancet). Rod–like filaments (Fig 2.5 b, c) are seen attached to euryteles from the tentacles. Small birhopaloids (Table 2.1, Fig 2.5f) found both in tentacle tips and tentacle base, but more abundant in tentacle base. Large, spherical holotrichus, isorhizas (Table 2.1, Fig 2.5g–h) found in exumbrellar (bell) warts and velarial lappet warts. Isorhizas (n=2) were occasionally seen in tissue from tentacle base. Only a single nematocyst was found associated with the gastric cirri: a small birhopaloid was found within tissue of the gastric cirri (L 17.4 μm, W 12.1 μm). No nematocysts were found associated with the manubrium despite thorough examination.

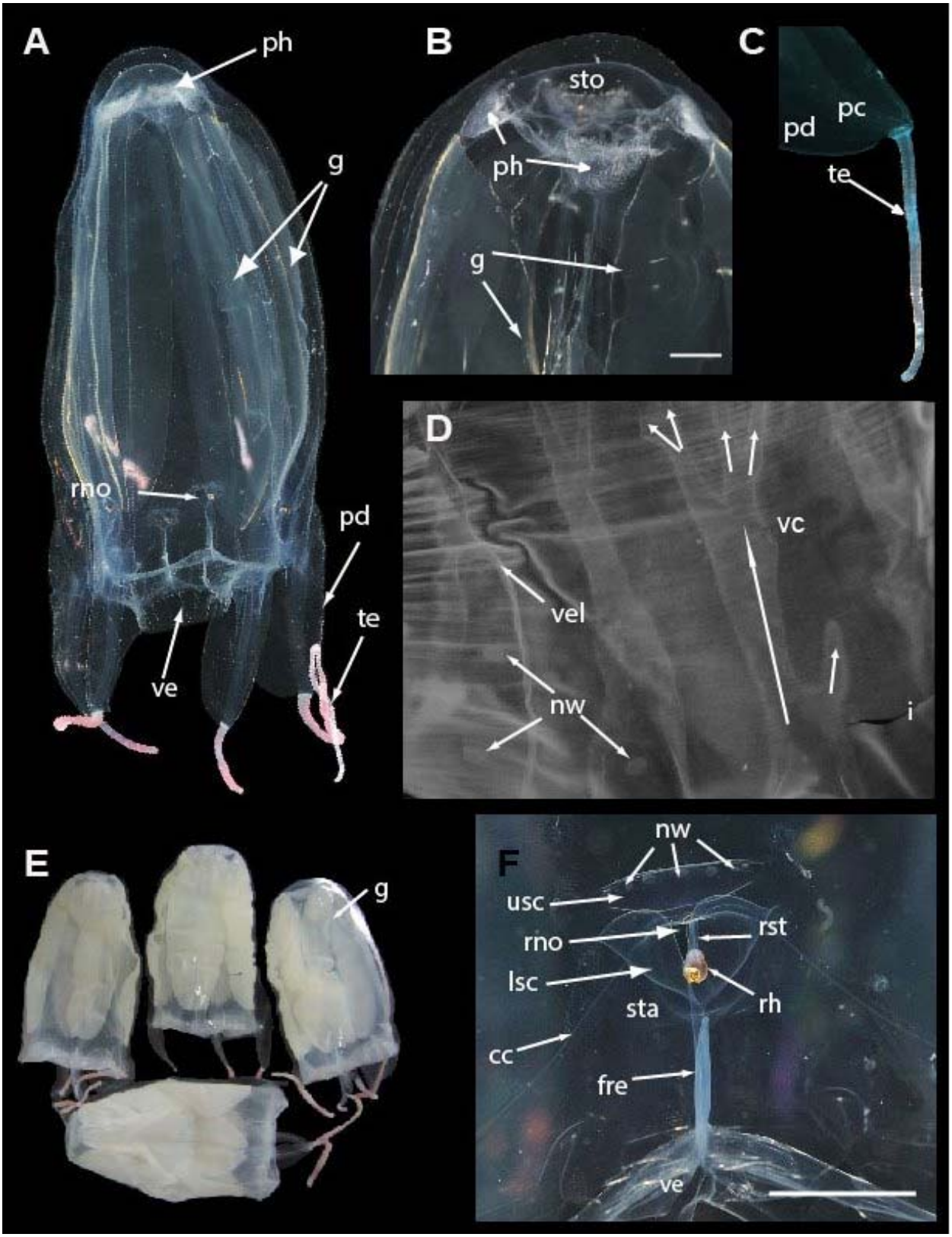

Figure 2.2 *A. alata* neotype (USNM 1195802 unless otherwise specified) from Kralendijk, Bonaire, The Netherlands.

A. Mature female medusa (live), whole body (live, BW=40, BH=70). B. Apical portion showing stomach and crescentic gastric phacellae, visible in each corner, and wide central manubrium opening into subumbrella in live medusa. C. Pedalium and tentacle of live medusa. D. Velarial canals (3 per octant) in preserved medusa: each bears a row of 3 nematocyst warts on proximal first half to two thirds (on the velarial lappets) (dashed arrows indicate the main three velarial canals and branches extending from the main base of each canal). E. USNM 1195803–1195806, preserved mature medusae from the same locality as the neotype. F. T–shaped rhopaliar niche opening with nematocyst warts on upper and lower scale coverings in live medusa. Live photographs by T. Peters (Fish Eye Photography, Bonaire) & A. Yanagihara. Abbreviations: cc=circular canal, fre=frenulum, g=gonads, i=interradius, lsc=lower scale of rhopaliar niche covering, nw=nematocyst wart, ph=gastric phacellae (comprises gastric cirri), pc=pedalial canal, pd=pedalium, rh=rhopalium, rno=rhopaliar niche opening, rst=rhopaliar stalk, sta=statocyst, sto=stomach, te=tentacle, usc=upper scale of rhopaliar niche covering. Scale bars: 5 mm (B & F).

Table 2.1 Cnidome of *Alatina alata*. Nematocyst measurements are provided as min–mean–max; L=length of capsule in µm, W=width of capsule at widest point in µm, SD=standard deviation, n=number of nematocysts measured. Nematocyst identification follows (Collins *et al.* 2011; Mariscal 1974; Östman 2000). See also Fig 2.5a–h.

Body Part	**Nematocyst Type**	**Length (µm)** min–*mean*–max	**SD**	**Width (µm)** min–*mean*–max	**SD**	**n**	**Fig No.**
Tentacle tip	oval, heterotrichus microbasic p–euryteles	19.6–*24.5*–30	3.59	10–*12.58*–20	2.45	49	Fig 2.5a–e
Tentacle base	oval, heterotrichus microbasic p–euryteles	20–*23.7*–30	4.71	10–*11.8*–20	3.49	16	Fig 2.5a–e
Tentacle tip	small birhapaloid	10.5–*13.4*–15.1	1.79	9.3–*12.2*–15.5	2.74	5	Fig 2.5f
Tentacle base	small birhapaloid	10–*11.4*–15.2	2.21	10–*10.4*–12.4	0.8	13	Fig 2.5f
Bell warts	large, spherical holotrichus, isorhizas	28.1–*29.5*–31	0.99	27.8–*29.8*–31.7	1.29	6	Fig 2.5g–h
Velarial lappet warts	large, spherical holotrichus, isorhizas	22.7–*26.9*–29.7	1.8	23.4–*26.8*–28.9	1.8	11	Fig 2.5g–h

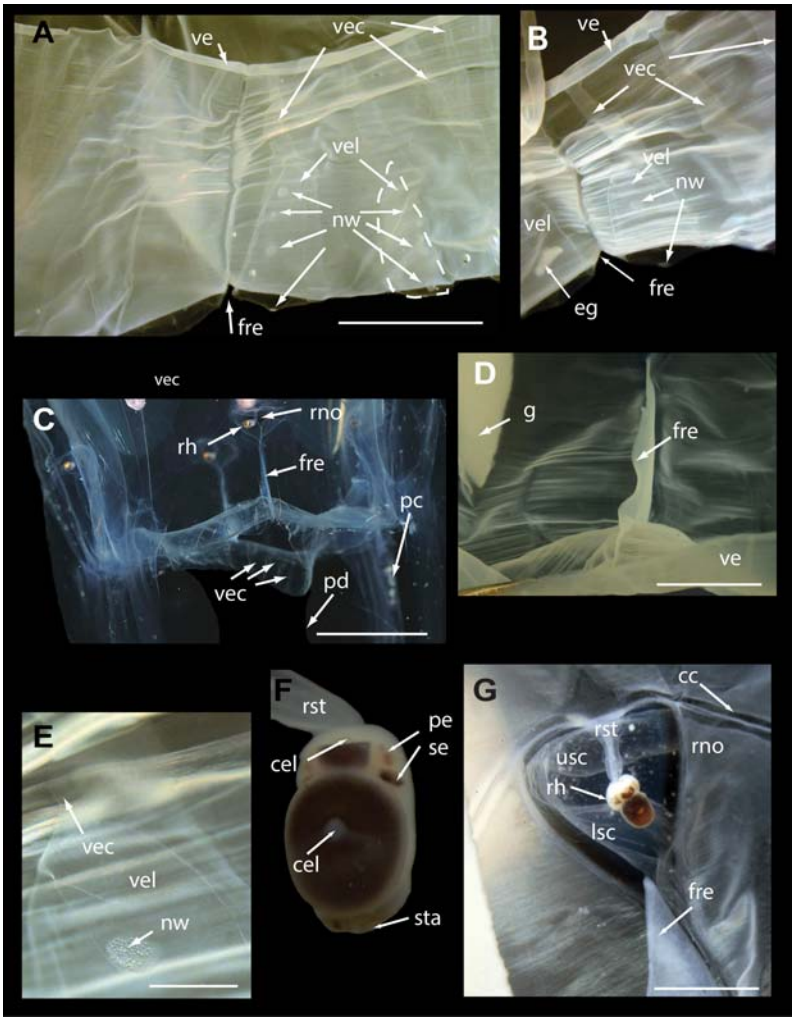

Figure 2.3 *Alatina alata* from Kralendijk, Bonaire, The Netherlands (same location as neotype).

A. & B. USNM 1195807, velarial canals (3 per octant) in preserved medusa, simple to branching (one outlined with dotted line in Fig 2.3A). Each canal bears a row of 3 nematocyst warts on the proximal first half to two thirds, i.e., on the velarial lappets. A clump of disassociated eggs is seen within the velarial canal on the left. C. USNM 1195802 (neotype), velarium of live medusa, partially pushed out revealing velarial canals. D. USNM 1195806, double layered frenulum anchoring the subumbrella to velarium at perradius. E. USNM 1195807, nematocyst wart on velarial lappet filled with isorhiza nematocysts. F. USNM 1195804, rhopalium dissected from preserved medusa, bears two median complex eyes, two upper lateral pit eyes, and two lower lateral slit eyes. G. USNM 1195806, subumbrella view of rhopaliar window of preserved medusa, outline of rhopaliar niche opening visible. Live photographs by T. Peters (Fish Eye Photography, Bonaire) & A. Yanagihara. Abbreviations: cel=complex eye lens, fre=frenulum, g=gonads, nw=nematocyst wart, pc=pedalial canal, pe=pit eye, pd=pedalium, rh=rhopalium, rno=rhopaliar niche opening, rst=rhopaliar stalk, se=slit eye, sta=statocyst, usc=upper scale of rhopaliar niche covering. ve=velarium, vec=velarial canal, vel=velarial lappet. Scale bars: 5 mm (A & D), 10 mm (C), 1 mm (E), 3 mm (G).

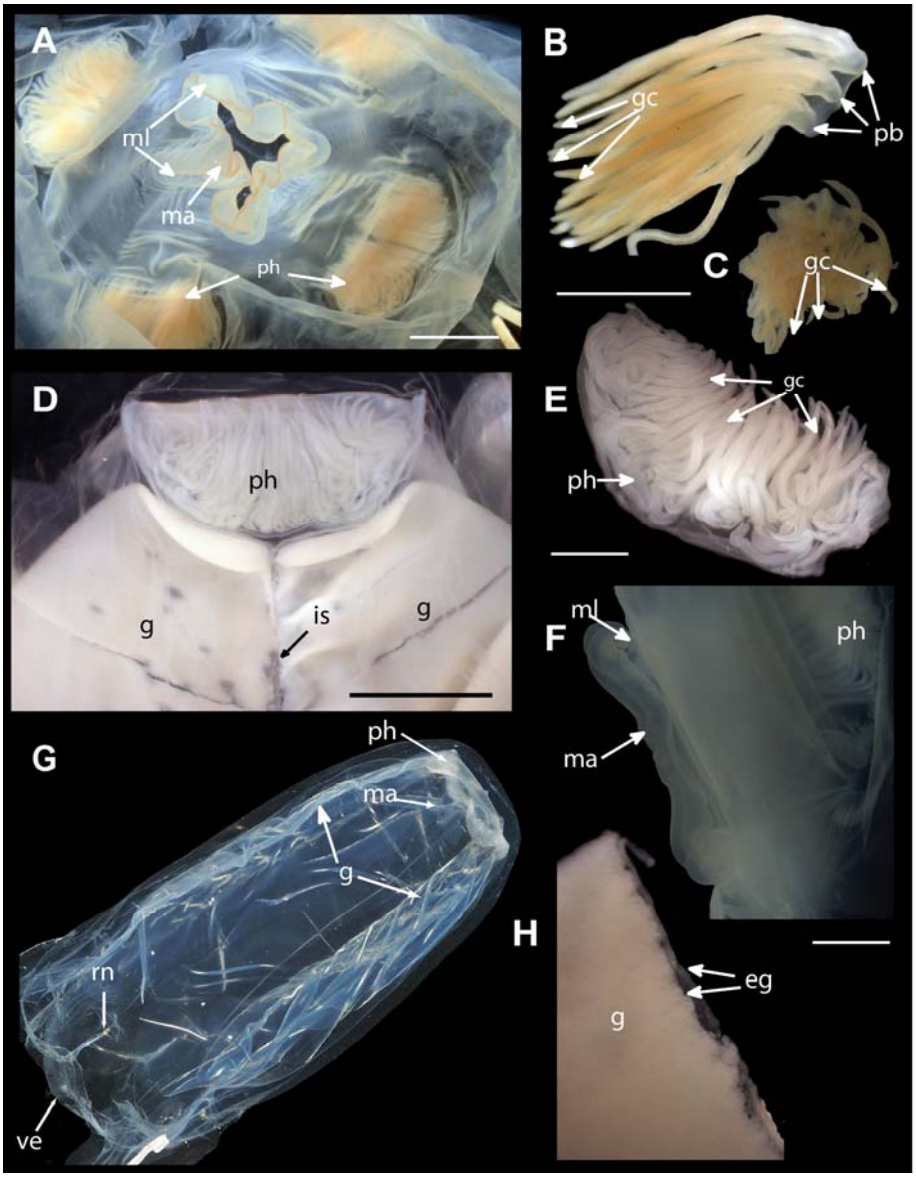

Figure 2.4 *Alatina alata* specimens, preserved and live (all except A–C from Kralendijk, Bonaire (The Netherlands) the same locality as the neotype).
A. USNM 53659, oral view of manubrium and four crescentic gastric phacellae in each corner of the stomach (BW=19 mm, BH=47 mm). B. USNM 53659, dissected tuft of gastric cirri from crescentic gastric phacella. C. USNM 53659, dissected gastric cirri, appearing as central mass in stomach of developing medusa (BW=10 mm, BH=22 mm). D. USNM 1195807, gastric phacellae and adjacent ripe paired gonads in preserved female medusa. E. USNM 1195807, dissected crescentic gastric phacella from mature medusa. F. USNM 1195806, lateral view of manubrium with mouth lips curled up. G. Live mature medusa, gonads beginning to rupture following a reproductive swarming episode (live, BW=60, BH=90). H. USNM 1195807, gonads dissected from mature female, bilayer of mature oocytes visible along adradial point of rupture. Live photographs by T. Peters (Fish Eye Photography Bonaire) and A. Yanagihara. (Fish Eye Photography, Bonaire). Abbreviations: eg=eggs, gc=gastric cirri, ma=manubrium, ml=mouth lips, pb=primary branch, ph=gastric phacellae (comprises gastric cirri), is=interradial septum. Scale bars: 5 mm (A & D), 2mm (B, C, E, F, H).

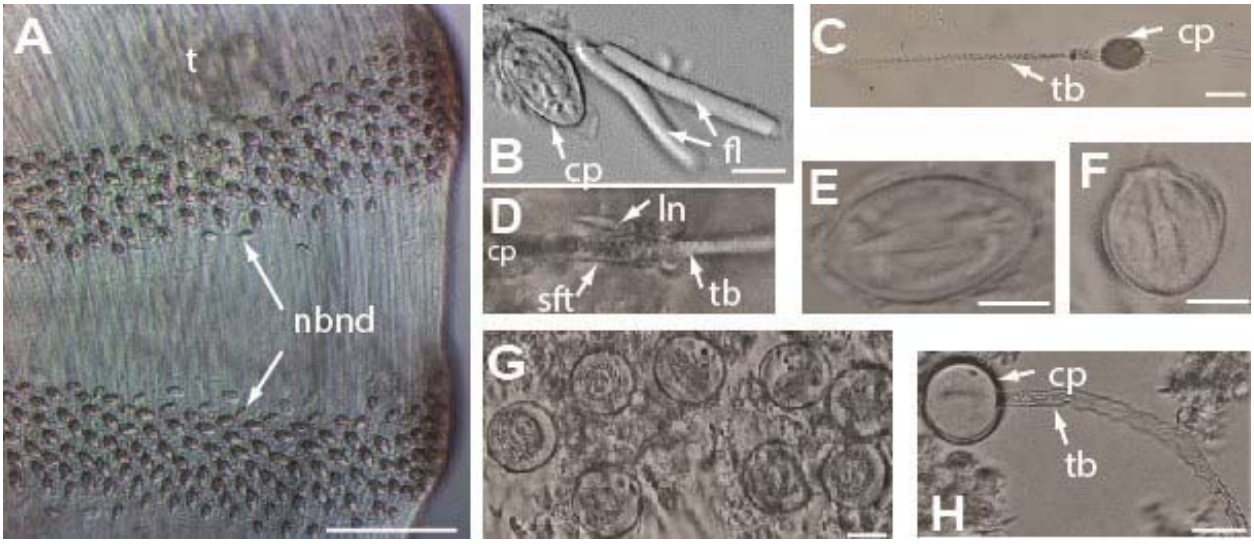

Figure 2.5 Cnidome of *Alatina alata* (all extracted from preserved material unless specified otherwise).

A. Nematocyst bands along the length of the tentacle of a live specimen collected in Bonaire (June 2013). B. & E. Undischarged, microbasic heterotrichus p–euryteles extracted from tentacles; filaments associated with euryteles in B (USNM 1195807). C. Discharged microbasic heterotrichus p–eurytele extracted from tentacles of live medusa. D. Lancet contiguous with discharged tubule of microbasic heterotrichus p–eurytele extracted from tentacles (USNM 1195805) F. Small undischarged birhopaloid found in tentacles of preserved specimen USNM 1195807. G. Large undischarged holotrichus isorhizas found in nematocysts warts (i.e., bell warts & velarial lappet warts (USNM 1195802, neotype). H. Large discharged holotrichus isorhizas found in bell warts (USNM 1195802). H. Abbreviations: cp=capsule, fl=filaments, ln=lancet, nbnd=nematocyst band, sft=shaft, tb=tubule. Scale bars: ~150 μm (A), 10 μm (B, E, F), 25 μm (C), 20 μm (G, H).

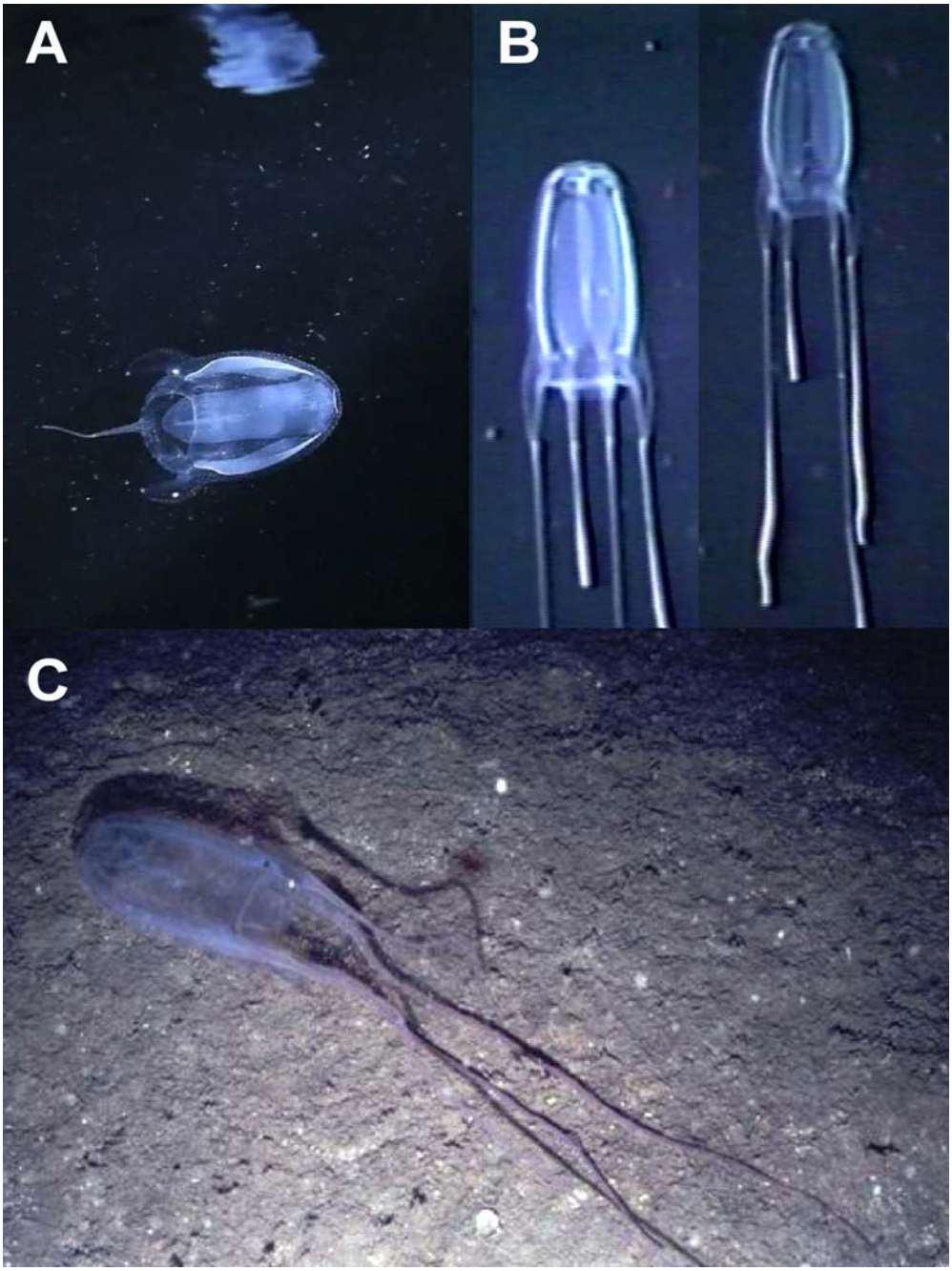

Figure 2.6 *Alatina alata* in situ images.
A. *A. alata* at a depth of 2 m from Kralendijk, Bonaire, The Netherlands. Reflection on the underside of the surface of the water seen above the medusa (frame grab from video by D. Karamehmedovic & A. Yanagihara). B. Series of digital frame grabs taken from video footage of a mature *A. alata* medusa actively ascending in the water column; filmed at a depth of 500–540 m, West off Gorda Cay, Bahamas, from Johnson Sea Link I manned submersible (video voucher USNM 1195809). B. Photograph of. *A alata* medusa oriented parallel to the ocean bottom, photographed by the Remotely Operated Vehicle SeaROVER at a depth of about 100 m in the Gulf of Mexico (USNM 1005621). No specimen was collected as a museum voucher for multimedia material.

Sexual reproduction and early life history

Both male and female gonads are comprised of a bilayer matrix (Fig. 4H). During spawning events witnessed in buckets in the CIEE lab, male gonads became cloudy, and ruptured in several spots along the distal axis releasing spermatozoa into the gastro–vascular cavity, which were then shed into the surrounding water via the manubrium. As females took up sperm into the gastro–vascular system, gonads became opaque, and like in males, they ruptured in several spots, and spherical eggs were ovulated into the gastric sacs, which were by then saturated with sperm (based on microscopic observations of contents). Within several hours embryos were seen circulating throughout the entire gastro–vascular system of the female medusae (in and out of tentacles, velarial canals, etc.). Blastulae were exuded, which developed into planulae that settled as polyps after several days. Later developmental stages were not observed in this study. All but two of the medusae collected during the swarming episode in this study had ruptured gonads with mature ova or spermatozoa spilling into the gastric sacs. In the two exceptions (USNM 1205449 and USNM 1205450), gonads were present as thin strips along the interradial septa, despite the large size of the medusae (BH=77 mm and BH=78 mm). Examination of an excised portion of the preserved gonads from these individuals using oil emersion light microscopy (100x objective (i.e., 1000x magnification) revealed no mature gametes.

Distribution and diet

In this study *A. alata* medusae were mostly observed in shallow waters near shore (Fig 2.6a), but the species is also encountered offshore in deeper waters in the Atlantic Ocean. For instance, seven of the museum specimens examined were collected at discreet depths between 55 m and 675 m. Additionally, we obtained video footage of a mature *A. alata* medusa swimming at a depth of about 540 m West off Gorda Cay, Bahamas (USNM 1195809) (Fig 2.6b), and an in situ photo voucher taken from the ROV SeaROVER (USNM 1005621) depicting *A. alata* close

to bottom at a depth of about 100 m in the Gulf of Mexico (Fig 2.6c). Among the USNM material, four *A. alata* lots were collected in open–net trawls between the surface and depths of up to 300 m. Fourteen were collected on the surface, in shallow and open ocean localities where bottom depth was up to 2282 m. The remaining 22 lots observed in this study had no associated depth record. Live medusae collected from Bonaire had hyperiid amphipods in their subumbrella; some museum specimens examined contained euphausids and small caridean shrimps in the stomach or subumbrella, and one individual had pelagic polychaete worms lodged in three of the six velarial canals in a quadrant, but most medusae collected had empty stomachs.

Discussion

At the time of *Alatina alata*'s original description, *Carybdea marsupialis* (Linnaeus 1758) was the only other cubozoan species known, and only a few morphological characters had been established for delineating the species. It may be argued that it is impossible to know that the specimens described herein as *Alatina alata* refer to the species that Reynaud referred to when he first described it as *Carybdea alata* in Lesson (1830). In the following discussion we recount why we are convinced that specimens from the Dutch Caribbean Island of Bonaire are appropriate for establishing a neotype for *Alatina alata* (Reynaud 1830), and ultimately stabilizing the species name. According to the definition and rules set forth under Article 75 of the International Code for Zoological Nomenclature (ICZN 1999): "A neotype is the name–bearing type of a nominal species–group taxon designated under conditions specified in this Article when no name–bearing type specimen (i.e., holotype, lectotype, syntype or prior neotype) is believed to be extant and an author considers that a name–bearing type is necessary to define the nominal taxon objectively".

The precise type locality of *A. alata* is unknown, and can never be positively determined. Reynaud's original description of this species in Lesson's (1830) La Centurie Zoologique states only that "this medusa inhabits the Atlantic Ocean." The "Centurie", also known under the title "Choix d'animaux rares, nouveaux ou imparfaitement connus" or "Selection of rare, new or imperfectly known animals" was dedicated to Geoffroy-Saint-Hilaire, then professor at what is now the National Museum of Natural History (Paris, France) (Reynaud 1830). In the preface, Lesson explains that the name Centurie was chosen because the intent was to publish 100 plates "drawn from nature" (i.e., still life) from the many collections in the Paris museum (ranging from terrestrial mammals, to birds, and marine invertebrates) (see Lesson 1830). However, due to political events of the time, only 80 plates were compiled to produce the Centurie, which was distributed in 20 issues over a 15–month period from 1830 to 1831 (see Postscript in Lesson 1830). The colored sketches in the plates are almost exclusively attributed to Prêtre (see inscription on front cover of Lesson 1830), but some illustrations were also done by Lesson, as indicated in signatures below each plate.

Many animals depicted in the Centurie were little–known species redescribed by Lesson, or other well–known naturalists (e.g., Cuvier, Peron, Smith, Geoffroy-Saint-Hilaire) (e.g., see Lesson 1830, Plate 2, pp. 14–17, etc.), who emphasized the necessity of an accompanying color plate to enrich each description (see Lesson 1830, Preface). Additionally, a number of new species were described from specimens collected by Lesson and others aboard the vessels La Coquille (1822–1825) and L'Astrolabe (1826–1829) as part of the combined French Scientific World Tour throughout the Pacific and Atlantic oceans (to Madagascar, New Guinea, Australia, Brazil, Mexico; see (Bauchot *et al.* 1990; Reynaud 1830) for additional localities and details). Some of the specimens were collected by Reynaud aboard the corvette La Chevrette to South

Africa and India (1827–1828) and described by him or other specialists (e.g., see Lesson 1830, Plate 1, pp. 11–13). Twelve of the 80 descriptions in the Centurie are of new species described by Reynaud (see Lesson 1830, plates 1, 15, 17, 20, 23, 25, 28, 33, 34, 37 and accompanying descriptions). Five of these species are described from specific localities in India, and one from Cap de Bonne–Espérance in South Africa. The remaining six species, including *A. alata* (as *C. alata*), are reported simply from the Atlantic Ocean suggesting that *A. alata* was observed in an undocumented locality or in several localities throughout the Atlantic Ocean. The entries in the Centurie are not compiled in chronological order or by geographic vicinity, and there is no date associated with the description of *A. alata*. Thus there is no way to determine where Reynaud was when he witnessed and described this new species. There is doubt about whether a specimen was ever collected for *A. alata*, as Reynaud doesn't mention one, unlike Lesson who specifies that many of his descriptions are of museum specimens adorning the galleries, or those found in the collection of the Paris museum and other museums (Lesson 1830). Furthermore, during a recent visit to the collections of the Paris museum we were unable to recover any *A. alata* specimens from the Atlantic Ocean.

Reynaud's new species was quickly adopted by other workers at the time (Agassiz 1862; Lesson 1843). In the absence of Reynaud's material Haeckel (1880) reported on *A. alata* (as *Carybdea alata*) from a specimen he received through Wilhelm Bleek from South Africa. In addition, he placed Reynaud's observation of *A. alata* into the South Atlantic without further explanation, even though Reynaud did not specify where in the Atlantic the medusa he described as *Carybdea alata* was collected. Furthermore, Haeckel (1880) described two additional species of *Alatina* from the tropical belt of the Atlantic, specifically A. pyramis from the Antilles and A. obeliscus from the Cape Verde Islands. Specimens of both species had been deposited in the now

defunct Museum Goddefroy, Germany and their whereabouts are unknown, A. pyramis had not been recognized for about a century until its resurrection by Gershwin (2005), who believes it would be recognizable by its "frilly lips and non– bifurcating velarial canals if ever reencountered. Its description appears distinct from the material we describe herein as *A. alata*, even though Haeckel reports A. pyramis from a region of the Atlantic adjacent to our proposed neotype locality for *A. alata*. A. obeliscus was described from the tropical Western Atlantic (as C. obeliscus), but has been considered unrecognizable (Gershwin 2005), and we concur. Gershwin (2005) suggested that a neotype should be established for *A. alata* when suitable material from the South Atlantic became available, presumably following Haeckel's (1880) reference to the South Atlantic. However, as detailed above, the specific locality of the original material cannot be confidently determined. *A. alata* has appeared in the literature for 183 years, and forty six USNM specimens exist documenting its morphology and distribution in the Western Atlantic (including the Gulf of Mexico and Caribbean Sea). Our investigations of these specimens convince us that the Western Atlantic specimens belong to the same species described by Reynaud in 1830 as *C. alata*.

Conclusions

Our redescription of *A. alata* from the island of Bonaire, Dutch Caribbean is provided with the purpose of stabilizing the identity of *A. alata* whose taxonomy in the absence of type material has been problematic and the cause of much confusion in cubozoan taxonomic studies. Thus by fixing the name to a neotype from Bonaire where fresh material can reliably be obtained, we hope to encourage future studies of the *Alatina* group.

Acknowledgements

I acknowledge my coauthors on this paper. We are thankful to the Bonaire CIEE Research Station staff (Dr. Rita Peachey, Caren Eckrich) for allowing us use of the laboratory during our stay in Bonaire; additional collectors Lauren and Cinde Wirth, Esther van Blerk, Jen, Brevan and Ayla Collins; Andy Ames; Geoff Keel and Katie Ahlfeld of the Smithsonian Museum Support Center (MSC) who helped with specimen curation, and Cheryl Bright (MSC Invertebrate Zoology Collections Manager) for hosting us during our work in the collections; Dr. Mike Vecchione of NOAA NMFS–NSL for notifying us of the in situ video footage; John Reed (Fort Pierce) and Dr. Richard Harbison (WHOI) for providing details of the video log; NMNH Library staff and the Biodiversity Heritage Library who provided us with access to rare and old manuscripts; STINAPA Bonaire Manager Kalli De Meyer and Dr. Frank van Slobbe of DROB; Bonaire Marine Park Manager Ramon de Leon who gave us collection permits and letters of authorization. Finally, we extend a special thanks to Karel of Karel's Pier, Kralendijk, for allowing our group to sample from the pier on which his restaurant is located during the annual Jellyfish Jamboree on Bonaire, The Netherlands. Jonathan Lawley assisted with photographing preserved specimens. We are also thankful to the reviewers of an earlier draft of this manuscript who provided excellent advice for improvement, and to the Zootaxa editors.

Chapter 3: 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

Abstract

Cubozoans (box jellyfish) are cnidarians that have evolved a number of distinguishing features. Many cubozoans have a particularly potent sting, effected by stinging structures called nematocysts; cubozoans have well-developed light sensation, possessing both image-forming lens eyes and light-sensitive eye spots; and some cubozoans have complex mating behaviors, including aggregations, copulation and internal fertilization. The cubozoan *Alatina alata* is emerging as a cnidarian model because it forms predictable monthly nearshore breeding aggregations in tropical to subtropical waters worldwide, making both adult and larval material reliably accessible. To develop resources for *A. alata*, this study generated a functionally annotated transcriptome of adult and larval tissue, applying preliminary differential expression analyses to identify candidate genes involved in nematogenesis and venom production, vision and extraocular sensory perception, and sexual reproduction, which for brevity we refer to as "venom", "vision" and "sex".

We assembled a transcriptome *de novo* from RNA-Seq data pooled from multiple body parts (gastric cirri and tentacle, rhopalium, and ovaries) of an adult female *A. alata* medusa and larval planulae. Our transcriptome comprises ~32K transcripts, after filtering, and provides a basis for analyzing patterns of gene expression in adult and larval box jellyfish tissues. Furthermore, we annotated a large set of candidate genes putatively involved in venom, vision and sex, providing an initial molecular characterization of these complex features in cubozoans. Expression profiles and gene tree reconstruction provided a number of preliminary insights into

the putative sites of nematogenesis and venom production, regions of phototransduction activity and fertilization dynamics in *A. alata*.

Our *Alatina alata* transcriptome significantly adds to the genomic resources for this emerging cubozoan model. This study provides the first annotated transcriptome from multiple tissues of a cubozoan focusing on both the adult and larvae. Our approach of using multiple body parts and life stages to generate this transcriptome effectively identified a broad range of candidate genes for the further study of coordinated processes associated with venom, vision and sex. This new genomic resource and the candidate gene dataset are valuable for further investigating the evolution of distinctive features of cubozoans, and of cnidarians more broadly.

Keywords: Cubozoa, expression patterns, pedalium, sting, embryo, gametogenesis, planulae, eye, spawning aggregations, sperm

Introduction

Background

Cubozoa (box jellyfish) is a class of Cnidaria with a suite of distinct features including a cuboid bell, lens eyes and a typically highly potent sting. Like many cnidarians, cubozoan life history includes a swimming planula larva that ultimately settles onto a substrate to become an asexually reproducing polyp. Polyps then give rise to medusae (jellyfish), which have separate sexes and are the sexually reproductive stage. Some cubozoan taxa have evolved complex sexual behavior including synchronous spawning aggregations, mating and internal fertilization (Garm *et al.* 2015b; Lewis & Long 2005; Lewis *et al.* 2013). Cubozoan medusae vary widely in the potency of their sting; in humans, cubozoan stings range from being harmless to causing deadly

envenomation (Bentlage & Lewis 2012; Brinkman *et al.* 2015; Gershwin *et al.* 2013; Lewis & Bentlage 2009). A particularly note-worthy character of cubozoan medusae is their image-forming lens eyes, which have been implicated in visually-guided behavior (Garm *et al.* 2007a; b; Nilsson *et al.* 2005; O'Connor *et al.* 2010).

Like all other cnidarians, cubozoans possess nematocysts (stinging organelles) essential for prey capture and defense. Nematocysts are remarkably complex subcellular structures that develop within specialized cells called nematocytes. Nematocysts are secreted from post-Golgi vesicles and consist of a double-walled capsule containing venom and a harpoon-like spiny tubule, and one to several different kinds can develop within a cnidarian throughout its life cycle (Mariscal 1974). Nematocysts are of several forms, broadly divided into penetrant (e.g., euryteles) and adherent (e.g., isorhizas). Penetrant nematocysts are primarily concentrated in the tentacles of cubozoan medusae where they are used for prey capture. In some species, nematocysts are also found in body parts with putative digestive roles, such as the gastric cirri (in the stomach), where they may further immobilize prey items inserted into the cubozoan mouth (manubrium) (Larson 1976). Adherent nematocysts are typically found on the exterior of the cubed-shaped bell and do not appear to function in predation (Gershwin 2006; Östman 2000). The location of nematocyst development (nematogenesis) is poorly known in most cnidarians; having only been well-characterized in the model hydrozoan polyp *Hydra*, where morphology and molecular studies reveal clusters of developing nematocysts within the body (David & Challoner 1974). In contrast, molecular studies of another hydrozoan medusa *Clytia,* suggest that nematogenic regions are found in the tentacle bulb, proximal to the tentacles in which mature nematocysts are found (Houliston *et al.* 2010). Transcriptomic and proteomic studies on the cubozoan *Chironex fleckeri*, the scyphozoans *Chrysaora fuscescens* and *Stomolophus meleagris*,

and the hydrozoan *Olindias sambaquiensis* have focused on characterizing venom components from tentacle components (Brinkman *et al.* 2015; Jouiaei *et al.* 2015a; Li *et al.* 2014; Ponce *et al.* 2016; Weston *et al.* 2013), but it is unknown whether nematogenesis and venom production occur in the medusa tentacles. In *A. alata*, tiny unidentified nematocysts occur within the tentacle base which is contiguous with the pedalium, but it is not clear if these represent an early developmental stage of the larger euryteles that are highly concentrated in the tentacle tips (Lewis *et al.* 2013). Studies comparing expression of "venom implicated genes" across medusa body parts can help identify additional putative site(s) of venom production and regions of nematogenesis in cubozoans.

Unique among cnidarians, only cubozoan medusae possess image-forming eyes implicated in visual-guided behavior (Nilsson *et al.* 2005). Two complex eyes, complete with lens and retina, are located on special sensory structures called rhopalia on each of the four sides of the medusa bell. Each rhopalium also possesses a statocyst (balance organ), and two pairs of ocelli (light receptors) (Conant 1898; Nilsson *et al.* 2005) that lack a lens, like other simple animal eyes (having a single pigment cell and at least two photoreceptors (Arendt *et al.* 2009; Land & Nilsson 2002; Nordström *et al.* 2003)). Molecular components of the opsin-mediated phototransduction pathway have been identified in the rhopalium of the cubozoans *Tripedalia cystophora* and *C. brevipedalia* (as *Carybdea rastonii*) (Bielecki *et al.* 2014; Koyanagi *et al.* 2008), as well as in non-rhopalium medusa tissue, and planulae, which have simple eye spots (Bielecki *et al.* 2014; Liegertová *et al.* 2015; Nordström *et al.* 2003). Cubozoan planulae eye spots (ocelli) studied in *T. cystophora* are single cell structures containing a cup of pigment and photosensory microvilli, serving as rhabdomeric photoreceptors (Arendt *et al.* 2009; Nordström *et al.* 2003). Opsins have also been documented in other cnidarians without lens eyes (Feuda *et*

al. 2012; Mason *et al.* 2012; Porter *et al.* 2012; Schnitzler *et al.* 2012; Shinzato *et al.* 2011; Speiser *et al.* 2014; Suga *et al.* 2008) suggesting a role in light perception independent of image formation. Studies comparing expression of "vision implicated genes" across medusa body parts with and without eyes, and planula larvae with simple eyes, can help identify molecular components of the opsin-mediated phototransduction pathway in the rhopalium and aide in discovery of putative areas of extraocular photoreception in cubozoans.

Although most cnidarians reproduce sexually by simple broadcast spawning of their gametes (sperm and/or eggs), many cubozoan species engage in complex sexual behaviors including synchronous spawning aggregations, mating and internal fertilization (Garm *et al.* 2015b; Lewis & Long 2005; Lewis *et al.* 2013). In species with internal fertilization, such as *Alatina alata* (Lewis *et al.* 2013) and *Copula sivickisi* (Lewis & Long 2005), sperm are taken up by the female (as a spermatophore in the latter species), and following fertilization blastulae or planulae are released into the water (Bentlage & Lewis 2012; Marques *et al.* 2015). Histological studies have detected a gametogenic differentiation gradient within the gonads of two cubozoan taxa (*Copula sivickisi* and *Carybdea xamachana*) (Conant 1898; Garm *et al.* 2015a; Lewis & Long 2005), but it is unknown how widespread this process is in cubozoans. Equally elusive is the location of fertilization in cubozoans, although it has been hypothesized to occur in the gastrovascular cavity adjacent to the ovaries in a few species (Lewis & Long 2005; Lewis *et al.* 2013). Comparing expression patterns of "sex implicated genes" in different body parts can help determine whether a gametogenic differentiation gradient is present in additional cubozoan species, and might also aide in pinpointing more precisely the site of fertilization.

The goals of this study were to identify candidate genes in box jellyfish that may be involved in nematogenesis and venom production, vision and extraocular sensory perception,

and sexual reproduction, which for brevity we refer to as "venom", "vision" and "sex" implicated genes. We focused on the species *Alatina alata,* which provides a number of advantages for molecular investigation of these traits. The distribution of nematocysts has been well-documented in this species (Gershwin 2005; Lewis *et al.* 2013; Yanagihara *et al.* 2002) , and its sting is potent, causing serious human envenomation; like other cubozoans it has both simple and compound eyes on the medusa rhopalia as well as eye spots in planulae (ciliated swimming larvae); and mature medusae of this species form monthly nearshore spawning aggregations at predictable times (8–10 days after the full moon) in Indo-Pacific and Atlantic localities (Arneson & Cutress 1976; Lawley et al. In Press; Lewis *et al.* 2013). *A. alata* medusae have also been documented (as *Carybdea* alata) in the open ocean at great depths (Lewis *et al.* 2013; Morandini 2003). The monthly predictability of mature medusae in nearshore waters (Carrette *et al.* 2014; Chiaverano *et al.* 2013; Crow 2015; Lewis *et al.* 2013) and the ability to obtain planulae *in vitro* make *A. alata* a particularly favorable candidate for a cubozoan model.

RNA-Seq transcriptomics provides a reasonably unbiased method of profiling putative candidate genes in different body parts and life stages. This approach has been used successfully in other cnidarians to identify putative genes involved in different stages of a scyphozoan life cycle (Brekhman *et al.* 2015) and in different polyp types in a colony hydrozoan (Sanders *et al.* 2014); to identify candidate venom genes in anthozoans (Macrander *et al.* 2015) and in the tentacles of venomous scyphozoans and cubozoans (Brinkman *et al.* 2015; Ponce *et al.* 2016); and for reconstructing evolutionary relationships within Cnidaria (Zapata *et al.* 2015). In the absence of a reference genome for *A. alata*, we generated a *de novo* transcriptome assembly pooled from RNA-Seq data from specific body parts (gastric cirri, tentacle – including the base of the adjoining pedalium), rhopalium, and ovaries) of a female medusa undergoing internal

fertilization during a spermcast mating event. We compared these transcripts to known eukaryote gene and protein databases, and identified genes implicated in venom, vision and sex based on homology and tissue-specific gene expression profiles. We also investigated the expression of these candidate genes in planulae. Presented here is the first functionally annotated transcriptome of *A. alata*, which serves as a valuable resource for understanding the molecular underpinnings of cubozoan biological processes and their mediation of complex behaviors.

Materials and Methods

Specimen vouchers

Male and female *A. alata* medusae were collected in Bonaire, The Netherlands. All proper collection and export permits were obtained. Medusae were kept within a glass aquarium in filtered seawater for several hours. Males shed sperm into the water that was taken up by the female manubrium. Using light microscopy (1000x) to observe the ovaries, which are located within the gastrovascular cavity, confirmation of ovulation and sperm and egg interaction was made (i.e., putative fertilization) (Fig 3.1c, d). No prey items were present within the stomach and associated gastric cirri or attached to the tentacle. A live female *A. alata* medusa undergoing internal fertilization was placed on ice, and using a sterile RNase-free disposable scalpel tissue samples were quickly excised from the gastric cirri, ovaries, tentacle, rhopalium (Fig 3.1). A fifth sample consisting of thousands of swimming planulae that had developed from blastulae released from different females in the lab was also collected. All samples were placed in 2 ml cryovials and flash frozen with liquid nitrogen. Frozen samples were shipped via Cryoport to the Smithsonian Biorepository. Additionally, a single spawning *A. alata* female medusa was collected at Oil Slick Leap, Kralendijk on April 22, 2014, relaxed in 7.5% Magnesium chloride,

fixed and preserved in 8% formalin, and deposited into the collection of the National Museum of Natural History, Washington, D.C. as a morphological voucher (USNM 1248604). No specific permissions were required from an ethics committee to conduct the research described herein as no humans or protected species were used.

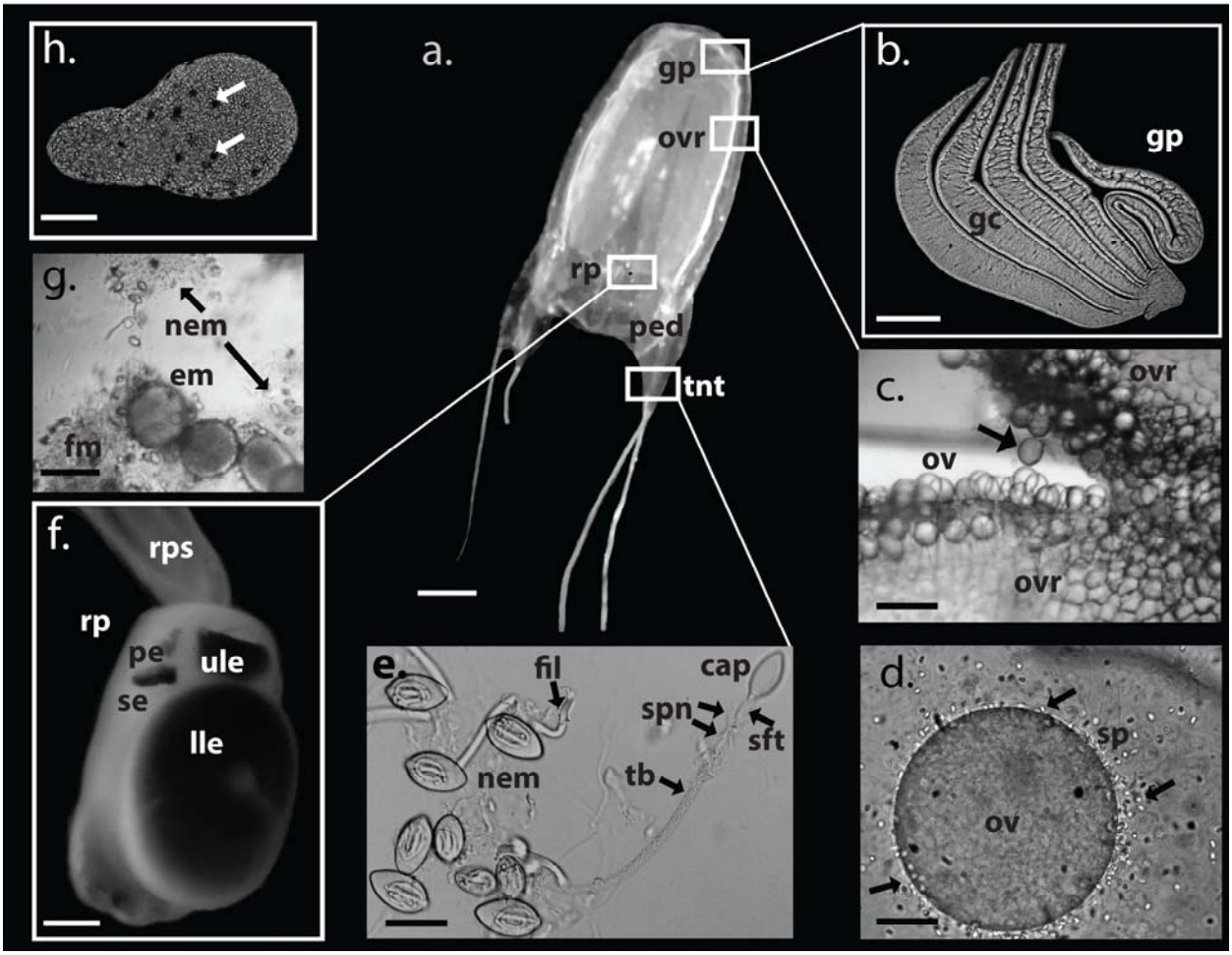

Figure 3.1 a-h. Morphology and life stages of *A. alata* subsampled for *de novo* transcriptome assembly.
White boxes correspond to the location of medusa body parts sub-sampled. a. Mature *A. alata* medusa (live). b. A portion of a gastric phacella removed from a live medusa, with five individual gastric cirri. c. Ovulation documented within the female gastrovascular cavity; arrow indicates imminent release of teardrop shaped ova. d. Interaction between recently ovulated egg and spermatozoa (arrows) in the fluid examined from the gastrovascular cavity (representing putative fertilization). e. Intact nematocysts (euryteles with associated filaments) on the left, and a discharged eurytele on the right, isolated from the tentacle. f. Rhopalium connected to the rhopaliar stalk showing upper and lower lens eyes, and lateral pit and slit eyes (one of each pair visible). g. Bundles of blastulae released by females entangled in fibrous material and intact

eurytele nematocysts. h. Swimming planula (within 20 hours of blastula release from the female); arrows indicating planulae eyes spots. Abbreviations: cap=capsule of nematocyst; em=embryos; fil=filaments; fm=fibrous material, gc=gastric cirri; gp=gastric phacella (comprised of numerous gastric cirri); nem=nematocyst (birhopaloids)s; ov=ovum (ova); ovr=ovaries; pe=pit eye; ped=pedalium; rp=rhopalium; rps=rhopaliar stalk; se=slit eye; sp=sperm; sft=shaft of nematocyst tubule; spn=spines of nematocyst shaft; tb=tubule of discharged nematocyst; tnt=tentacle. Scale bars: a=~15 mm, b=~0.5mm, c=~250 μm, d=~30 μm, e=~20 μm, f=~2 mm, g= ~100 μm, h= ~30 μm.

Sequencing

The five frozen tissue samples (gastric cirri, ovaries, tentacle (and adjoining pedalium base), rhopalium, and planulae) were sequenced at the University of Kansas Medical Center – Genomics Core (KUMC), where total RNA (0.5ug) was used for library preparation for each sample. Illumina HiSeq 2500 Sequencing System was used to generate FASTQ files, which were de-multiplexed into individual sequences for further downstream analysis.

Transcriptome assembly and post-assembly analyses

The 278M paired end raw reads from five samples were analyzed on the Smithsonian Institution High Performance Cluster, SI/HPC, and filtered using the program TrimGalore! (Krueger 2012) with the adaptor trimming tool Cutadapt (Martin, 2011) and FastQC (Andrews 2014) (--quality 30 --phred33 --length 25) to remove Illumina lane and multiplex adaptors (overlapping by 1bp). ALLPATHSLG error correction software (Gnerre *et al.* 2011) was used on the 265M trimmed paired end reads (PAIRED_SEP option was set to 100), and unpaired reads following trimming to predict and correct sequencing errors (see (Macmanes & Eisen 2013)) and mitigate potential errors in transcriptome assemblies. All five samples (i.e., gastric cirri, ovaries, tentacle, rhopalium, and planulae) were pooled and assembled *de novo* into a reference transcriptome (FASTA format) for *A. alata* using Trinity (version trinityrnaseq_r20131110)

(Grabherr *et al.* 2011; Haas *et al.* 2013), with the following additional flags: --no_bowtie --normalize_reads --path_reinforcement_distance 75.

Differential expression estimates and analyses

RNA-Seq by Expectation Maximization (RSEM) was run on each of the five samples separately to estimate transcript abundance (read counts). A single matrix was generated corresponding to expression values for all samples as normalized Trimmed Mean of M-values (TMM) (Robinson & Oshlack 2010). EdgeR was then used to identify differentially expressed genes in the counts matrix (--dispersion 0.1) (Robinson *et al.* 2010); followed by differential expression analysis (i.e., detection of variation in read abundances) to extract all genes most significantly expressed, i.e., with p-values<=0.005 and with at least a fourfold change of differential expression (--matrix iso_r123456.TMM.fpkm.matrix -P 1e-3 -C 2). This EdgeR step generated a single expression matrix of the results of all pairwise comparisons between the five samples. Further, hierarchical clustering generated a heatmap indicating clustering of similarly expressed genes (vertical axis) plotted by sample type (horizontal axis), while maintaining column order by sample (--order_columns_by_samples). This was done for all differentially expressed transcripts for all five samples (Additional 3.IV); just for the medusa samples (Fig 3.3a); and for each of the three subsets of differentially expressed candidate genes (Fig 3.6a, Fig 3.7a and Fig 3.8a). Color-coding on the vertical access of each heatmap indicates gene clusters with similar mean expression levels. Gene cluster patterns were further subdivided into 10 K-mean subclusters, which were visualized as subcluster profile plots (Fig 3.3b-k; Fig 3.6-8b-k). In the absence of biological replicates in this study, the specific significance of fold-change expression levels of each of the differentially expressed genes was of limited value, and we therefore chose to not further filter transcripts based on additional statistical analyses. Instead, all

differentially expressed genes were targeted as candidates for narrowing our search for genes of interest by sample type. Furthermore, redoing the hierarchical clustering analysis on just the three subsets of candidate genes (putative venom, vision and sex genes) allowed us to hone in on gene clusters that were relevant to transcriptome functional annotation and profiling of *A. alata* samples types.

Additional analyses

Venn diagrams were constructed using Venny (Oliveros 2007). Trinotate reports for each of the three sets of candidate genes investigated in this study (venom, vision and sex) were generated by filtering the original *A. alata* Trinotate report using this Python script: https://github.com/pbfrandsen/SI_scripts/blob/master/cheryl_trinotate.py.

Gene tree reconstruction

Amino acid sequences corresponding to predicted ORFs (TransDecoder), or translated nucleotide sequences, from the *A. alata* transcriptome were aligned using MUSCLE (default parameters with 5 iterations) against other cnidarian homologs from NCBI Genbank for the respective candidate genes of interest. ProtTest v. 3.2 was used to determine the most appropriate model of amino acid evolution (i.e., LG+G or WAG+G+F, or BLOSS62 I-G-F) for each alignment. Shimodaira-Hasegawa-like branch support indices (Shimodaira & Hasegawa 1999) and are shown at each node of the ML topology. Alignments and gene trees are available in FigShare.com: https://figshare.com/articles/Supplemental_Information_for_A_new_transcriptome_and_transcriptome_profiling_of_adult_and_larval_tissue_in_the_box_jellyfish_Alatina_alata_an_emerging_modelfor_studying_venom_vision_and_sex/3471425

Availability of data and materials

The datasets supporting the conclusion of this article are available in the following repositories: Raw sequence data have been deposited into the NCBI Sequence Read Archive as BioProject PRJNA312373, as well as corresponding BioSamples each for gastric cirri, ovaries, tentacle, rhopalium, and planulae (SAMN4569893-4569897). The Transcriptome Shotgun Assembly project has been deposited at DDBJ/ENA/GenBank under the accession GEUJ00000000. The filtered *A. alata* transcriptome described in this paper is the first version, GEUJ01000000. Additionally, BioProject PRJNA263637 and BioSample SAMN03418513 correspond to RNA-Seq data used for a cnidarian phylogenomic study (Zapata *et al.* 2015) that we generated from the same *A. alata* medusa and cohort of planulae in the current study. Additional datasets supporting the conclusions of this article are included within the article and its additional files.

Results

Sample collections

Alatina alata material was collected during a spermcasting aggregation in Bonaire, The Netherlands (April 23-25, 2014, 22:00-01:00) according to the methods in (Lewis *et al.* 2013). A single ovulating female medusa (Fig 3.1a) was dissected to provide four tissue samples, namely: i. gastric cirri (Fig 3.1b), ii. ovaries (Fig 3.1c) (within gastrovascular cavity filled with sperm (Fig 3.1d)), iii. tentacle (containing nematocysts Fig 3.1e) and adjoining pedalium base, which we refer to collectively as the "tentacle sample" below (Fig 3.1a), and iv. rhopalium (including rhopaliar stalk) (Fig 3.1f). A fifth sample consisted of planulae (Fig 3.1h) that developed from

blastulae (Fig 3.1g) released by females in the lab. Transcripts putatively involved in venom were characterized through their apparent up-regulation in the tentacle and gastric cirri samples; transcripts putatively involved in vision were targeted through analysis of the rhopalium sample; and transcripts putatively involved in sex and embryogenesis were investigated through analysis of the ovaries and planulae samples. We sought to identify the possible onset of expression of candidate genes in the larval planula stage. Given the microscopic size (~150 μm) of the planulae, multiple individuals were pooled to obtain sufficient tissue for RNA isolation and sequencing. Detailed protocol provided in Methods.

RNA-Seq and bioinformatics

Table 3.1 *A. alata* pooled transcriptome assembly statistics. *De novo assembly* began with 264,505,922 trimmed raw reads generated from medusa gastric cirri, ovaries, tentacle, rhopalium, and planulae samples.

	Whole transcriptome	**Filtered transcriptome (fpkm=1.5)**
No. of transcripts	126,484	31,776
No. of genes	84,124	20,173
Total assembled bases (bp)	125,647,941	48,556,932
Avg (mean) transcript length (bp)	993	1,528
Median transcript length (bp)	456	989
Max transcript length (bp)	9,993	9,993
N50	1,994	2,545
GC content (%)	39	40
Percent proper pairs	77	78
Samtools percent mapped and paired	75	77

De novo transcriptome assembly

RNA-Seq was performed on five different tissues using the Illumina HiSeq2500 Sequencing System (see Methods). Using Trinity software (Grabherr *et al.* 2011; Haas *et al.* 2013) the ~264 million raw paired-end sequence reads were assembled *de novo* into a pooled transcriptome yielding ~126KTrinity transcripts corresponding to ~84K Trinity genes with an N50 of 1994 (Table 3.1). Throughout this paper, we use the term "gene" to refer to each transcriptome component represented by a unique Trinity gene id, and the term "transcript" to refer to additional transcriptome components that Trinity assigned as multiple putative "isoforms" of a single unique gene id (Grabherr *et al.* 2011; Haas *et al.* 2013). Transcriptome completeness was assessed using the subset of 248 widely conserved eukaryote core genes (with low frequency of gene family expansion) using the program CEGMA (Core Eukaryotic Genes Mapping Approach) (Parra *et al.* 2007; Tatusov *et al.* 2003). We retrieved 242 complete CEGs (98%) and an additional three partial CEGs (1%), resulting in 99% CEG representation. We sought to identify highly expressed transcripts which, by definition, are represented by more reads (than lowly expressed transcripts), and thus have a better chance of being contiguously assembled. Therefore we generated a filtered transcriptome comprising a subset of transcripts expressed above a minimum threshold of 1.5 fragments per kilobase per million fragments mapped (fpkm) (Brekhman *et al.* 2015), based on read quantification and alignment accuracy using RSEM (Li & Dewey 2011). This filtering step is consistent with our aim of identifying highly expressed transcripts for potential candidate genes across different samples types. The filtered transcriptome, hereafter referred to simply as the *A. alata* transcriptome, yielded ~32K transcripts corresponding to ~20K genes; N50 = 2545 (Table 3.1). The percentage of sequences above 1000 bp doubled and the percentage of short genes (200–500bp) was reduced by half

(Additional 3.I). Fig 3.2 illustrates the workflow used in this study (modelled after (Sanders *et al.* 2014).

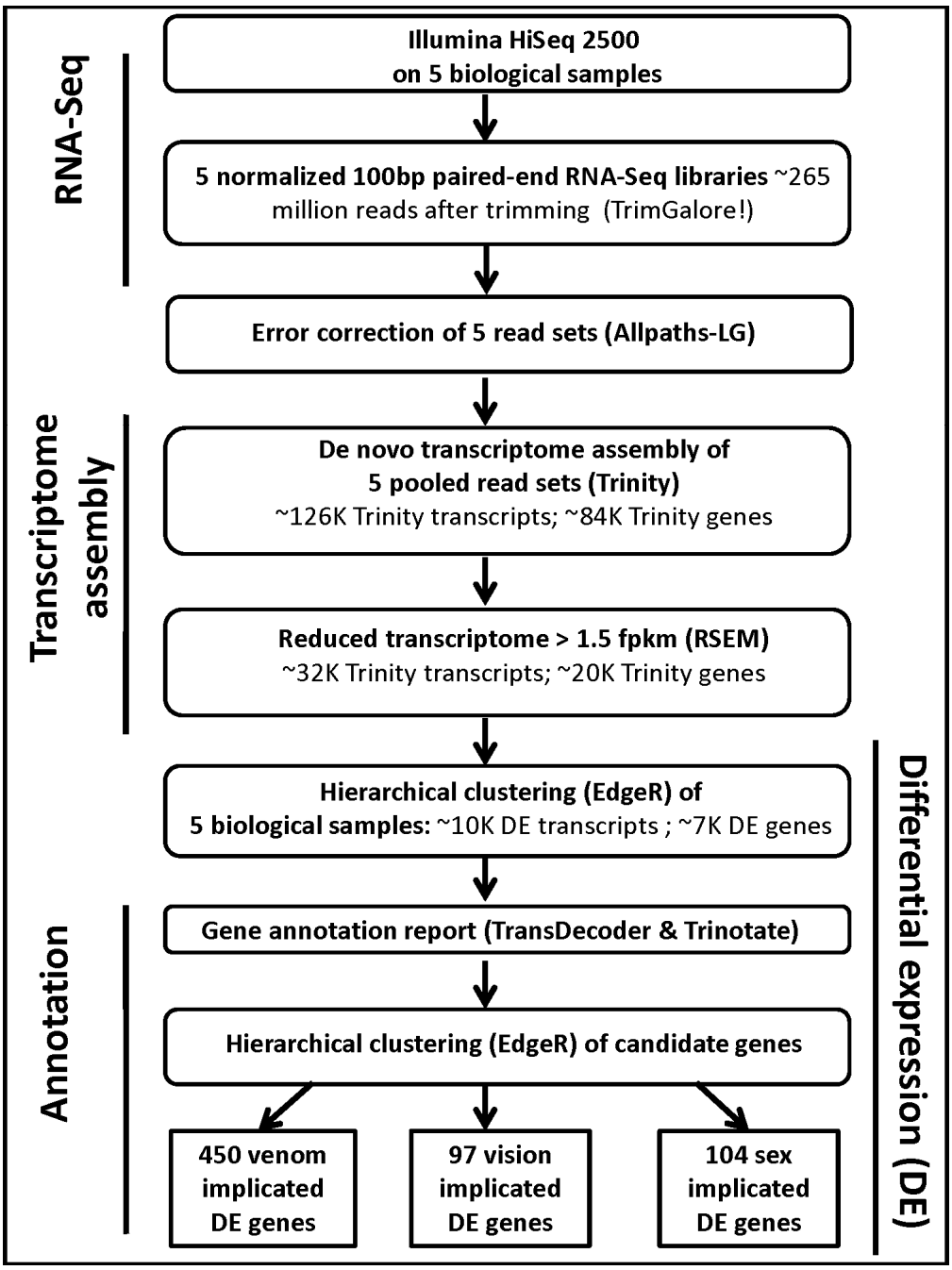

Figure 3.2 RNA-Seq analysis flowchart.
A flowchart outlining the methodology used in transcriptome assembly, gene annotation and differential expression of the five *A. alata* samples analyzed in this study. Additional details provided in Results and Methods. Figure modelled after (Sanders *et al.* 2014).

Functional annotation

Our first objective was to annotate the *A. alata* transcriptome. The longest open reading frames (ORFs) were predicted for transcripts using TransDecoder (Haas *et al.* 2013) and subsequently annotated with Trinotate (Grabherr *et al.* 2011), which compiles results of homology searches of reliable databases (i.e., Uniprot; NCBI; eggNOG/GO; HMMER/PFAM, SingalP) to capture Basic Local Alignment Search Tool (BLAST) protein and gene homologies. The resulting Trinotate report for the ~32K *A. alata* transcripts contained 12,317 BLASTP top hits from TrEMBL and 10,627 BLASTP top hits from SwissProt, from which 656 candidate genes were examined in this study for their putative roles in venom, vision and sex (see Candidate gene profiling below). In total 96 of the top 100 most up-regulated genes in the transcriptome (based on normalized counts) were assigned at least one Trinotate annotation category: 85% of those had BLAST top hits; and 63% corresponded to candidate genes we explored as implicated in either venom, vision or sex in this study (Additional 3.II). The Trinotate report listed 14,551 transcripts corresponding to peptides based on TransDecoder predicted ORFs (open reading frames); 2,098 transcripts with transmembrane protein domains (TMHMM database); 1,610 transcripts containing the classical secretory signal peptide (SignalP database); and 5,252 TrEMBL BLASTP top hits corresponded to cnidarian proteins (Additional 3.III).

Gene expression patterns and profiles

We then sought to detect gene expression patterns across the five samples (gastric cirri, ovaries, tentacle with pedalium base, rhopalium, and planulae) with the aim of providing a descriptive analysis of the top expressed gene clusters by sample. In order to estimate transcript

abundance we aligned each set of reads back to the *A. alata* transcriptome and generated an RNA-Seq fragment counts matrix for each sample using RSEM (Li & Dewey 2011). We subsequently identified differentially expressed genes (see Methods) which we clustered according to their expression profiles using hierarchical clustering analyses within the framework of the EdgeR Bioconductor software package (Robinson *et al.* 2010), a preferred methodology for studies lacking biological replicates (Brekhman *et al.* 2015; Chen *et al.* 2016). Of the ~32K Trinity transcripts (~21K Trinity genes) identified in the *A. alata* transcriptome ~10K transcripts (6,676 genes) were found to be differentially expressed, within a broad range, across the five samples (Additional 3.IV). EdgeR takes the normalized gene counts for all samples (generated from the initial RSEM counts matrix), and then clusters genes with similar mean expression rates across samples (Chen *et al.* 2016). Gene clusters are visualized in the form of a heatmap, permitting pinpointing of genes up-regulated in certain samples that might be of interest as candidate genes. The results of hierarchical clustering were consistent with our initial RSEM evaluation of up-regulated genes by sample.

In an attempt to identify genes that were specifically abundant in each medusa sample, we conducted subsequent hierarchical clustering with the planulae sample excluded. EdgeR analyses of genes from just the four medusa samples (gastric cirri, ovaries, tentacle and rhopalium) revealed 2,916 differentially expressed genes. To identify patterns of highly expressed gene clusters by medusa sample, we generated a heatmap of the subset of 2,916 genes and further partitioned the subset into 10 gene subcluster profiles based on mean expression patterns across samples (Fig 3.3 a-k; Additional 3.V). Furthermore, redoing the hierarchical clustering and subcluster profiling analyses for all five samples using the three subsets of candidate genes (putative venom, vision and sex genes) allowed us to hone in on gene clusters

that were relevant to transcriptome functional annotation and profiling of different *A. alata* samples.

Tissue-specific "core genes"

Our next objective was to identify specific genes that are highly expressed in particular body parts (up-regulated compared to other samples), and subsequently identify functional categories informed by the Trinotate report. We constructed a Venn diagram of the 2,916 differentially expressed genes from the four medusae samples (Fig 3.4a), which revealed that 76% (2,228 genes) were expressed to some degree in all medusa samples (Additional 3.I). A subset corresponding to the top 50 most highly expressed genes per sample was plotted in a second Venn diagram (Fig 3.4b), and genes unique to each sample's top 50 were respectively designated that sample's "core genes" (Additional 3.VI). Only one gene (lacking a Trinotate annotation) among each sample's top 50 most highly expressed genes overlapped in all four samples.

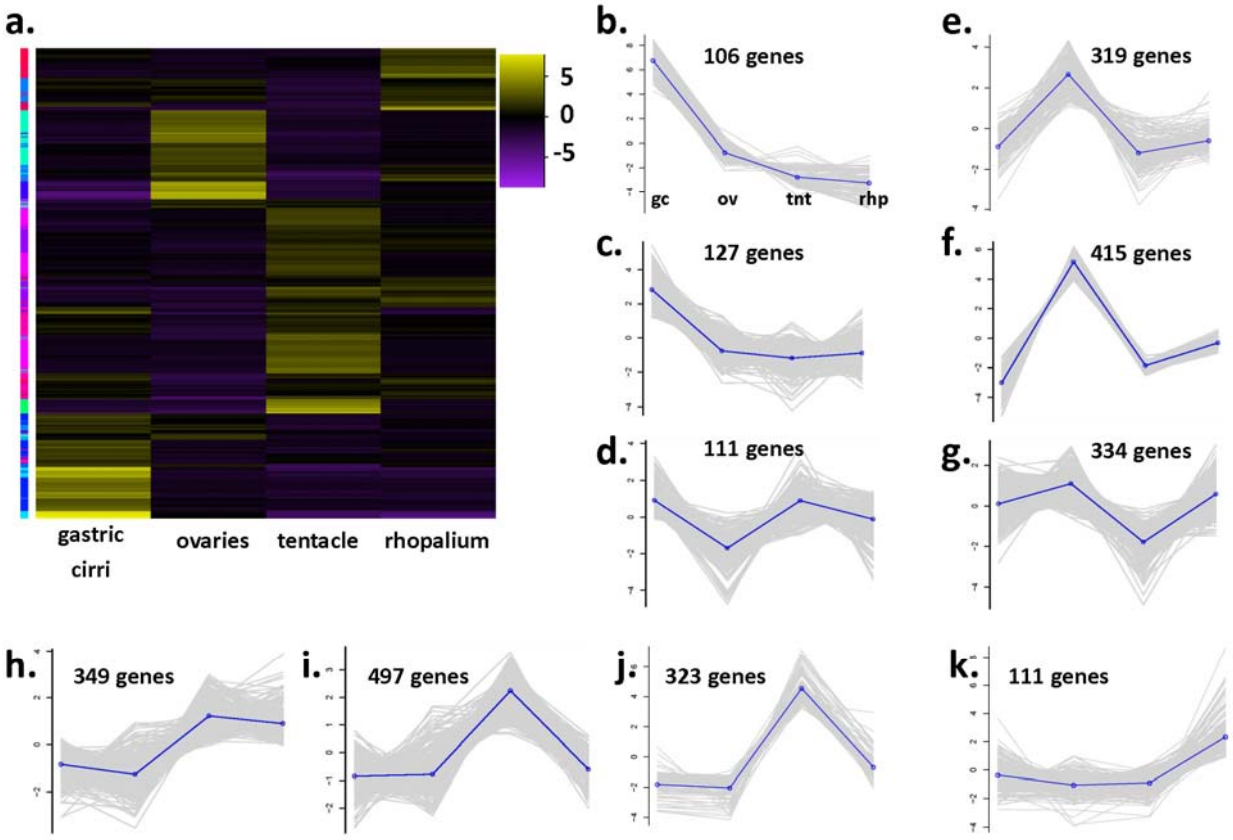

Figure 3.3 a-k. Heatmap of differentially expressed genes in *A. alata* medusa samples. Hierarchical clustering (EdgeR) and corresponding ten subcluster profiles for the 2,916 genes differentially expressed across *A. alata* medusa samples (gastric cirri, ovaries, tentacle, rhopalium) (of the identified ~20K Trinity genes . Intensity of color indicates expression levels for each of the ten hierarchical clusters (vertical access). Bright yellow patches (highly up-regulated) correspond to the highest peaks for each k-mean subcluster profile. K-mean profiles (b-k) match the order of column names in a., representing the mean expression of gene clusters highly up-regulated in each sample (centroid demarcated by the solid line). Three bright yellow transcript clusters in the gastric cirri column correspond to each of the peaks seen in plots b, c and d, in the ovaries column correspond to plots e, f, g, and in the tentacle column correspond to h, i, j, while the two bright yellow clusters in the rhopalium column correspond to peaks in plots k, and to the less prominent peaks in plots g and h. The vertical colored bar on the left of the heatmap (a.) indicates distinct patterns corresponding to the ten subcluster profiles (sc=subcluster number), for which the number of genes each comprises is indicated. Abbreviations: gc=gastric cirri, ov=ovaries, tnt=tentacle (and pedalium base), rhp=rhopalium, and pln=planulae. (Original matrix in Additional file 3.V).

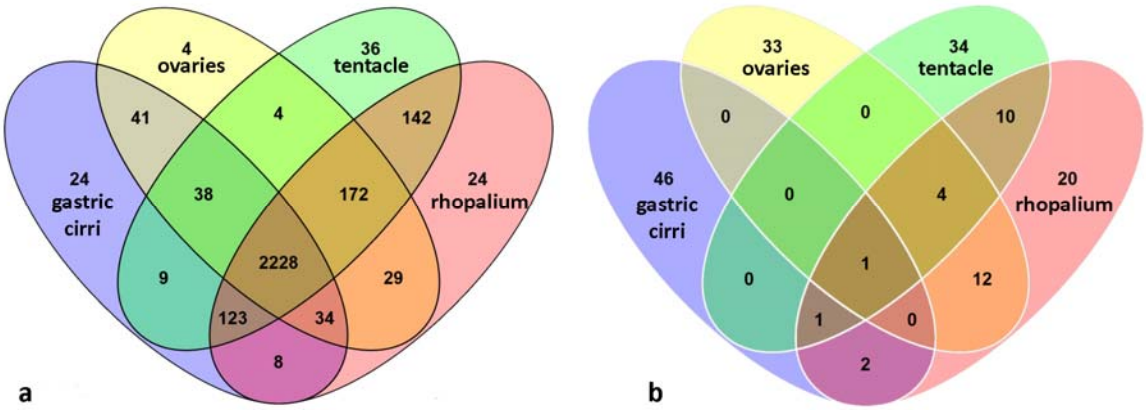

Figure 3.4 a, b. Venn diagrams showing overlap of genes differentially expressed exclusively in *A. alata* medusa samples (gastric cirri, ovaries, tentacle, rhopalium). a. Shows that of the 2,916 total genes differentially expressed across the four samples 2,228 (76%) are expressed in all samples, 24 are unique to gastric cirri, 4 are unique to ovaries, 38 are unique to the tentacle, and 24 are unique to the rhopalium. b. Shows that of the top 50 most highly differentially expressed genes by sample type, a single gene is expressed in all four samples. The subset of genes unique to the top 50 most up-regulated genes by sample, called the "core genes" herein, comprises 46 genes in the gastric cirri, 33 in the ovaries, 34 in the tentacle, and 20 in the rhopalium. Core gene annotations provided in Additional file 3.VI and summarized in histograms Fig 3.6a-d.

An assessment of the core genes revealed that only about 40% had Trinotate-generated annotations (Fig 3.5a-d). Gastric cirri core genes (n=46) corresponded mostly to putative proteins (26 annotations) implicated in venom and digestion, including metalloproteinases (Fig 3.5a) (Brinkman *et al.* 2015; Casewell *et al.* 2012; Jouiaei *et al.* 2015b); ovaries core genes (n=33) corresponded to putative proteins (22 annotations) involved in gametogenesis, including Vitellogenins (Fig 3.5b), which, although well studied in bilaterians (Aungsuchawan *et al.* 2011), are newly reported in cubozoans herein; tentacle core genes (n=34) corresponded to many putative proteins (19 annotations) associated with nematocyst development, including minicollagens (Fig 3.5c) (Balasubramanian *et al.* 2012; David *et al.* 2008; Kurz *et al.* 1991); and rhopalium core genes (n=20) corresponded to several putative proteins (12 annotations) identified in vision and the phototransduction pathway, including J-crystallins (Fig 3.5d) (Kozmik *et al.* 2008b; Liegertová *et al.* 2015; Speiser *et al.* 2014).

Figure 3.5 a-d. Abundances of annotated "core genes" in the *A. alata* transcriptome according to medusa sample.

Column headings correspond to rank(s) among the top 50 of each core gene (or gene family) by sample according to the Venn diagram in Fig 3.4b, protein annotation from UniProtKB SwissProt (Sprot) and TrEMBL (separated by a back slash) and fpkm values in a. gastric cirri, b. ovaries, c. tentacle, d. rhopalium. Genes with putative functions in sperm motility are indicated with asterisks in 5b. Genes lacking Trinotate annotations are not included. Detailed statistics (fpkm, counts, DE values for top 50 ranked genes by sample with annotations) provided in Additional file 3.VI).

Candidate gene profiling

Using the tissue-specific core gene annotations (see above) as a starting point for identifying candidate genes in *A. alata* medusa body parts we compiled three candidate-gene lists comprising 148, 109 and 39 terms (related to genes and/or gene functions) broadly associated with "venom", "vision", and "sex", informed primarily by relevant genes and proteins documented in the scientific literature (Additional 3.VII – IX: provides lists of terms, references and EdgeR expression matrices for all putative candidate genes). Subsequently, we queried the *A. alata* Trinotate report separately using each of the lists to identify matching terms among BLASTX , BLASTP and PFAM (protein family) top hits from the Trinotate transcriptome annotation report. This generated three additional targeted Trinotate annotation reports consisting of gene subsets we categorized as "venom implicated genes" (Additional 3.VII), "vision implicated genes" (Additional 3.VIII), and "sex implicated genes" (Additional 3.IX). Subsequently, EdgeR hierarchical clustering (see above) was conducted separately on each subset of corresponding candidate genes, generating three new heat maps (Fig 3.6, Fig 3.7 and Fig 3.8); profiling the respective patterns of expression of each set of candidate genes across all five samples. In all cases, gene cluster patterns were divided into 10 subclusters based on mean expression patterns for genes with potential implication in venom (n=450) (Fig 3.6b-k); vision (n=97) (Fig 3.7b-k); and sex (n= 104) (Fig 3.8b-k). By comparing planulae samples with samples from body parts of the mature female medusa, an additional aim of this study was to potentially

detect the onset of expression of these possible candidate genes within planulae. EdgeR expression matrices for all candidate genes and their respective annotations are provided in Additional 3.VII – IX. We did not detect expression of most candidate genes in the planulae sample. Instead, we found that transcripts most highly expressed in the planulae comprised mainly: histones (core and early embryonic), *nanos* transcription factor, and genes implicated in neurogenesis, mitosis, microtubule, and protein processing (Additional 3.X).

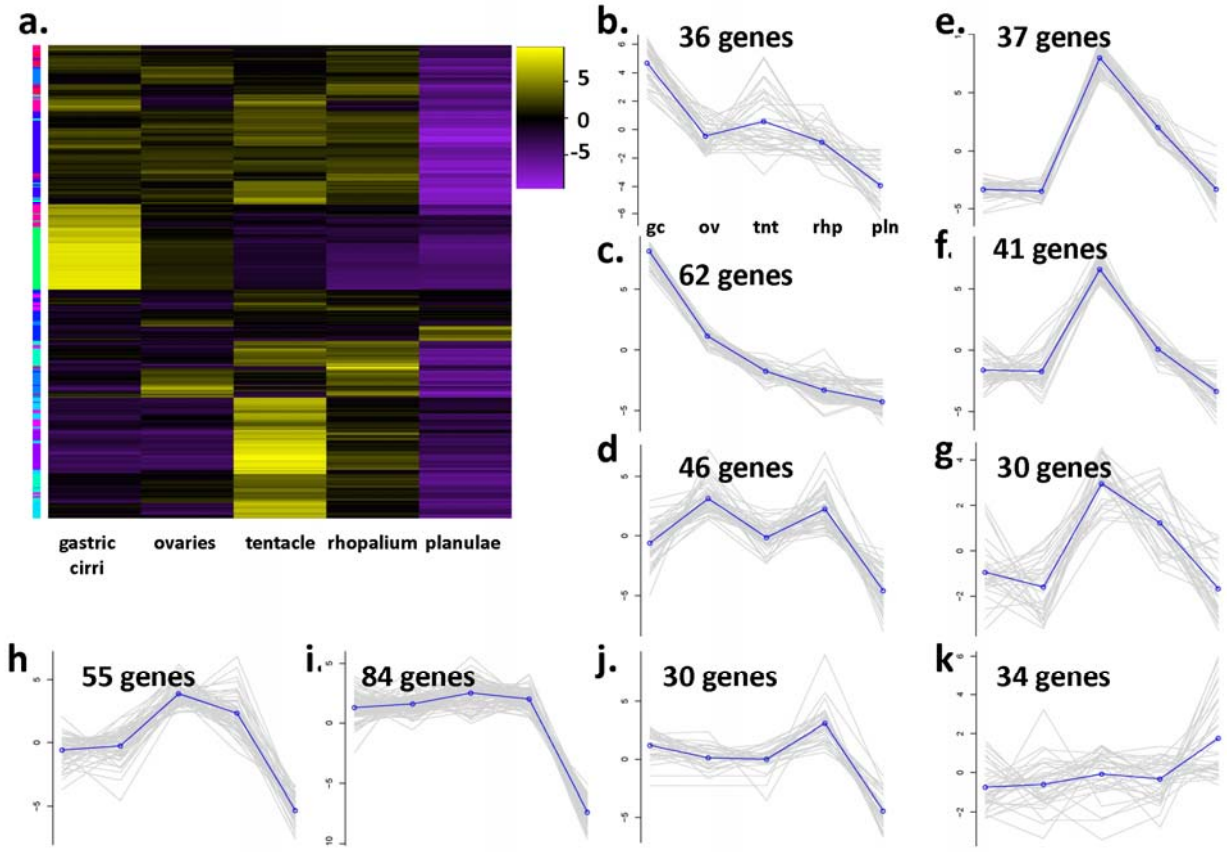

Figure 3.6 a-k. Venom Heatmap for *A. alata*.
Hierarchical clustering (EdgeR) and corresponding ten subcluster profiles for the 455 genes implicated in venom differentially expressed across *A. alata* medusa (gastric cirri, ovaries, tentacle, rhopalium) and planulae samples. Intensity of color indicates expression levels for each of the ten hierarchical clusters (vertical access). Bright yellow patches (highly up-regulated) correspond to the highest peaks for each k-mean subcluster profile. K-mean profiles (b-k) match the order of column names in a., representing the mean expression of gene clusters highly up-regulated in each sample (centroid demarcated by the solid line). Two bright yellow transcript clusters in the gastric cirri column correspond to peaks in plots b and c; one cluster in the ovaries column corresponds to plot d; four clusters in the tentacle column correspond to plots e, f, g and h; one cluster in the rhopalium corresponds to plot j, and to the less prominent peak seen in plot I; one cluster in the planulae column corresponds to plot k. The vertical colored bar on the left of the heatmap (a.) indicates distinct patterns corresponding to the ten subcluster profiles (sc=subcluster number), for which the number of genes each comprises is indicated. Abbreviations: gc=gastric cirri, ov=ovaries, tnt=tentacle (and pedalium base), rhp=rhopalium, and pln=planulae. Gene annotations by subcluster provided in Additional file 3.VII.

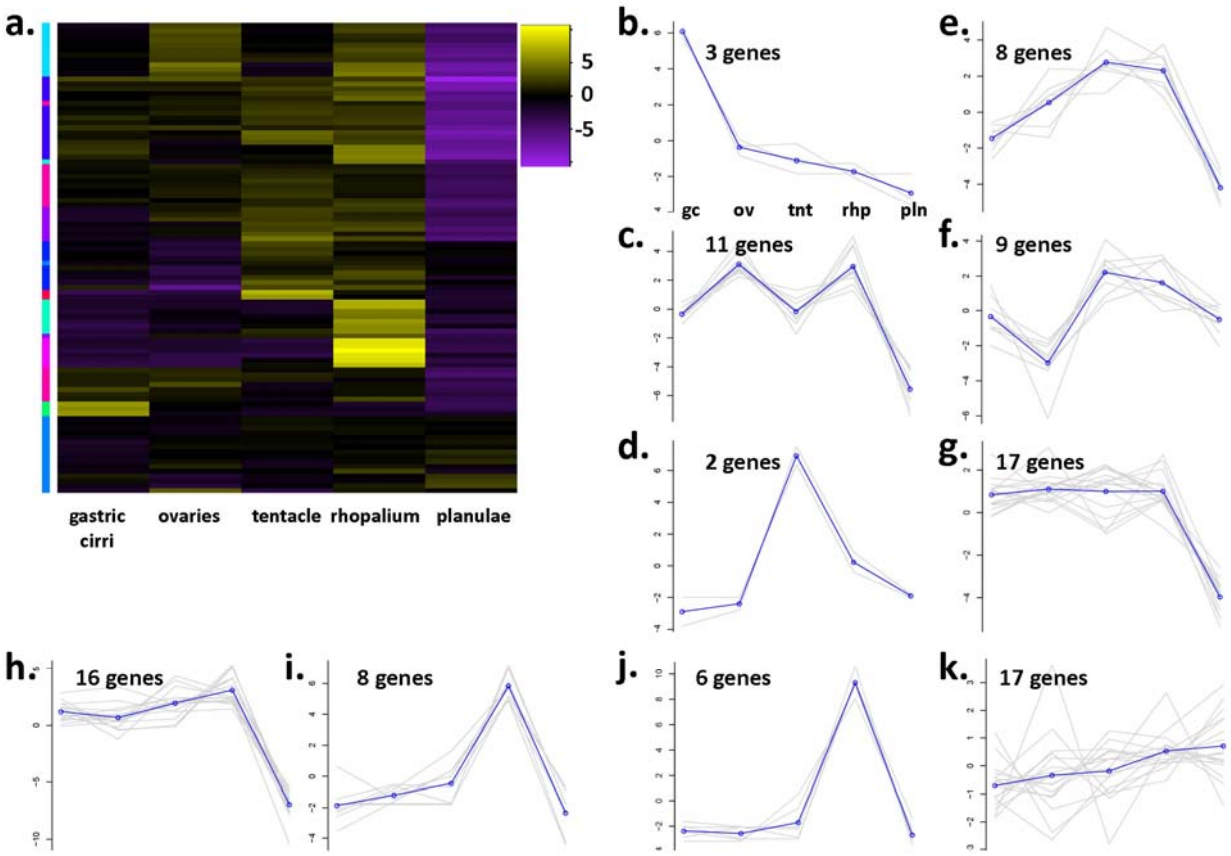

Figure 3.7 a-k. Vision Heatmap for *A. alata*.
Hierarchical clustering (EdgeR) and corresponding ten subcluster profiles for the 97 genes implicated in vision and the phototransduction pathway differentially expressed across *A. alata* medusa (gastric cirri, ovaries, tentacle, rhopalium) and planulae samples. Intensity of color indicates expression levels for each of the ten hierarchical clusters (vertical access). Bright yellow patches (highly up-regulated) correspond to the highest peaks for each k-mean subcluster profile. K-mean profiles (b-k) match the order of column names in a., representing the mean expression of gene clusters highly up-regulated in each sample (centroid demarcated by the solid line). One bright yellow transcript clusters in the gastric cirri column correspond to a peak in plot b; one cluster in the ovaries column corresponds to plot c; three clusters in the tentacle column correspond to plots d, e and f; three bright yellow clusters in the rhopalium column correspond to peaks in plots h, I and j, and two less intense clusters correspond to peaks in plots g and k, and a slightly intense yellow gene cluster in the planulae column corresponds to the peak in plot k. The vertical colored bar on the left of the heatmap (a.) indicates distinct patterns corresponding to the ten subcluster profiles (sc=subcluster number), for which the number of genes each comprises is indicated. Abbreviations: gc=gastric cirri, ov=ovaries, tnt=tentacle (and pedalium base), rhp=rhopalium, and pln=planulae. Gene annotations by subcluster provided in Additional file 3.VIII.

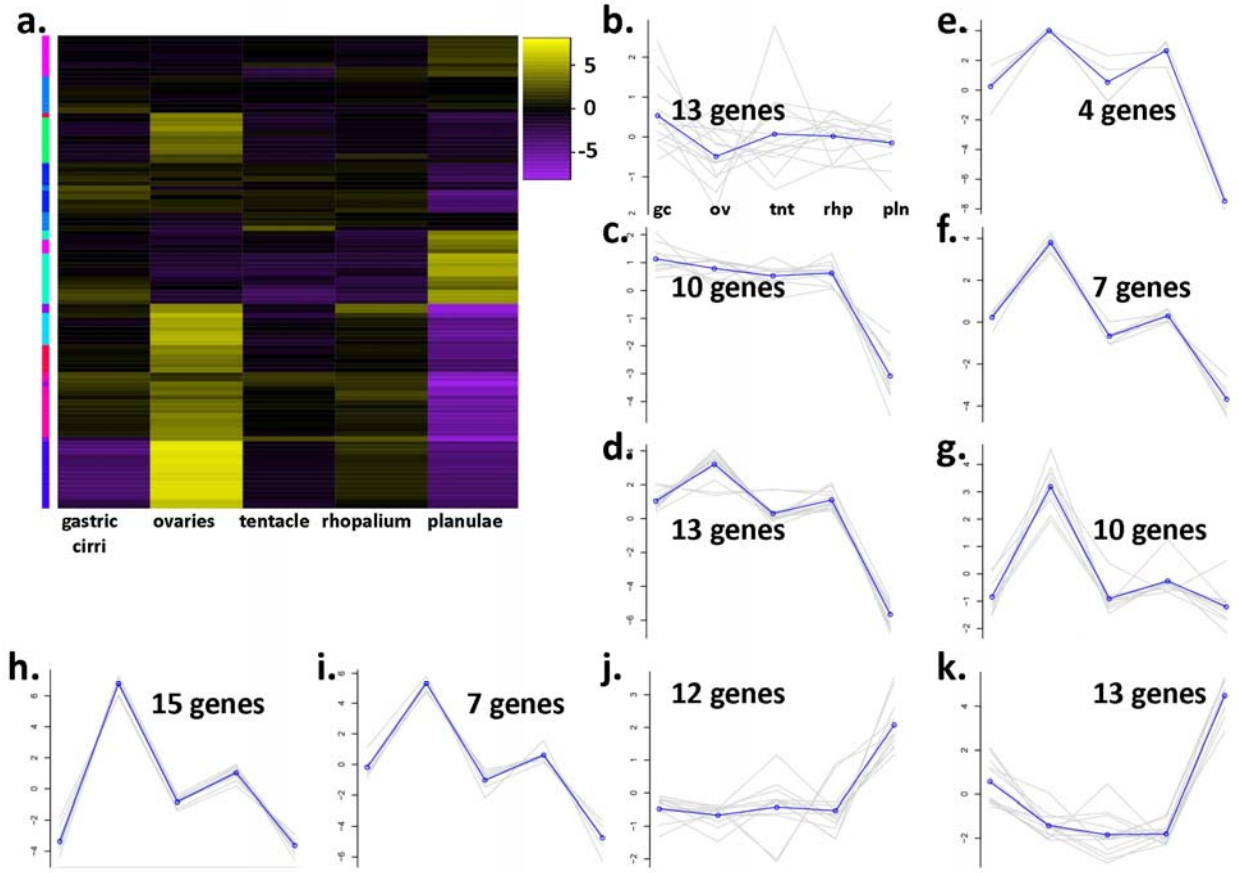

Figure 3.8 a-k. Sex Heatmap for *A. alata*.
Hierarchical clustering (EdgeR) and corresponding ten subcluster profiles for the 104 genes implicated in sex and early development differentially expressed across *A. alata* medusa (gastric cirri, ovaries, tentacle, rhopalium) and planulae samples. Intensity of color indicates expression levels for each of the ten hierarchical clusters (vertical access). Bright yellow patches (highly up-regulated) correspond to the highest peaks for each k-mean subcluster profile. K-mean profiles (b-k) match the order of column names in a., representing the mean expression of gene clusters highly up-regulated in each sample (centroid demarcated by the solid line). Two yellow gene clusters in the gastric cirri column correspond to peaks in plot b and c; six bright yellow clusters in the ovaries column corresponds to profiles d-i; no major up-regulated gene clusters were detected in the tentacle column; four less intense clusters in the rhopalium column correspond to peaks in plots b, d, e, and h; two bright clusters in the planulae column correspond to peaks in the subcluster profiles j and k. The vertical colored bar on the left of the heatmap (a.) indicates distinct patterns corresponding to the ten subcluster profiles (sc=subcluster number), for which the number of genes each comprises is indicated. Abbreviations: gc=gastric cirri, ov=ovaries, tnt=tentacle (and pedalium base), rhp=rhopalium, and pln=planulae. Gene annotations by subcluster provided in Additional file 3.IX.

Putative venom implicated genes

Here we highlight our findings of the 450 differentially expressed transcripts we broadly refer to as "venom implicated genes" based on preliminary candidate gene profiling (above). This effort focused specifically on identifying genes highly expressed in the tentacle (used in prey capture) and gastric cirri (used in digestion). By comparing a body part with tissue abundant in penetrant nematocysts (tentacle and adjoining pedalium base) (Fig 3.1e) with one lacking nematocysts (gastric cirri Fig 3.1b) our aim was to identify putative site(s) of venom production and nematogenesis in *A. alata*. Although nematocysts are typically abundant in the gastric cirri of many box jellyfish species (Gershwin 2006), only a single individual nematocyst has been documented in *A. alata* gastric cirri (Lewis *et al.* 2013) despite examination of hundreds of samples by several independent studies (Arneson & Cutress 1976; Gershwin 2005, 2006; Lewis *et al.* 2013). Conversely, nematocysts (Fig 3.1e) are primarily concentrated in the tentacle (contiguous with the pedalium) of *A. alata* (Lewis *et al.* 2013), as is the case with all box jellyfish (Gershwin 2006). Hierarchical gene cluster profiling (Fig 3.6 a-k) revealed that many of the putative venom implicated genes were fittingly highly expressed in the tentacle (Fig 3.5c; Fig 3.6e-i), but were surprisingly also highly up-regulated in the gastric cirri (Fig 3.5a; Fig 3.6b,c; Additional 3.VII).

CaTX/CrTX toxin family genes. We identified eleven different homologs of the CaTX/CrTX toxin family (also annotated individually as CrTX-A or CaTX-A). These were either up-regulated almost exclusively in the tentacle (n=4) (Fig 3.5c; Fig 3.6e, f) or in the gastric cirri (n=7) (Fig 3.5a; Fig 3.6b, c). This gene family consists of pore-forming toxins that cause pain, inflammation and necrosis during human envenomation (Brinkman & Burnell

2009) and prior to this study has exclusively been associated with venom from nematocysts (Jouiaei *et al.* 2015a; Nagai 2003; Ponce *et al.* 2016). The CaTX/CrTX toxin family (Nagai *et al.* 2000a; b), previously called the "box jellyfish toxin family" (Brinkman & Burnell 2009; Nagai *et al.* 2000a; b), was thought to be restricted to medusozoans (Brinkman *et al.* 2015), but recently a homolog was also identified in the coral *Acropora* (Gacesa *et al.* 2015). Gene tree reconstruction (Fig 3.9) of the eleven CaTX/CrTX gene homologs up-regulated in either the gastric cirri or the tentacle confirmed homology of the *A. alata* transcripts with CaTX/CrTX toxin family genes in other cnidarians (Brinkman *et al.* 2015; Gacesa *et al.* 2015; Ponce *et al.* 2016). The analysis recovered four well-supported groups of *A. alata* CaTX/CrTX genes, each exclusively containing genes with tissue-specific expression patterns, either in tentacle or gastric cirri (Fig 3.9).). One group includes three *A. alata* homologs (annotated as CaTX/CrTX, CaTX or CrTX) specific to the tentacle that group with several non-cubozoan medusozoans including the coral *Acropora.* Three additional groups are all within a well-supported cluster of CaTX/CrTX genes identified from cubozoan taxa. One of these nested groups includes the homolog (CaTX-A) reported more than a decade ago (Nagai 2003) in *A. alata* (as *Carybdea alata*). Homologs of this gene are sister to a sub-group comprised of *Chironex fleckeri* homologs, which have been exclusively identified from tentacle tissue. The two additional nested groups (Fig 3.9) include the homolog (CrTX-A) reported more than a decade ago in *Carybdea brevipedalia* (reported as *C. rastonii*), comprising transcripts only identified in our gastric cirri sample, in a well-supported cluster of homologs derived from *Carybdea brevipedalia* and *Malo kingi*. Ours is the first report of expression of the CaTX/CrTX toxin family in a medusozoan body part that lacks nematocysts.

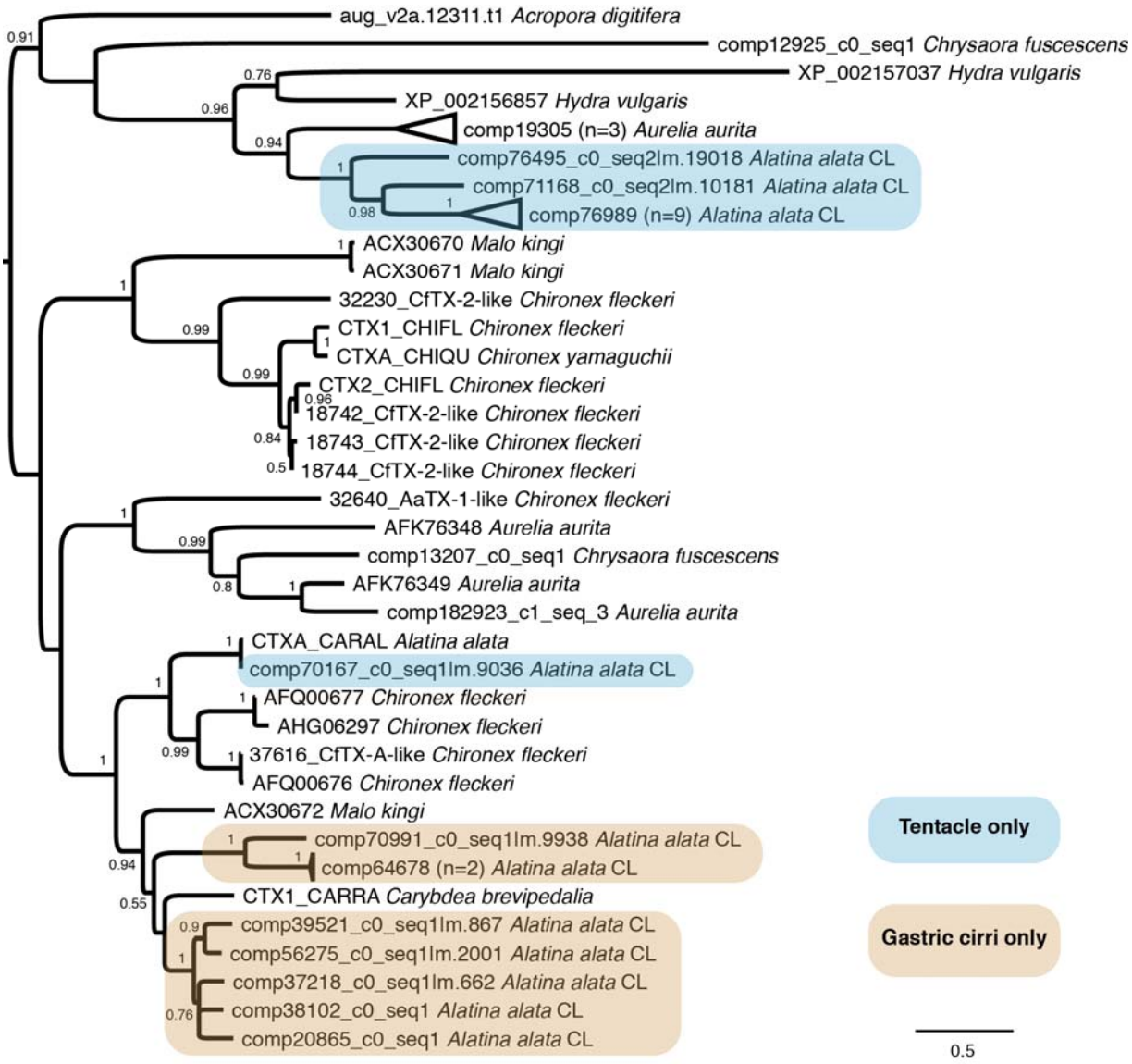

Figure 3.9 Cnidarian CaTX/CrTX toxin family gene tree.
ML topology of all known homologs of the CaTX/CrTX toxin family in cnidarian taxa from NCBI Genbank and transcriptome components of *A. alata* in this study. Assumes the WAG+G+F model of amino acid evolution, as specified as most appropriate by ProtTest v. 3.2. Shimodaira-Hasegawa-like branch support indices are shown at each node. Tissue-specific expression patterns correspond to transcripts primarily enriched in the tentacle (in blue) and gastric cirri (in beige). In *A. alata*, the expression of genes annotated as the CaTX-A homolog is specific to the tentacle sample (comp70167, comp71168), while genes annotated as the CrTX-A homolog are specific to gastric cirri and a single tentacle gene (comp76495). Additionally, the expression of a single gene (comp76989) annotated as both CrTX and CaTX (i.e., from TrEMBL and SwissProt databases) is tentacle specific.

Venom components. We report that up-regulation of "cysteine-rich secretory protein family" (CRISPs) occurred almost exclusively in either the gastric cirri or the tentacle. Some examples include: "serine protease coagulation factor vii", "chymotrypsin-like elastase family" homologs, and "serine protease inhibitor" (Fig 3.5a; Fig 3.6c). Likewise, multiple homologs of the "zinc metalloproteinase/astacin (peptidase family m12a)" (Fig 3.5a, c; Fig 3.6b, c, f, i, k) were primiarily up-regulated in the gastric cirri, but with high expression in the tentacle as well. Zinc metalloproteinases are peptidases with known roles in venom maturation in spiders and snakes, and were recently identified as tentacle venom components of some jellyfish taxa (Brinkman *et al.* 2015; Jouiaei *et al.* 2015a; b). Conversely, homologs of well-known bilaterian venom proteins (e.g., pit viper (*Croatulus*)/zinc metalloproteinase nas-4/venom factor (Fig 3.6i); scorpion (*Lychas*), venom protein 302 (Fig 3.6h); the venom prothrombin activator pseutarin-c non-catalytic subunit from the eastern brown snake (*Pseudonaja textilis*) (Fig 3.6i); and "alpha-2-macroglobulin family N-terminal region" (Fig 3.6i) were most up-regulated in the gastric cirri and tentacle in this study, but were also expressed in the ovaries and rhopalium samples.

Nematocyst structural genes. Genes encoding putative nematocyst structural proteins were also characterized in this study. This is in line with our aim to characterize molecular components of cubozoan nematocysts and pinpoint putative regions of nematogenesis, given the current view that venom deployment in medusozoans is exclusively controlled by nematocysts. Our findings revealed that three minicollagens, key components in nematocyst capsule development (David *et al.* 2008), were up-regulated almost exclusively in the *A. alata* tentacle (with adjoining pedalium base) (Fig 3.5c; Fig 3.6e, h), with slight expression signal in all other medusa samples. This is the first description of minicollagen homologs in the class Cubozoa. All three of the minicollagen genes identified from *A. alata* possess a collagen-like domain of

repeated tripeptides of the form GlyXY flanked on both sides by proline repeats all within N- and C-terminal cysteine rich domains (CRD). The CRDs for two of the three genes are of the regular form (CXXXCXXXCXXXCXXXCC) and thus would be classified as Group 1 minicollagens (Shpirer *et al.* 2014). The third has a regular CRD at the C-terminus but a variant form at the N-terminus. Gene tree reconstruction of minicollagen genes for cnidarian taxa (Fig 3.10) show that the minicollagens identified for *A. alata* cluster mostly with other Group 1 minicollagens. Of the additional nematogenesis related genes (Balasubramanian *et al.* 2012; David *et al.* 2008; Eckert 2012; Engel *et al.* 2002; Forêt *et al.* 2010; Horibata *et al.* 2001; Houliston *et al.* 2010; Milde *et al.* 2009) differentially expressed in this study, most were almost exclusively up-regulated in the tentacle (with adjoining pedalium base). They include "nematocyst outer wall antigen" (NOWA) (Fig 3.6e), "chondroitin proteoglycan 2" (Fig 3.6f), "nematoblast-specific protein nb035-sv2/nb035-sv3/nb012a" (Fig 3.6e, f), "nematogalectin-related protein" (Fig 3.6e), and "Dickkopf-related protein 3" (Fig 3.6 e, f, h). Five different "Dickkopf-related protein 3" homologs were up-regulated in *A. alata* tentacle sample; with only two having notable expression in other body parts (gastric cirri, ovaries or rhopalium).

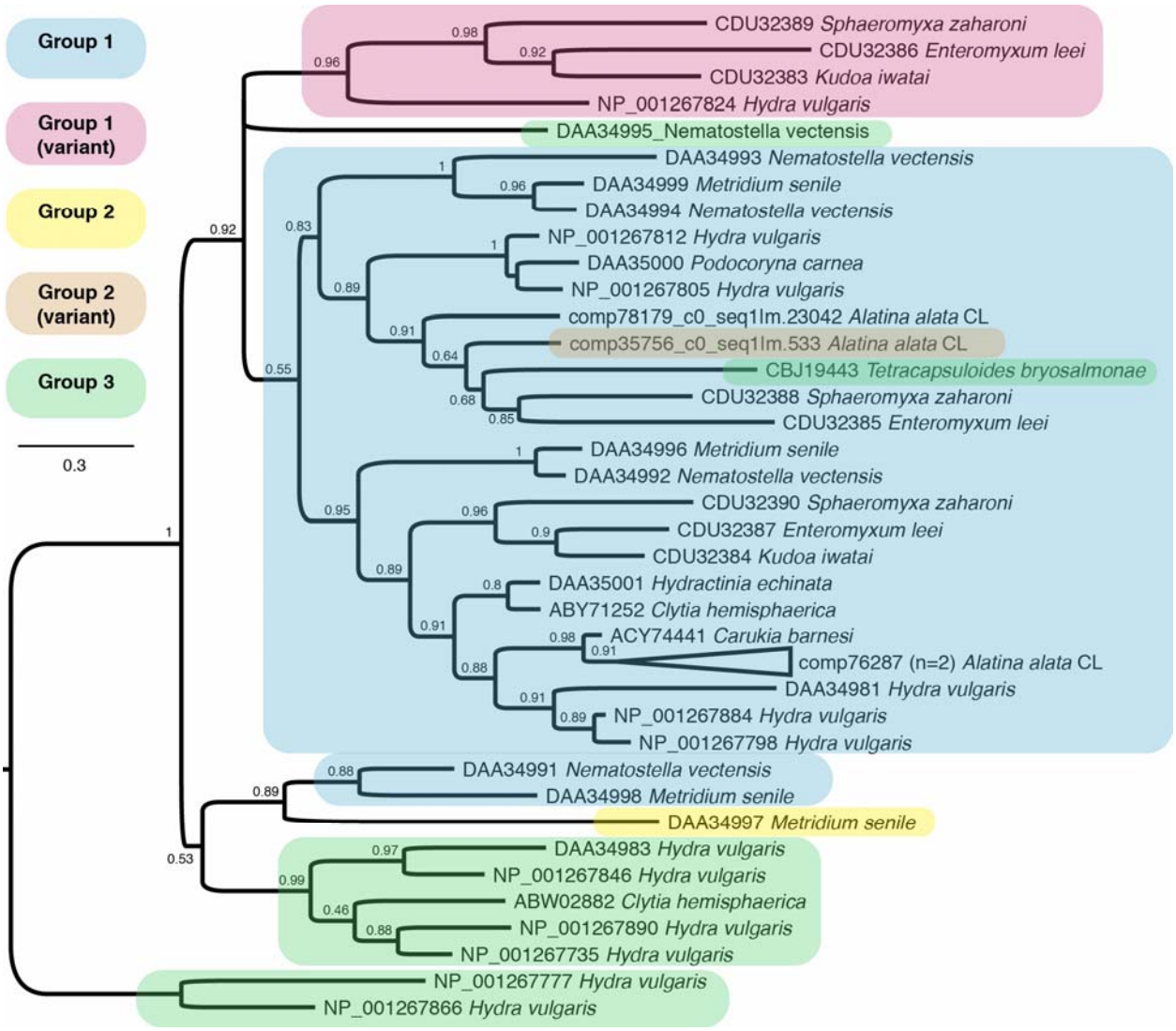

Figure 3.10 Cnidarian minicollagen gene tree.
ML topology of minicollagen gene family in cnidarian taxa from NCBI Genbank and transcriptome components of *A. alata* in this study. Assumes the BLOSS62I-G-F model of amino acid evolution, as specified as most appropriate by ProtTest v. 3.2. Shimodaira-Hasegawa-like branch support indices are shown at each node. All *A. alata* minicollagen types were more closely related to non-cubozoan homologs. Following Shpirer et al. 2014, Group 1 minicollagens possess N- and C-terminal cysteine rich domains (CRDs) of a regular form (CXXXCXXXCXXXCXXXCC). We use the label Group 1 (variant) to identify those minicollagens having three or more regular CRDs. Group 2 minicollagens possess one regular and one irregular CRD, with the variant form having the regular CRD at the C-terminus, whereas Group 3 minicollagens possess irregular CRDs at both termini. One (comp76287_c0) of the three recovered minicollagen homologs from *A. alata* was differentially expressed across all five samples, but most up-regulated in the tentacle.

Putative vision implicated genes

Here we highlight our findings of the 97 differentially expressed transcripts we broadly refer to as "vision implicated genes" based on preliminary candidate gene profiling (above). This effort focused specifically on genes expressed in the rhopalium of *A. alata*, which bears a pair of lens eyes with cornea and retina, two pairs of simple ocelli-comprising photoreceptors, and a statocyst (Fig 3.1f). By comparing the rhopalium with its visual capabilities and planulae with its known photoreceptors, against the medusa samples that lack known photoreceptors (gastric cirri, ovaries and tentacle), our aim was to identify the expression of opsins and other vision implicated genes in the rhopalium of *A. alata*, as well as in putative extraocular photoreceptors in *A. alata*. Hierarchical gene cluster profiling (Fig 3.7a–k) revealed that most of the 97 putative vision implicated genes (see "gene-profiling" above) were up-regulated in the rhopalium (Fig 3.5d), but in many cases they were more highly expressed in other samples, in particular in the medusa (Fig 3.7c, e-j; Additional 3.VIII).

Opsins. In this study the Trinotate report corresponding to the *A. alata* transcriptome contained a total of 41 transcripts with PFAM annotations corresponding to homologs of the "7 transmembrane receptor (rhodopsin family)". Of the rhodopsin family, opsins are considered universal light sensitive proteins associated with photoreceptor cells of animal retinas. We found eleven opsin homologs to be variably expressed across *A. alata* medusa samples, with only six homologs most up-regulated in the rhopalium (Fig 3.6g, h; Fig 3.7 i, h). A gene tree reconstruction of all rhodopsin family cnidarian genes (Fig 3.11) , which we rooted on the group that includes all previously known medusozoan opsins, recovered two of the three previously identified cnidarian opsin groups (Feuda *et al.* 2012). Group A includes only anthozoan taxa, while Group B includes all opsins previously known from medusozoans, eleven transcripts from

Alatina and a few from anthozoans. However, cnidarian opsins previously identified as Group C fell into two groups – one of which includes thirty *A. alata* transcripts and is sister to group B plus C. This appears to be the first example of medusozoan opsins outside group B cnidarian opsins.

Among the top 10 most abundant genes in the rhopalium of *A. alata* was a homolog of the cubozoan lens-eye opsin for *Carybdea rastonii* (=*C. brevipedalia*) (Fig 3.5d; Fig 3.7i). The *Carybdea* lens eye opsin was also highly expressed in all *A. alata* samples, including planulae which have eye spots (Fig 3.1h). Normalized counts revealed three additional opsin genes that were not differentially expressed across all samples, were almost exclusively found in the planulae sample (Fig 3.11). Only two *A. alata* opsin homologs were expressed almost exclusively in the rhopalium (Trinotate top BLAST hits: "dopamine receptor 2" and "visual pigment-like receptor peropsin") (Fig 3.11). Among the putative "rhodopsin family" genes expressed in medusa samples were those annotated in the Trinotate report as non-opsin based photoreceptors, such as: "compound eye opsin bcrh1/d(1b) dopamine receptor" (Fig 3.7e) and "opsin rh1/mu-type opioid receptor" (Fig 3.7h), "blue-sensitive opsin" (Fig 3.7k), "visual pigment-like receptor peropsin" (Fig 3.7d), "melanopsin-b" (Fig 3.7b; Fig 3.8b; Additional 3.VIII, 3.IX).

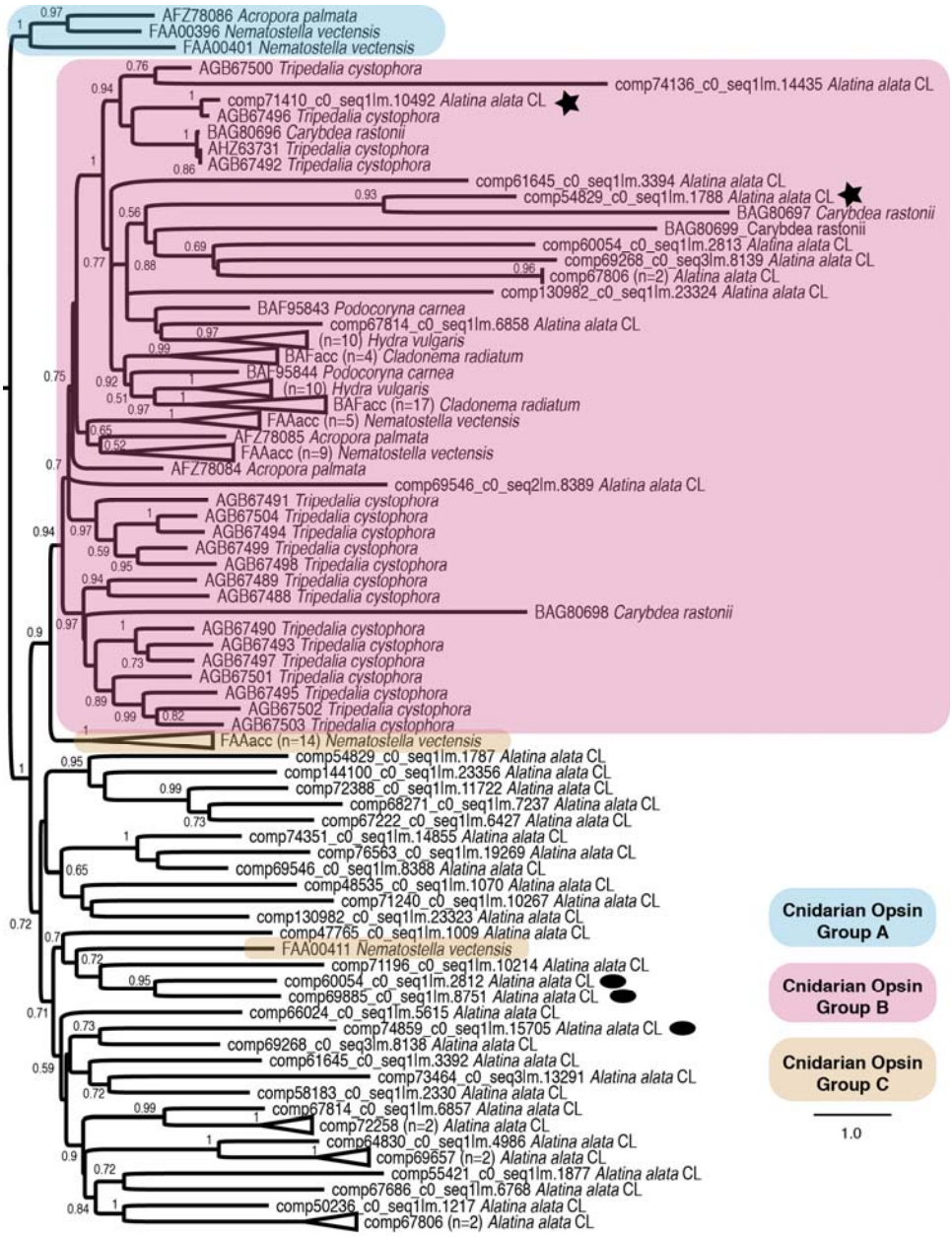

Figure 3.11 Cnidarian opsin gene tree.
ML topology of all known homologs of the opsin gene family in cnidarian taxa from NCBI Genbank and transcriptome components of *A. alata* in this study. Assumes the LG+G model of amino acid evolution, as specified as most appropriate by ProtTest v. 3.2. Shimodaira-Hasegawa-like branch support indices are shown at each node. Blue, pink and brown shading correspond to cnidarian opsin groups A, B and C, respectively, recognized by Feuda et al. (2012). Group A is arbitrarily chosen for rooting the topology, and Group B is monophyletic. Group C may not be monophyletic, as some Group C homologs cluster within a clade that is sister to the B plus C group. Stars denote opsin transcripts almost exclusively up-regulated in the *A. alata* rhopalium sample (comp54829, comp71410) and ovals denote the same in the planulae sample (comp60054, comp69885, comp74859).

Chromophores. We found that several isozymes of *cis*-retinol dehydrogenase, members of the retinoic acid signaling pathway, which convert retinol (vitamin A) to retinal, were expressed in the rhopalium, and in other samples including planulae (Fig 3.7c, g, k). In animals retinal (i.e., 11-cis retinal) is bound to opsin on the photoreceptors of the retina (Mason *et al.* 2012), and is thought to be a universal chromophore (light-activated pigment), though various chromophores are used across Metazoa (Mason *et al.* 2012). Carotenoid oxygenase beta, beta-carotene 15,15'-monooxygenase (BCDO1) (Fig 3.7g,h) and beta, beta-carotene 9',10'-oxygenase (BCDO2), known to irreversibly cleave carotenoids to produce the essential visual pigments retinal and retinoic acid respectively (Wyss 2004), were expressed in the four *A. alata* medusa samples (Fig 3.7e), suggesting that the catalytic components are present to make retinal. We also report the expression of a putative blue-sensitive photoreceptor protein and circadian clock regulator "cryptochrome-1" in the rhopalium, but with high expression in the gastric cirri (Fig 3.8b). This suggests the presence of an additional putative chromophore in *A. alata* that functions in extraocular blue-light mediated behaviors (e.g., phototaxis), previously documented in coral and other metazoans (Liedvogel & Mouritsen 2010; Reitzel *et al.* 2013; Rivera *et al.* 2012, 2013).

***Crystallins*.** We found transcripts showing similarity to all three known J-crystallin groups (J1, J2, J3), and all were highly expressed in the rhopalium (Fig 3.5d; Fig 3.7i, j). The J2 crystallin homolog was expressed in all samples including planulae (Fig 3.7i); J3 crystallin was almost exclusively expressed in the rhopalium (Fig 3.7j); as were all but a single J1 crystallin homolog that was also abundant in the ovaries (Fig 3.6d; Fig 3.7j). Crystallins are water-soluble stable structural proteins that provide transparency and increase the refractive index of eye lenses, though most also have roles unrelated to lens function. Numerous types of crystallins are

found across Metazoa, and many are identical (or closely related) to commonly expressed metabolic enzymes or stress proteins (Kozmik *et al.* 2008a; b; Piatigorsky *et al.* 1988). J-crystallins are classified in three evolutionarily independent groups, J1, J2 and J3, and thus far have only been reported in cubozoans (Kozmik *et al.* 2008b; Piatigorsky *et al.* 1989, 1993). A study on *T. cystophora* showed that the promotors of all three J-crystallin genes can be activated by the paired domain transcription factor PaxB, but the promotor sites are non-homologous among the three J-crystallin types (Kozmik *et al.* 2008b). We aligned all known cubozoan J-crystallins (J1, J2, J3) with the respective *A. alata* homologs identified in this study (Fig 3.12). The resulting three alignments illustrate the similarity between *A. alata* transcripts identified in this study and homologs of the three distinct *T. cystophora* J-crystallin types. Additionally, Alpha-crystallin B chain" (vertebrate lens heat-shock proteins) (Fig 3.7i) was highly up-regulated in the rhopalium, but expressed in all five samples. Conversely, we report transcripts annotated as S-type crystallin (cephalopod lens protein) variably expressed across samples: S-crystallin 2 up-regulated in the rhopalium and absent in ovaries (Fig 3.7f), S-crystallin 3 up-regulated almost exclusively in the tentacle (Fig 3.7d), and S-crystallin 4 most highly expressed in planulae and gastric cirri (Fig 3.7k).

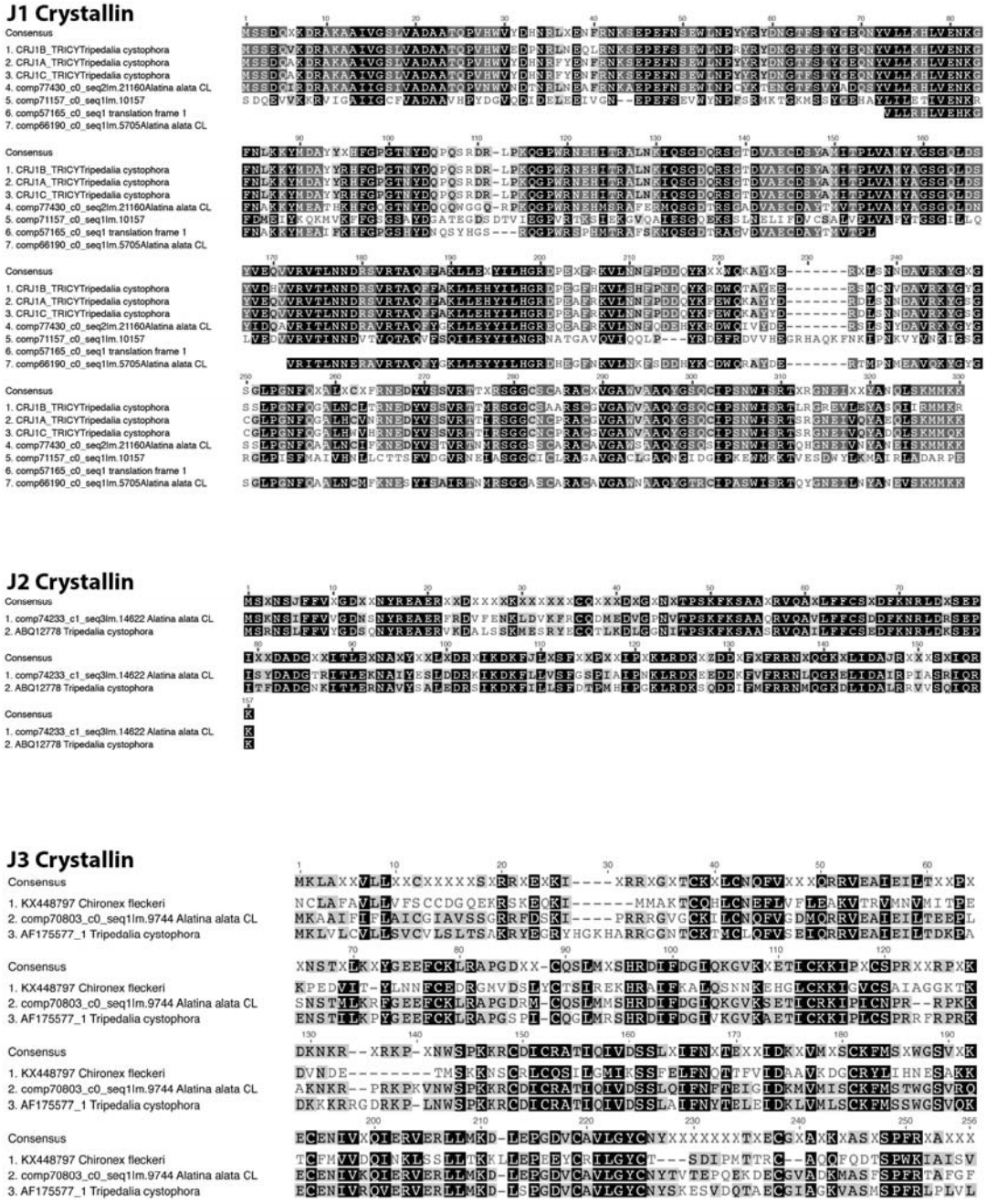

Figure 3.12 Cubozoan J-Crystallins alignment.
Figure corresponding to the amino acid sequence alignments of the three non-homologous cubozoan J-crystallins (J1, J2, J3). Proteins corresponding to *T. cystophora* J1, J2 and J3 crystallins and to *C. fleckeri* J3 crystallins from NCBI were aligned against the respective J1, J2 and J3 crystallin homolog for *A. alata* in this study. Boxes around residues indicate similarity in amino acid sequence, with black boxes corresponding to consensus regions. Sequences were aligned using MUSCLE (default parameters with 5 iterations). *A. alata* amino acid sequences correspond to predicted ORFs (TransDecoder), except for comp57165 which is a frame 1 translation of the Trinity transcript.

Homeobox genes and transcription factors. Expression of putative homeobox proteins "Six1b" and "Six4" and the "Six" transcriptional co-activator "eyes absent" (Eya) homolog occurred in all *A. alata* medusa samples, with the highest expression in the gastric cirri. Across Metazoa, the Six-Eya complex functions downstream from certain Pax homeobox genes in a diversity of developmental processes including early eye development (Kumar 2009). The putative "retinal homeobox proteins rx1b" and "rx3" were expressed in all samples, except for in the ovaries in the case of "rx1b". Retinal homeobox proteins ("rax" or retina and anterior neural fold homeobox) are essential for early eye-development and in regulation of stem cell proliferation in vertebrates (Furukawa *et al.* 1997), but have not previously been reported in cnidarians.

Putative sex and development implicated genes

Here we highlight our findings of the 104 differentially expressed transcripts we broadly refer to as "sex implicated genes" based on preliminary candidate gene profiling (above). This effort focused mainly on genes expressed in the ovaries of A. alata during ovulation and internal fertilization. By definition, the ovaries are the site of oogenesis, and are situated within the gastrovascular cavity in A. alata (Lewis et al. 2013). Microscopic examination of the gastrovascular cavity of the female A. alata medusa in this study revealed ovulation (Fig 3.1c) and internal fertilization (Fig 3.1d) occurring within this cavity. By comparing our ovaries sample, which also contained zygotes and embryos, with other body parts predicted to lack reproductive material (gastric cirri, tentacle, rhopalium, and planulae), our aim was to identify genes involved in gametogenesis as well as to determine more precisely the location of internal fertilization within A. alata, which we expected would occur adjacent to the ovaries, in the sperm-saturated gastrovascular cavity (Lewis et al. 2013)). Hierarchical gene cluster profiling

(Fig 3.8a-k) revealed that the highest expression of the 104 putative sex and developmental genes (see "gene-profiling" above) occurred in the ovaries (Fig 3.8d-i), but that many were also expressed across all samples (Additional 3.IX).

Oogenesis and embryogenesis. We found that homologs for the putative large lipid transfer protein Vitellogenin-2 were most up-regulated in A. alata ovaries (Fig 3.6i; 8e), though highly expressed in all medusa samples. Conversely, genes annotated as Vitellogenin-1 or simply Vitellogenin were expressed in all medusa samples, but most up-regulated in the tentacle (Fig 3.6i,g). Likewise, genes annotated as Apolipophorin or Apolipoprotein B-100, the other major animal protein group involved in lipoprotein processing (Hayward et al. 2010), were expressed across all medusa samples but most abundant in the tentacle (Fig 3.6g-i). Genes of the Vitellogenin family are responsible for lipid transfer from ovarian follicle cells to oocytes, providing nutrition during embryogenesis in bilaterians and some cnidarians (Hayward et al. 2010; Levitan et al. 2015), and have a documented role as egg yolk protein precursors in the ovaries of animals, including anthozoans (Shikina et al. 2013). Gene tree reconstruction of A. alata transcripts with known cnidarian Vitellogenin and Apolipophorin-like proteins (Fig 3.12) confirmed their homology with other cnidarian large lipid transfer proteins.

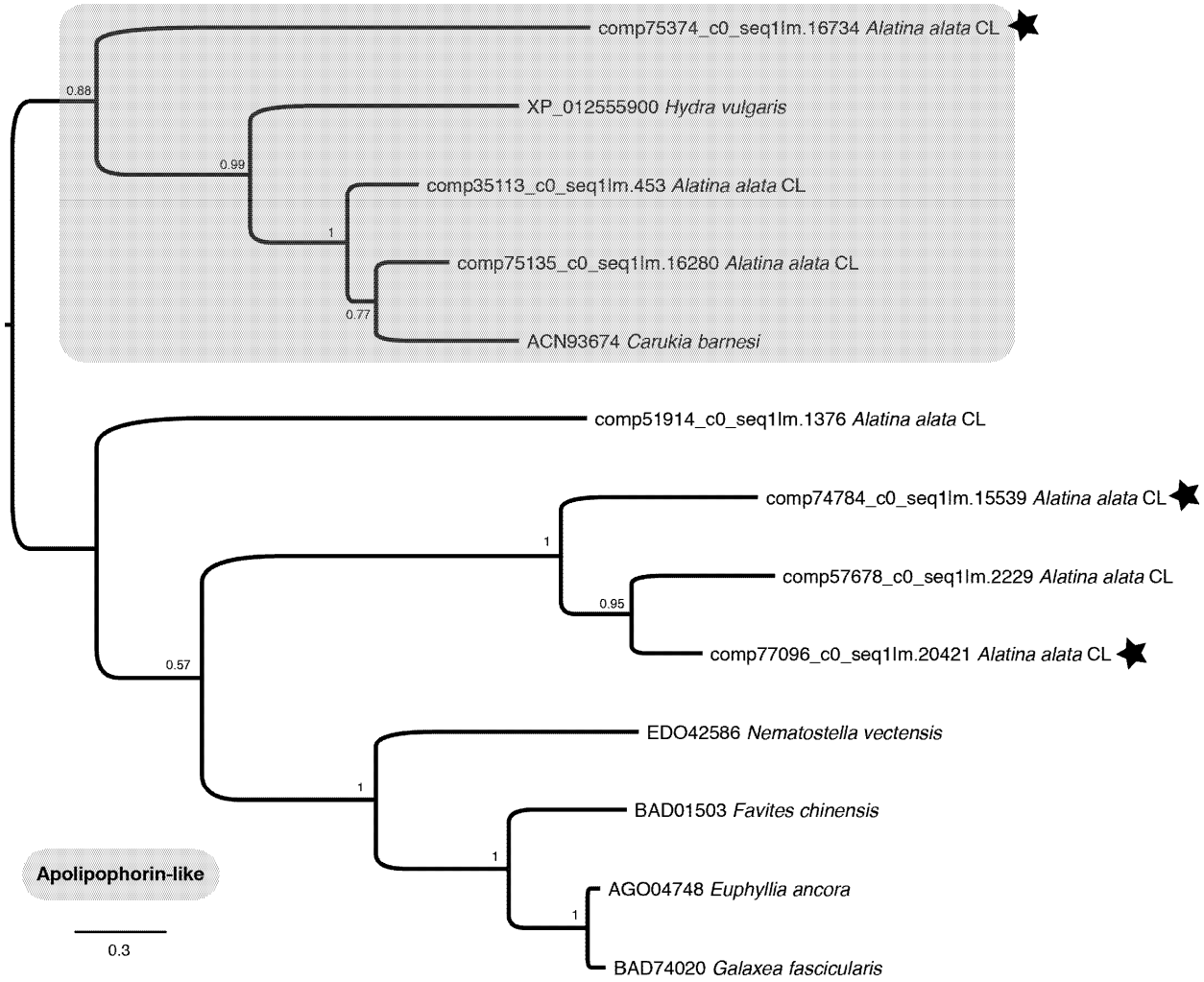

Figure 3.13 Cnidarian Vitellogenin gene tree.
ML topology of all known homologs of the Vitellogenin gene superfamily and apolipophorin-like putative Vitellogenin precursor in cnidarian taxa from NCBI Genbank and transcriptome components of *A. alata* in this study. Assumes the LG+G model of amino acid evolution, as specified as most appropriate by ProtTest v. 3.2. Shimodaira-Hasegawa-like branch support indices are shown at each node. Grey (top) highlights a clade of apolipophorin-like putative Vitellogenin precursor homologs for cubozoan *Carukia barnesi* and *A. alata* from this study and hydrozoan *Hydra vulgaris*, while the non-highlighted clade (bottom) highlights the relationships among known Vitellogenins of cnidarian taxa. Stars designate transcripts most upregulated in the ovaries (comp74784, comp75374 & comp77096); all others are either most highly expressed in the tentacle (comp35113 & 51914) or equally high in the gastric cirri and tentacle samples (comp57678 & comp75135).

We report the upregulation of multiple creatine kinase isozymes in the ovaries (Fig 3.5b), all of which were expressed to some degree in all samples including planulae (Fig 3.6d; Fig

3.8h). Creatine kinase activity has a documented role in oogenesis and early embryogenesis in mammals (Iyengar *et al.* 1983). More broadly though, creatine kinase is important in cells with variable rates of energy turnover, such as muscle, neurons, photoreceptors, and primitive spermatozoa (Pineda & Ellington 1999), which is consistent with its expression in all samples in this study.

Sperm motility. Creatine kinase isozymes have also been documented in mediating high energy phosphate transport between sperm mitochondria and sperm flagellar tail (Pineda & Ellington 1999; Wallimann & Hemmer 1994). Among the genes most up-regulated in the ovaries (Fig 3.5b) were homologs of genes functioning in sperm tail development and motility: "parkin coregulated gene protein homolog" (Fig 3.8i), "outer dense fiber protein 3/sperm-tail pg-rich repeat/shippo-rpt" (Fig 3.8h), and multiple putative creatine kinase isoenzymes including "testis isozyme/protoflagellar creatine kinase" (Fig 3.8j, h). Likewise multiple putative "serine/threonine-protein kinase" isoenzymes including "testis-specific serine/threonine-protein kinase 1" (Fig 3.8g, h) were primarily up-regulated in the ovaries, but also variably expressed in all five samples. Expression in all samples of these sperm-related genes is consistent with the presence of ubiquitous sperm documented within the female gastrovascular cavity (where the ovaries are found) facilitating internal fertilization in this study (Fig 3.1d). Sperm were also abundant in the surrounding seawater, and were undoubtedly adhered to the tentacles and exterior of the medusa bell when all tissue samples were excised from *A. alata*.

Sperm capacitation. We report the exclusive upregulation in the ovaries of the putative sperm hyperactivation and acrosomal vesicle reaction promotor protein "cation channel sperm-associated protein" (CatSper2) (Fig 3.8g). CatSper genes belong to the family of voltage-gated Ca^{2+} channels that are crucial for sperm fertility in mammals (Rahman *et al.* 2014). In particular,

capacitation, which typically occurs within the female reproductive tract, involves the destabilization of the acrosomal sperm head membrane allowing greater binding between sperm and oocyte during fertilization due to an increased permeability of Ca^{2+} (Rahman *et al.* 2014). Recent studies have showed that sperm of several invertebrate species also undergo capacitation (for a review see (Aungsuchawan *et al.* 2011)). However, until now the possibility of capacitation occurring in sperm of non-bilaterian invertebrates has not been investigated. Unexpectedly, the "CUB and zona pellucida-like domain-containing protein" was most highly up-regulated in the tentacle and gastric cirri (Fig 3.6h, j), despite one of its known roles in ova of attracting sperm to eggs for fertilization in mammals (Adonin *et al.* 2012). The CUB and zona pellucida-like domain-containing protein is also associated with trypsinogen activation and was previously found in box jellyfish tentacles (Brinkman *et al.* 2015). Overall, the abundance of transcripts related to sperm dynamics identified in the *A. alata* transcriptome permitted unforeseen profiling of molecular components involved in putative sperm capacitation and fertilization for the first time in a cubozoan.

Discussion

This study has generated the first annotated transcriptome from multiple tissues of the cubozoan *Alatina alata*, focusing on both the adult (medusa) and larvae (planulae). Our transcriptome significantly adds to the genomic resources available for this emerging cubozoan model. This transcriptome, based primarily on multiple adult body tissues, complements a recently published transcriptome for the same species that focused primarily on early developmental stages (Zapata *et al.* 2015). Furthermore, in this study we annotated a large set of genes, allowing for an initial characterization of the molecular complexity of this cubozoan. We

also compared transcript abundance across samples to identify genes putatively involved in several key features of cubozoans, namely nematogenesis and venom production, vision and sensory perception, and sexual reproduction. These quantitative data should be considered preliminary, due to lack of replication, but they are suggestive of interesting candidate genes that will be useful for future study. Below we highlight some of the major findings from this initial comparison across samples focusing specifically on genes relevant to i) prey capture and defense, ii) vision and the phototransduction pathway and iii) sexual reproduction and embryogenesis.

Prey capture and defense

In cubozoans, and more broadly in all cnidarians, prey capture and defense are based on nematocyst (stinging organelles) and associated venom. By comparing a body part abundant in penetrant nematocysts (tentacle and adjoining pedalium base) with one lacking nematocysts (gastric cirri) our aim was to identify putative site(s) of nematocyst development (nematogenesis) and venom production in *A. alata*.

We found that transcripts corresponding to a number of putative nematocyst structural proteins (minicollagens, nematogalectin, NOWA, chondroitin, and Dickkopf homologs) were up-regulated in the tentacle (and adjoining pedalium base). Although, putative nematogenic transcripts were detected primarily in the tentacle, some were also detected in non-tentacle medusa samples. We expect that this signal stems from the abundant adherent nematocysts covering the medusa bell. Together these findings are consistent with nematogenesis in *A. alata* occurring primarily, but not solely, in the region comprising the tentacle and adjacent pedalium

base. Future *in situ* hybridization studies employing genes identified in nematogenesis in this study can help pinpoint more precisely nematogenic regions in *A. alata*.

Venom is a complex cocktail of bioactive compounds (e.g., protein and/or peptides called toxins, salts and neurotransmitters) secreted by one animal that is delivered to another animal by an infliction (Casewell *et al.* 2012; Fry *et al.* 2009). Venom disrupts physiological and biochemical molecules of prey and predators, thus facilitating feeding and defense (Casewell *et al.* 2012). Nematocysts have long been considered the sole secretory structure for venom deployment in cnidarians (Galliot 2012). However, we found preliminary evidence for venom production in the gastric cirri, where nematocsyts are lacking in *A. alata*. Futhermore, we found that the gastric cirri and tentacle express distinct groups of homologs of a major family of medusozoan venom proteins, the CaTX/CrTX toxin family (Nagai *et al.* 2000a; b). This suggests that venom plays an important, and possibly different role in the gastric cirri and tentacle. Venom components likley differ between the nematocyst-bearing tentacle, with a primary role in immobilizing prey and warding off predators, and the gastri cirri, with a primary role in killing and digesting prey (Larson 1976).

Based on our findings, we hypothesize that *A. alata* has gland cells that secrete toxins associated with the gastric cirri. Evidence was recently presented for toxin-secreting gland cells in the ectoderm of the sea anemone *Nematostella* in regions containing nematocysts as well as areas that may lack nematocysts (Moran et al. 2012; Moran et al. 2013), but our findings represent the first putative case in a cubozoan. Future morphological studies examining the ultrastructure of the stomach and gastric cirri, and *in situ* hybridization methods to identify the expression of venom implicated genes, will permit testing of the hypothesis of toxin-secreting gland cells associated with the gastric cirri of *A. alata*.

Although differences exist in the exact complement of putative bioactive toxins between the gastric cirri and tentacle sample, the venom cocktail in each body part includes transcripts from similar digestive enzyme families. A recent review of jellyfish toxins lists a number of toxin-like digestive enzymes that are deployed as components of nematocyst venom to disable homeostatic processes in prey or predators (Jouiaei *et al.* 2015b), as has been noted in animals possessing venom glands (Calvete *et al.* 2009; Casewell *et al.* 2012; Junqueira-de-Azevedo *et al.* 2015). These bioactive proteins function in cytolytic, paralytic and hemolytic roles, thereby facilitating prey digestion (Brinkman & Burnell 2009; Jouiaei *et al.* 2015b; Zhang *et al.* 2003). Specifically we note the up-regulation of several enzyme groups primarily in either the tentacle or gastric cirri in *A. alata* that have been well studied in venomous animals (Casewell *et al.* 2012; Junqueira-de-Azevedo *et al.* 2015; Möhrlen *et al.* 2006; Trevisan-Silva *et al.* 2010), namely astacin-like metalloproteinase and serine proteinase (and inhibitors), and more broadly cysteine-rich secretory proteins (CRISPs). Metalloproteinase and serine proteinase (and inhibitors) are a common component of the venom of animals with venom glands either activating toxins or acting as toxins themselves (Calvete *et al.* 2009; Jouiaei *et al.* 2015a; Junqueira-de-Azevedo *et al.* 2015). In particular, cysteine-rich secretory proteins identified in snake venoms are thought to inhibit smooth muscle contraction in bite victims (Yamazaki & Morita 2004). Both metalloproteinase and CRISPs have previously been characterized in the tentacles of cubozoan (Avila Soria 2009; Brinkman *et al.* 2015) and other cnidarians (Balasubramanian *et al.* 2012; Gacesa *et al.* 2015; Ponce *et al.* 2015, 2016; Weston *et al.* 2013). The abundance of multiple isozymes of astacin-like metalloproteinase and serine proteinase (and inhibitors) and CRISPs in the gastric cirri and tentacle of *A. alata* suggest a dual role in venom and digestion. Further studies are required to test this hypothesis given the broad involvement of

these bioactive proteins in other biological processes (Brinkman *et al.* 2015; Möhrlen *et al.* 2006).

Vision and the phototransduction pathway

Cubozoans are the earliest diverging animal clade to have image-forming lens eyes, which are part of specialized sensory organs called rhopalia. By comparing a medusa body part bearing conspicuous eyes (the rhopalium) and planulae with eye spots (rhabdomeric photoreceptors) against the medusa samples that lack documented photoreceptors (gastric cirri, ovaries and tentacle), our aim was to profile the molecular components of the opsin-regulated phototransduction pathway and identify additional regions of putative extraocular sensory perception in *A. alata*.

We found that many transcripts with conserved roles in vision (opsins and crystallins) were up-regulated in the rhopalium. Although transcripts with putative roles in light-mediated phototransduction pathway were detected primarily in the rhopalium where eyes are present, their expression was broadly detected across the medusa samples, and in some cases in planulae. We expect this signal stems from the presence of additional photoreceptors (yet undescribed) throughout the body of this cubozoan. Together these findings are consistent with a vision-related role for opsins and crystallins in the lens-eye of the rhopalium, as well as a role in putative photoreceptors within non-rhopalium tissues and in planulae eye spots.

The animal phototransduction pathway is mediated by photopigments in photoreceptors consisting of two parts: a membrane protein (apoprotein) “opsin” and a chromophore “retinal” (vitamin A derivative) (Shichida and Matsuyama 2009). Opsins mediate light as phototypical G protein-coupled receptors in both visual and non-visual systems (Terakita 2005). Currently more than 1000 types of opsin are known across Metazoa, with three subfamilies recognized in

bilaterians: rhabdomeric (r-opsins), Go-coupled plus retinochrome retinal G protein-coupled receptor (Go/RGR) and ciliary (c-opsins) (Shichida & Matsuyama 2009). Studies characterizing opsins in cnidarians have raised the possibility that cnidarian opsins form a monophyletic clade referred to as "cnidops" that is sister to the c-opsins (Bielecki *et al.* 2014; Koyanagi *et al.* 2008; Liegertová *et al.* 2015; Plachetzki *et al.* 2007, 2010). Other metazoan-wide analyses of opsins have categorized cnidarian opsins into three groups, A, B and C, in which each of these cnidarians opsin groups has been found, albeit with limited support, to be sister to each of the respective bilaterian opsin groups (Feuda *et al.* 2012). One study has revealed support for r-opsins in cnidarians (Feuda *et al.* 2014), which is consistent with the identification of planulae eye spots as rhabdomeric photoreceptors (Arendt *et al.* 2009; Nordström *et al.* 2003). Although a consensus is lacking about the relationships between cnidarian opsins and other metazoan opsins, our study identified a number of transcripts with molecular characters corresponding to the rhodopsin family, adding to the known diversity of this gene family within cnidarians.

This study identified a diversity of metazoan opsin types in *A. alata*. Our opsin gene tree only included the known cnidarian opsins and thus does not address the question of cnidarian opsin monophyly. However, our opsin gene analysis recovered homologs within two of the three previously identified cnidarian opsin groups, namely group A and B (Feuda *et al.* 2012). Our analysis also recovered many *A. alata* opsin homologs within a large group that also contains cnidarian representatives of opsin group C, previously thought to be strictly anthozoan opsins. The presence of this opsin group in a clade with both anthozoans and medusozoans suggests that this sister group of the group B plus C clade was present in the cnidarian ancestor.

We found several opsin genes in *A. alata* were highly expressed in samples other than the rhopalium sample, and similar results have been reported for opsins in another cubozoan, *T.*

cystophora (Bielecki *et al.* 2014; Liegertová *et al.* 2015). Based on these findings we hypothesize that cubozoans have opsin-mediated extraocular photoreception activity possibly related to phototaxis, circadian rhythm or light-mediated spawning, such as has been demonstrated in other animals, including anthozoans (Levy *et al.* 2007; Mason *et al.* 2012; Rivera *et al.* 2012; Suga *et al.* 2008). These findings are also suggestive of extraocular photosensitivity (Arendt *et al.* 2009; Plachetzki *et al.* 2010, 2012a; Terakita 2005) that has a documented role in rhythmic behaviors and physiological processes in vertebrates and invertebrates, including nematocyst firing in cnidarians (Feuda *et al.* 2012, 2014; Liegertová *et al.* 2015; Plachetzki *et al.* 2010, 2012b; Porter *et al.* 2012). Such suggested extraocular photoreceptor cells may also comprise anatomically dispersed light sensitive neurons, in addition to ciliary or rhabdomeric morphotypes, possibly functioning in dispersed photoreception, also called the "dermal light sense" (for a review see (Desmond Ramirez *et al.* 2011)). Future characterization of the absorbance spectra for different opsin types in cubozoans, and visualization of the precise locality of expression using *in situ* hybridization, will help elucidate their potential functions in different medusa body parts and planulae.

In this study the expression of some of the components of the retinal photoisomerization pathway in all samples including planulae suggests that *A. alata* metabolizes the universal chromophore retinal (Mason *et al.* 2012; Shichida & Matsuyama 2009). However, transcripts for a putative blue-sensitive photoreceptor protein and circadian clock regulator cryptochrome homolog suggest an additional putative chromophore in *A. alata* that might function in non-rhopalium related blue-light mediated processes (e.g., phototaxis); such a function has previously been documented in other metazoans (Mason *et al.* 2012; Rivera *et al.* 2012). Determining the precise chromophore utilized by *A. alata* must await future functional studies.

Crystallins are multifunctional proteins often related to stress or metabolic enzymes that serve as important lens components controlling optical properties (Piatigorsky *et al.* 2001). The dual role crystallins play in eye lens as well as non-eye related tissues is known as "gene sharing" (Piatigorsky *et al.* 1988). We found that all three types of J-crystallins previously reported in *T. cystophora* (Piatigorsky *et al.* 2001) were present in *A. alata* and that these were typically most highly expressed in the rhopalium, with J2 crystallins showing more variable expression across samples.

We also identified transcripts corresponding to the developmental transcription factors Six and eyes absent (Eya), representing the first homologs of these genes identified from cubozoans. Genes in the Six-Eya homolog complex have known functions in eye development, including during embryogenesis and regeneration, in both non-bilaterians and bilaterians [79,84,116]. Six-Eya complex genes have been shown to act downstream of Pax genes [84], and PaxB expression has been reported in both adult and larval eyes of *T. cystophora*, where it is inferred to promote J-crystallin expression [28]. We also identified transcripts corresponding to the developmental transcription factors Six and eyes absent (Eya), representing the first homologs of these genes identified from cubozoans. Genes in the Six-Eya homolog complex have known functions in eye development, including during embryogenesis and regeneration, in both non-bilaterians and bilaterians [79,84,116]. Six-Eya complex genes have been shown to act downstream of Pax genes [84], and PaxB expression has been reported in both adult and larval eyes of *T. cystophora*, where it is inferred to promote J-crystallin expression [28]. Conversely, in the scyphozoan *Aurelia*, development of simple eyes is mediated by Six-Eya complex genes independent of PaxB expression [117]. Although the Trinotate report for the filtered *A. alata* transcriptome did not contain any transcripts annotated as PaxB, we identified two transcripts

(comp95018 and comp20156) annotated as other homeobox genes that appear to be putative PaxB homologs based on sequence identity (tBLASTx) with *Nematostella vectensis* PaxB mRNA. Whether eye development and Six-Eya expression in *A. alata* are dependent or independent of PaxB expression remain open questions. Future studies determining the spatial localization of gene expression during eye development in *A. alata* may be useful for further elucidating the gene regulatory networks functioning in eye development in cubozoan rhopalia and planulae eyes spots.

Sexual reproduction and embryogenesis

Cubozoan lifecycles alternate between an asexually reproducing sessile polyp stage and a sexually reproducing motile medusa stage. By profiling the transcripts from an adult body part abundant in developing oocytes (ovaries), our aim was to characterize the molecular components of oogenesis and early embryogenesis in *A. alata*. Additionally, because we found that sperm are internalized and interact with newly ovulated eggs within the gastrovascular cavity of *A. alata* females, our ovaries tissue sample also provided the opportunity to identify genes that might be involved in fertilization.

We identified several apparent homologs of Vitellogenin and Apolipoprotein, which have documented roles in oogenesis and embryogenesis, (Hayward *et al.* 2010; Iyengar *et al.* 1983; Levitan *et al.* 2015; Shikina *et al.* 2013) and found these to be most up-regulated in the ovaries of *A. alata*. Vitellogenin is an animal egg yolk protein that is synthesized in somatic cell lineages and subsequently incorporated into developing oocytes (by receptor mediated endocytosis), eventually serving as a nutrition source during embryogenesis (Hayward *et al.* 2010; Levitan *et al.* 2015; Shikina *et al.* 2013). In medusozoans little is known about the characteristics of

Vitellogenins as they have only been documented as egg yolk proteins in two coral species (Hayakawa *et al.* 2007; Shikina *et al.* 2013) and the model sea anemone *Nematostella vectensis* (Levitan *et al.* 2015). Vitellogenin proteins are expressed in both ovarian (or putative ovaries in anthozoans, e.g., (Levitan *et al.* 2015)) and extra-ovarian somatic cells, consistent with their important roles in processing large lipoproteins in a broad range of complex biological processes among metazoans (Hayward *et al.* 2010), including its disparate role as honey bee venom allergen (Danneels *et al.* 2015). Consistent with this, we found that in *A. alata* an apparent homologs of Vitellogenin-2 were expressed most highly in the ovaries, yet they and other Vitellogenins and Apolipoprotein-like homologs were detected in all medusa samples. In this study we also found a number of creatine kinase genes to be primarily up-regulated in the ovaries, but many were also detected (though at much lower expression levels) in all of our samples. Creatine kinases play an important role in oogenesis and early embryogenesis in mammals (Iyengar *et al.* 1983), having a broad enzymatic function in yielding ATP by catalyzing the reversible transfer of phosphate from creatine phosphate to ADP in cells with high activity (e.g., photoreceptors, primitive-type spermatozoa) (Pineda & Ellington 1999; Wallimann & Hemmer 1994).

We did not recover any genes characteristic of meiosis in the ovaries sample of *A. alata,* suggesting that the tissue was composed exclusively of mature ova at the time of sampling. It is also possible that the expression levels of putative meiosis transcripts were too low to be detected by our analyses, given our conservative transcriptome analysis protocol (see Methods). However, few studies exist that characterize the molecular aspects of sexual reproduction in cnidarians (Hayakawa *et al.* 2007; Shikina *et al.* 2013), limiting the number of potential gametogenic candidate genes targeted in this study. Future transcriptome and proteome profiling

studies of the gonads of *A. alata* and other cubozoans during medusa maturation are needed to shed light on the molecular underpinnings of the processes controlling gametogenesis in cubozoans.

In this study we also detected the expression of genes with putative roles in sperm flagella activation, pro-acrosomal vesicles and sperm capacitation, with many of these being most up-regulated in the ovaries sample of *A. alata*. These morphological and biochemical changes to the sperm are necessary for the sperm to reach and fertilize an oocyte, and their occurrence has been documented within the female reproductive tract in many animals (Rahman *et al.* 2014). Sperm capacitation was previously thought to occur exclusively in mammals, but more recently it has been documented in several invertebrates (Aungsuchawan *et al.* 2011). Our study is the first to suggest that sperm capacitation might occur within the gastrovascular cavity (putative female reproductive tract) of a cnidarian. We note that although sperm storage structures have been reported in a single family of cubozoans (Tripedaliidae) (Lewis & Long 2005; Straehler-Pohl *et al.* 2014), we do not expect sperm storage to occur in *A. alata*. Morphological observations during the course of this study as well as previous studies in *A. alata* have identified no structure(s) with a putative role in sperm storage in either male or female medusae (Gershwin 2005; Lewis *et al.* 2013; Mayer 1906). Future histological studies of *A. alata* medusae undergoing internal fertilization should elucidate the ultrastructure of the female reproductive tract and provide further insight into fertilization dynamics in this species.

Based on our observations in this and a prior study (Lewis *et al.* 2013), monthly spermcasting aggregations of *A. alata* medusae consist entirely of males and females with mature gonad morphology. We therefore hypothesize that gonad development occurs offshore in response to environmental and molecular cues related to the lunar cycle that may instigate

inshore migrations. During these monthly nearshore aggregations, which span three to four consecutive days, both sexes exhaust their entire gamete reserves in a process known as "controlled gonad rupture" (Lewis *et al.* 2013; Miller 1983). Male gonads completely disintegrate over the course of just several hours, and females simultaneously ingest massive quantities of sperm for internal fertilization. The interaction of sperm and eggs witnessed in the gastrovascular cavity, followed by release of blastulae into the surrounding water by females within hours, along with the abundance of sperm and fertilization-related transcripts detected in the ovaries sample, corroborate previous observations (Lewis *et al.* 2013) that fertilization occurs immediately following sperm ingestion and ovulation, adjacent to the ovaries within the gastrovascular cavity. Future molecular studies characterizing expression in sperm and eggs prior to and during fertilization will provide further insight into the dynamics of fertilization in cubozoans.

Conclusions

Whereas most cubozoans are difficult to study in their natural settings, *Alatina alata* is becoming a useful model for evolutionary and molecular studies because mature adults can be found predictably in near-shore waters. In this study, we generated a new genomic resource for *A. alata*, a transcriptome of multiple adult tissues and planulae, and characterized patterns of expression of transcripts across several body parts of a female medusa and larval planulae. We identified a large suite of candidate genes implicated in predation and defense, vision and the phototransduction pathway, and sexual reproduction and embryogenesis. This new genomic resource and the candidate genes we have identified will be valuable for further investigating the evolution of distinctive features of cubozoans, and the evolution of cnidarians more broadly.

Additional Files

Additional 3.I Transcriptome statistics. Summary table **(.xls)** of Trinity gene and transcript length distribution in whole transcriptome and filtered (fpkm=1.5) transcriptome. Summary table **(.xls)** of EdgeR and Trinotate results.

Additional 3.II Trinotate report of candidate genes. A filtered Trinotate annotation report **(.xls)** corresponding to annotations for all 651 candidate genes investigated for their putative role in venom, vision and sex (and early development) in the *A. alata* transcriptome.

Additional 3.III Trinotate report of Cnidaria genes. A filtered Trinotate annotation report **(.xls)** corresponding to *A. alata* transcripts whose top BLASTX/BLASTP hits corresponded to cnidarian genes/ proteins; summary table and pie chart included.

Additional 3.IV Heatmap and transcript DE matrix for five samples. Heatmap and corresponding matrix file **(.xls)** used in hierarchical clustering (EdgeR) of ~10K transcripts differentially expressed across *A. alata* medusa (gastric cirri, ovaries, tentacle, rhopalium) and planulae samples (columns) (of the identified ~32K Trinity transcripts).

Additional 3.V Heatmap and gene DE matrix for four medusa samples. Corresponds to Fig 3a-k in the text. Matrix file **(.xls)** corresponding to the heatmap showing hierarchical clustering (EdgeR) of the 2,916 genes differentially expressed across *A. alata* medusa samples (gastric cirri, ovaries, tentacle, rhopalium) (of the identified ~20K Trinity genes).

Additional 3.VI Top 50 DE genes of medusa samples. Corresponds to Fig 3.4b (Venn) in the text. Spreadsheet **(.xls)** corresponding to the Venn diagram of the top fifty most highly expressed genes across the four medusa samples (from matrix in Additional 3.V). Core genes are identified for each sample (gastric cirri, ovaries, tentacle and rhopalium). Detailed statistics (fpkm, counts, DE values for top 50 ranked genes by sample) summarized in histograms in Fig 3.5a-d. Annotations are provided from the Trinotate report.

Additional 3.VII Venom heatmap. Corresponds to Fig 3.6 a-k in the text. Matrix file **(.xls)** corresponding to matrix used for hierarchical clustering to generate the heatmap (EdgeR) of the 450 genes putative genes implicated in venom in *A. alata* medusa (gastric cirri, ovaries, tentacle and rhopalium) and planulae samples. Annotations are provided from the Trinotate report.

Additional 3.VIII Vision heatmap. Corresponds to Fig 3.7 a-k in the text. Matrix file **(.xls)** corresponding to matrix used for hierarchical clustering to generate the heatmap (EdgeR) of the

97 genes putative genes implicated in vision in *A. alata* medusa (gastric cirri, ovaries, tentacle and rhopalium) and planulae samples. Annotations are provided from the Trinotate report.

Additional 3.IX Sex heatmap. Corresponds to Fig 3.8 a-k in the text. Matrix file **(.xls)** corresponding to matrix used for hierarchical clustering to generate the heatmap (EdgeR) of the 104 genes putative genes implicated in sex in *A. alata* medusa (gastric cirri, ovaries, tentacle and rhopalium) and planulae samples. Annotations are provided from the Trinotate report.

Additional 3.X Top 50 DE genes of planulae sample. Spreadsheet **(.xls)** corresponding to the top fifty most highly expressed transcripts in the planulae sample of those differentially expressed (DE) across all five *A. alata* samples (gastric cirri, ovaries, tentacle, rhopalium, planulae) (from matrix in Additional 3.IV). Annotations are provided from the Trinotate report.

Acknowledgements

I acknowledge all coauthors on this paper (AG Collins, AE Bely, JF Ryan, P Cartwright). We thank A. van Dorsten, J. van Blerk, A. Lin, R. Peachy and CIEE Bonaire staff for collection and lab assistance; S. Pirro, M. Shcheglovitova, M. Falconer, D. Brinkman, and B. Bentlage for assistance related to the study; Smithsonian Biorepository staff and Museum Support Center collections staff; KUMC Genome Core staff; Trinityrnaseq-users forum and B. Haas; Smithsonian Institution High Performance Computing Cluster and Laboratories of Analytical Biology staff: DJ Dajiang, M. Kweskin, V. Gonzales, and P. Frandsen for bioinformatics support. We express our gratitude to two anonymous reviewers whose comments and suggestions helped improve the manuscript. Funding for field work (CLA) was provided by a University of Maryland Biological Sciences Eugenie Clark Scholarship and Smithsonian Peter Buck Predoctoral research grant; and for RNA-Seq by Paulyn Cartwright through NSF grant DEB–095357. JFR was supported by startup funds from the University of Florida DSP Research Strategic Initiatives #00114464 and University of Florida Office of the Provost Programs. AGC acknowledges the Mary & Robert Pew Public Education Fund, which supported the capture of some of the imagery in Figure 1.

Chapter 4: Evidence for an alternative mechanism of toxin production in the box jellyfish *Alatina alata*

Abstract

Cubozoans (box jellyfish) are some of the most venomous marine animals. Like all cnidarians, the sting of a box jellyfish originates from structures called nematocysts produced from a post-Golgi vesicle. When tentacles come in contact with prey or would-be predators, numerous nematocysts deploy a cocktail of toxins at a rapid speed via a long spiny tubule immobilizing the target organism. The implication has long been that toxin peptides and proteins making up the venom within the nematocyst capsule are secreted directly by nematocytes during nematocyst production (nematogenesis). However, we challenge the idea that in cnidarians venom is only made within nematocytes, presenting molecular and morphological data suggesting a dual role certain toxic-like enzymes might have internally (digestion of prey) as well as externally (envenomation) in cubozoans. Herein we provide a review of cubozoan prey capture and digestion informed by the scientific literature, and in an effort to uncover evidence for a central area enriched in gland cells within the stomach associated with the gastric cirri we provide a comparative description of the morphology of the digestive tract of *Alatina alata* and *Carybdea* box jellyfish species. Finally, we also conduct a rigorous comparative analysis of the gene ontology (GO) of venom implicated homologs in the tentacles and gastric cirri with a particular emphasis on variation in abundance of zinc metalloprotease homologs and other venom associated genes comprising the *A. alata* transcriptome.

Keywords: Venom, digestion, *Carybdea*, prey capture, sting, defense

Introduction

Box jellyfish (Cnidaria: Cubozoa) are among the most venomous marine animals in the world (Barnes 1966; Bentlage *et al.* 2010; Gershwin *et al.* 2010). The box jellyfish sing originates from cellular structures called nematocysts (Fig 4.1a), which are produced by specialized cells called nematocytes through a secretory pathway (Ozbek 2011). These microscopic venom-filled capsules are highly concentrated in the tentacles and are vital for prey capture and defense (Gershwin 2006; Hessinger & Lenhoff 1988). Tissue specific transcriptome analyses of the venomous box jellyfish *Alatina alata* revealed that candidate venom and nematocyst development (nematogenesis) genes were upregulated in the tentacles, while additional venom implicated genes were also highly expressed in the gastric cirri (Lewis Ames et al. In Press). Box jellyfish digestion is presumed to be facilitated by gastric cirri which in most cubozoans are coalesced into four distinct structures (i.e. gastric phacellae) lining the edge of the quadrangular stomach (Bentlage & Lewis 2012; Conant 1897, 1898; Gerswin 2005; Larson 1976). Nematocysts in the gastric cirri are thought to aid in incapacitation and digestion of prey items (Larson 1976). However, the emerging cubozoan model *Alatina alata*, lacks nematocysts in the gastric cirri (Lewis Ames et al. In Press; Lewis *et al.* 2013) suggesting that an alternative secretory mechanism may be responsible for venom produced for prey incapacitation and possibly digestion in this species. In Cnidaria it has been proposed that nematocysts synthesize and store toxic peptides prior to envenomation (Beckmann & Özbek 2012), however, nematogenesis is not the only pathway in which venom is produced (Moran *et al.* 2012, 2013). Additional examination of the gastric cirri ultrastructure and digestion activity in cubozoans will aid in determining the prevalence and potential role of glandular cells in venom secretion.

Many enzyme protein families have been recruited to aid in prey capture and digestion in metazoans, ultimately as secretions of the venom gland during envenomation (Bottrall *et al.* 2010; Calvete *et al.* 2009; Casewell *et al.* 2012; Fry *et al.* 2009; Junqueira-de-Azevedo *et al.* 2015; Mali *et al.* 2004; Nevalainen *et al.* 2004). Due to the potential dual role of these digestive enzymes as toxic-like enzyme components of venom, demarcation between envenomation and digestion is often attributed to alternative origins of venom production linked to a single point of envenomation (Bottrall *et al.* 2010; Junqueira-de-Azevedo *et al.* 2015). In Cnidaria, discernment between venom and non-venom digestive peptides is difficult as no central gland-like structure has been characterized that might serve in either venom or digestion enzyme production, or in both. More stringent morphological and molecular analyses are needed to test hypotheses about the dual role certain toxic-like enzymes might have internally (digestion of prey) as well as externally (envenomation). To better understand venom diversity and mechanisms controlling venom synthesis in *A. alata* we describe the morphology of the gastric cirri and digestion associated structures, review prey capture and digestion in cubozoans, analyze the publically available *A. alata* transcriptome for potentially overlooked venom-like transcripts, and conduct a rigorous comparative analysis of the gene ontology (GO) of venom implicated homologs in the tentacles and gastric cirri for differentially expressed toxin-gene candidates that were not evaluated previously.

Materials and Methods

Morphological data

The gross morphology, life history, sexual behavior, and worldwide distribution of *A. alata* are well-documented (Lawley et al. In Press; Lewis Ames et al. In Press; Arneson and Cutress 1976; Lewis et al. 2013; Chiaverano et al. 2013; Carrette et al. 2014). In this study,

Alatina alata material was collected in Bonaire, The Netherlands during monthly spawning aggregation (8 – 10 days after the full moon) in 2011, 2015 and 2016. No prey items were observed within the stomach or adhered to the tentacle at any time during this study. Photographs and video were taken using light microscopy of the gastric phacella (composed of cirri) excised from a live *A. alata* box jellyfish (bell height ~70 mm) (Fig 4.2; Supplemental File 4.I). Additionally, the gastric phacella was excised from a preserved (8% formalin) *A. alata* museum voucher in the Smithsonian National Museum of Natural History (USNM1195804), homogenized tissue was examined using light microscopy (600x and 1000x magnification) and photographed (Fig 4.2c, e).

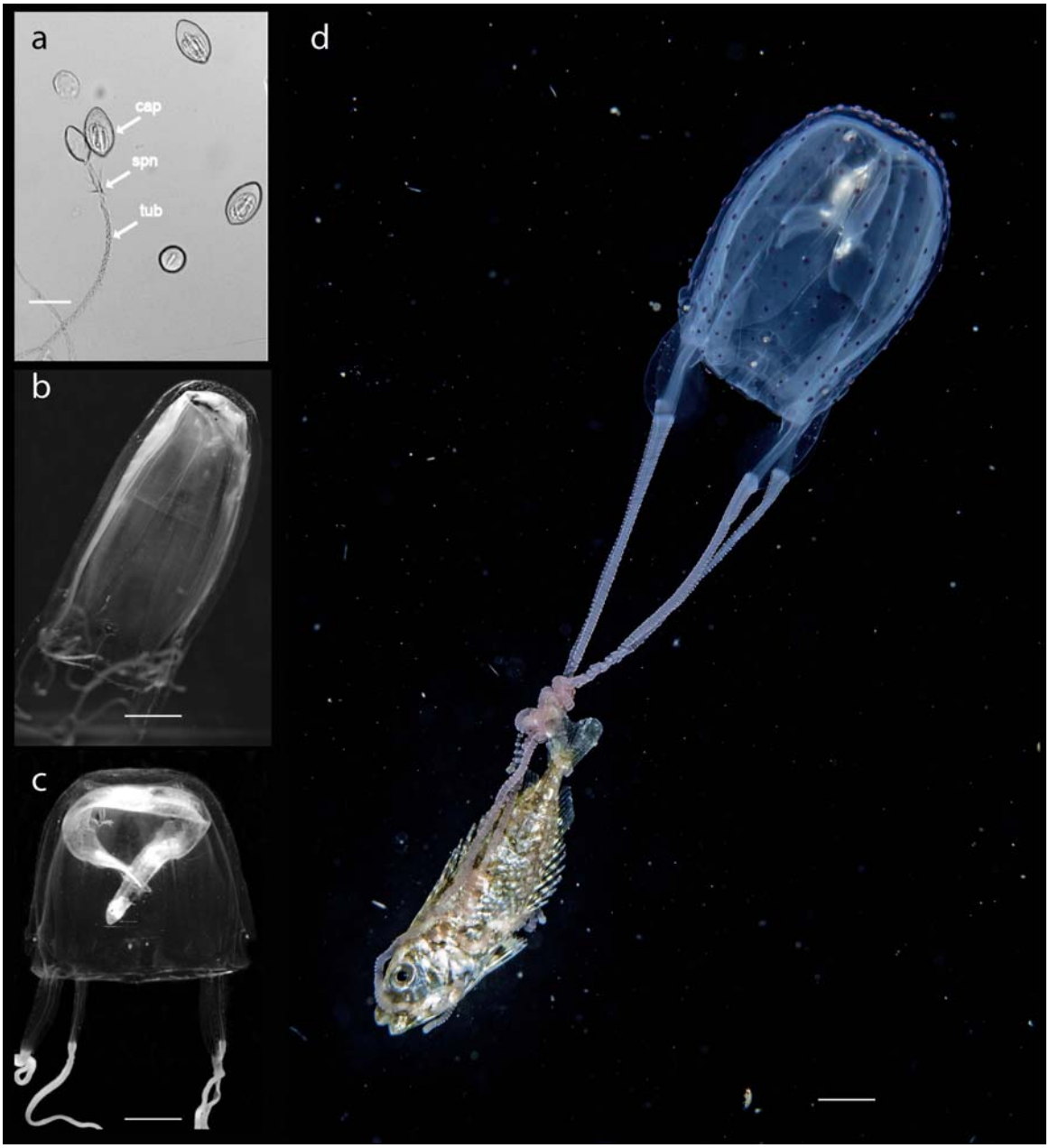

Figure 4.1 a-d Feeding in cubozoans. a. Penetrant nematocysts (euryteles) isolated from the tentacles of *Alatina alata* in this study. Four intact nematocysts seen on the right and a single discharged nematocyst seen on the left revealing the long spiny penetrant tubule. Bonaire, the Netherlands. b. *A. alata* feeding on a larval fish corresponding to the small thin dark black object lying horizontal in the stomach. Puerto Rico. Photo courtesy of R. N. Larson. c. *Carybdea xamachana* feeding on a larval eel almost twice the bell width, folded up inside the stomach with head and tail pushed down within the central feeding tube (manubrium). Puerto Rico. Photo courtesy of R. N. Larson. d. *Malo filipina* seconds after having captured and killed a young Rabbitfish. Aurora, The Philippines. Photo courtesy of S. Tuason. Abbreviations cap=capsule; spn=spines; tub=tubule. Scale bar a=30 µm, b-d=10mm.

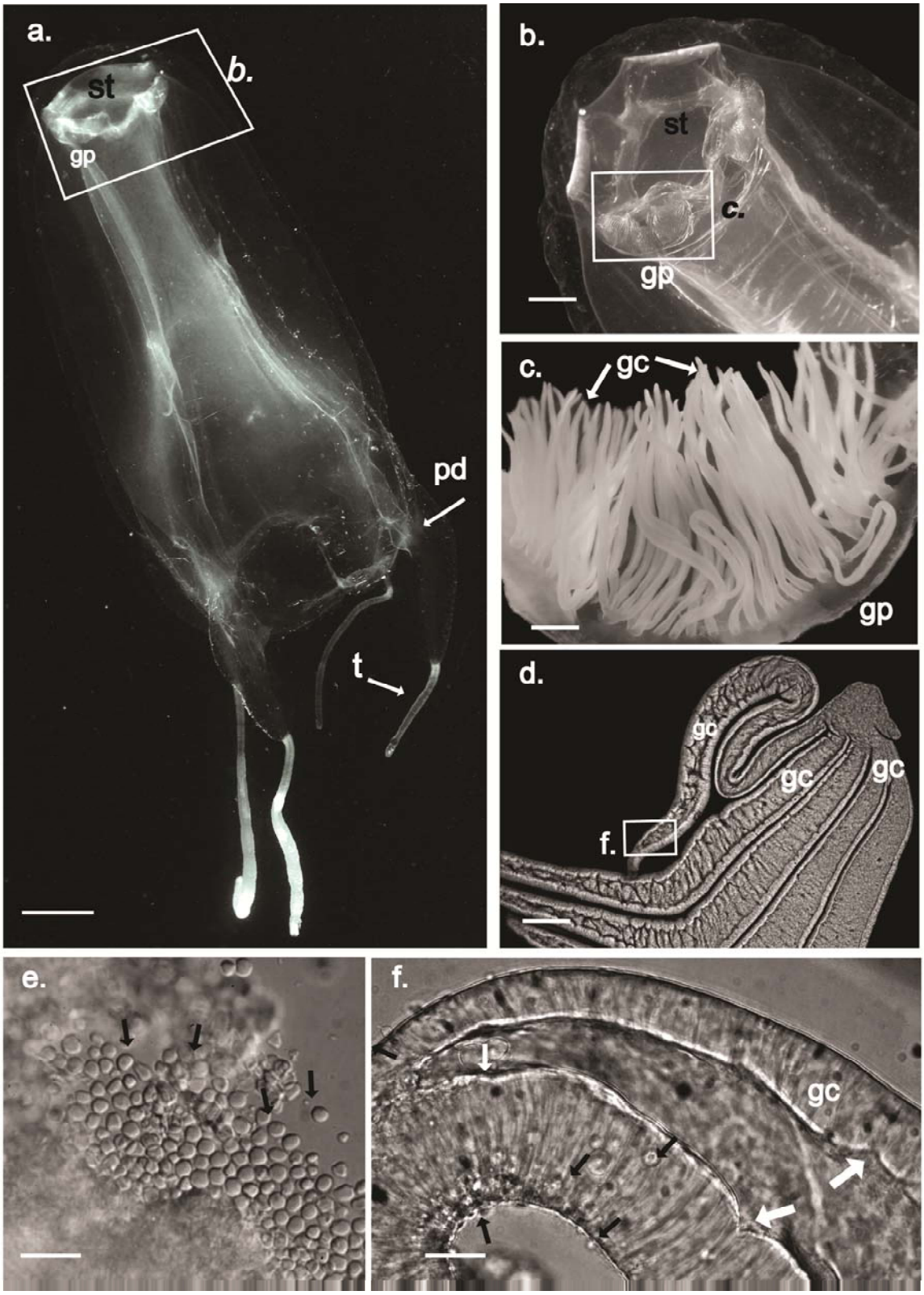

Figure 4.2 a-e. *Alatina alata* box jellyfish from Kralendijk, Bonaire, The Netherlands (October 29, 2016; 12:00).

a. Live mature *A. alata* box jellyfish swimming in the water nearshore, during a monthly spermcasting aggregation. The white box identifies the area on the transparent apex where the four gastric phacellae can be seen in each corner of the stomach. b. A magnification of the area highlighted in 1a of the *A. alata* apex. The white box highlights one of the four gastric phacellae found at each corner of the stomach. c. A single gastric phacella removed from the corner of the stomach of a preserved *A. alata* specimen (USNM 1195804). d. A tuft from a single gastric phacella removed from the corner of the stomach of a live *A. alata* box jellyfish. A membrane outlines each of the five individual gastric cirri forming the tuft. The appearance of horizontal ridges within the lumen corresponds to the somewhat regular pattern of constrictions along the length that connect with channels extending through the outer membrane to the stomach. e. Black arrows point to the remains of a macerated gastric phacella composed almost entirely of granules excised from the stomach of a preserved (8% formalin) *A. alata* medusa. f. A magnification of the area highlighted in 1d. of a single gastric cirrus. White arrows indicate areas where the lumen extends into the outer membrane of the gastric cirrus. Black arrows indicate tiny non-nucleated granules circulating within the lumen and transported across the outer membrane via channels (putative duct) into the stomach. Abbreviations gc=gastric cirrus, gp=gastric phacella, pd=pedalium, st=stomach, t=tentacle. Scale bars: a, 10 mm; b. 3 mm; c. 1 mm; d 0.5 mm; e. 130 µm; f. 0.10 mm. Fig 4.2 a, b, d, f courtesy of Allen G. Collins.

Molecular data

Sequencing, library preparation and *A. alata* transcriptome assembly were reported previously (Lewis Ames et al. In Press). Briefly, total RNA sequenced from multiple body parts (including tentacle and gastric cirri) was prepared using the paired end 100bp sequencing kit for the Illumina HiSeq 2500, trimmed using TrimGlaore! (Krueger 2012) and filtered using Allpaths Error Correction (Gnerre *et al.* 2011), and assembled *de novo* using Trinity (Grabherr *et al.* 2011; Haas *et al.* 2013). The resulting ~126K Trinity transcripts (CEGEMA completeness score = 99%) were further reduced to subset of ~32K transcripts at a 1.5 fpkm (fragments per kilo base per million fragments mapped) threshold identified in the *A. alata* transcriptome, publically available at DDBJ/ENA/GenBank GEUJ01000000.

Evaluation of de novo assembly

Candidate toxin genes were screened to determine if multi-copy isoforms from candidate toxins were incorrectly assembled into a single transcript (see Macrander et al. 2015) during *de*

novo assembly of the *A. alata* transcriptome (Lewis Ames et al. In Press). The original paired-end raw reads (BioProject PRJNA312373; BioSample SAMN4569893 and SAMN4569895) were mapped to candidate toxin transcripts in bowtie (v. 1.0.0) with default settings, except for trimming the first 5 bases from the raw reads (Langmead *et al.* 2009). The candidate toxin transcripts and mapped reads were then imported into Geneious version (Kearse *et al.* 2012) and visually inspected for polymorphisms that may be indicative of multiple isoforms. If only a single polymorphism was recovered consistently among the mapped reads it was treated as a potentially heterozygous allele. If more than one polymorphism was recovered, a *de novo* assembly for those transcripts was done using the Geneious assembler with 20% gaps per read, 5 max. gap size, 75 min. overlap, word length 12, a reanalyze threshold of 4, and maximum ambiguity of 16. The *de novo* assembled contigs were then reexamined for polymorphisms in the assembly.

RSEM and differential expression (EdgeR)

Relative gene expression levels were calculated for the gastric cirri and tentacle samples using the program RSEM (Li and Dewey 2011) by mapping raw reads from the gastric cirri and tentacle samples to the *A. alata* transcriptome (Lewis Ames et al. In Press). Differential expression (DE) analysis was determined as differences in gene abundances between the gastric cirri and tentacle samples using the EdgeR Bioconductor package (Robinson et al. 2010). The transcripts were then used as query sequences in a tBLASTn search against the UniProt database. Blast hits with gene ontology (GO) information were used in combination with a custom python script (Supplemental File 4.II) to average DE expression values derived from the matrix across all GO groups. This allowed us to control for overly abundant GO groups, while still considering

differences in gene expression in our gene ontology assignments. Upregulated GO groups and calculated GO specific expression values were visualized in the program REVIGO (Supek et al. 2011) for the biological process and molecular function GO domains in the gastric cirri and tentacles.

Identification of unknown toxins

In addition to characterizing the abundance and diversity of proteins documented in nematogenesis and box jellyfish venom components (Supplemental File 4.III), we identified additional uncharacterized toxins or toxic-like transcripts which contained signaling region and shared high sequence similarity to functionally characterized toxins from other taxa. Transcripts were translated into protein sequences in the program (http://transdecoder.github.io) TransDecoder, with open reading frames (ORFs) of 50 or more amino acids in length. The program SignalP (Petersen *et al.* 2011) was used to identify transcripts containing a signaling region. Signal peptides have been identified in transcripts encoding peptides secreted within nematocytes as structural elements of nematocysts (i.e., complex secretory organelles) (David *et al.* 2008; Zenkert *et al.* 2011). Here we also sought to identify signal peptides in the gastric cirri which lack nematocysts, given the predicted presence of putative venom gland cells associated with the gastric cirri (see Lewis Ames et al. In Press). Toxic-like sequences were identified in tBLASTn using annotated toxin peptides against the ToxProt data (http://www.uniprot.org/program/Toxins). To test the homology of the *A. alata* candidate zinc metalloprotease transcripts, amino acid sequences (as estimated ORFs) were aligned using MAFFT (Katoh & Standley 2013) using the L-INS-i algorithm, BLOSUM62 scoring matrix, 1.53 gap open penalty, and 0.123 offset value, alongside zinc metalloproteases from several

venomous and non-venomous animal species (obtained from NCBI Genbank). A zinc metalloprotease gene tree was constructed using the program FastTree (Price *et al.* 2010) using default settings and 1000 bootstrap replicates using seqboot in the PHYLIP package (Felsenstein 2009) (Supplemental Figure 4.I).

Results and Discussion

Prey capture and incapacitation in cubozoans

Few accounts of cubozoan prey capture have been documented in natural settings, but from the limited available accounts it is evident that box jellyfish feed on a variety of prey items (Table 4.1). An ontogenetic shift in diet has been shown to be correlated with modifications in nematocyst composition in some cubozoans (Carrette *et al.* 2002; Júnior & Haddad 2008), suggesting that different nematocyst types may be required to capture and incapacitate bigger prey items in adults (Acevedo *et al.* 2013; McClounan & Seymour 2012). Free amino acids being released by the potential prey item may trigger a response by the tentacles resulting in prey capture and immobilization (Larson 1976). When tentacles encounter potential prey item nematocysts are released ejecting the harpoon-like tubule (Fig 4.1a) into the target organism (Beckmann & Özbek 2012; David & Challoner 1974; David *et al.* 2008; Özbek *et al.* 2009). As nematocysts penetrate the epidermis, bioactive proteins are injected, many with non-specific binding properties which disrupt homeostasis in the potential prey (Badré 2014). Once the prey is attached to the tentacle, the tentacle contracts bringing it towards the bell margin (Fig 4.1d) and into the bell cavity where, in a coordinated manner, prey is inserted into the central prehensile manubrium (feeding tube) (Larson 1976; Stewart 1996) (Fig 4.3), and transported via

ciliary action and muscle contraction into the transparent stomach (Fig 4.3) located at the apex of the box jellyfish (Fig 4.1b, c).

Depending on the box jellyfish species, prey items coming in contact with tentacles are either immobilized upon contact, but remain writhing until they are inserted into the manubrium, or almost immediately paralyzed, becoming lifeless, after being trapped by one or more tentacles. Fig 4.1b-d compares three species of box jellyfish varying in size from ~30 mm *Malo filipina* (Fig 4.1d) and *Carybdea xamachana*. (Fig 4.1c) to ~100 mm *A. alata* (Fig 4.1b), with the two smaller species catching and ingesting prey items proportionately much larger than the bell size in comparison to the bigger species. Although no explanation exist for this discrepancy between prey sizes, we suspect that both the strength of the tentacle and the type of nematocysts within the tentacles are key factors dictating size limitations in prey capture among species. Cubozoan tentacles are composed of both circular and longitudinal muscle fibers arranged around a longitudinal canal that runs the length of the tentacle lumen made up of rings of nematocysts (Southcott 1967). Some box jellyfish species also have nematocysts along the mouth lips (Gershwin 2006), which are thought to aid in securing large prey in the manubrium (Larson 1976) (Fig 4.1c). Additionally, the holding capacity (dimensions) of the transparent stomach cavity varies drastically between cubozoan species, with some species have a deep broad stomach, while others, such as *A. alata* have a shallow stomach (~15 mm wide) consisting of a short (~5 mm long) wide manubrium (feeding tube), that opens up to almost the same width as the stomach (Bentlage & Lewis 2012) (Fig 4.2b). Compared to the deeper stomach and narrower manubrium of *Carybdea* (Fig 4.3), it appears that the broad shallow manubrium of *A. alata* is not conducive to supporting large prey items such as those ingested by *Carybdea*.

Human envenomation

Nematocyst type likely plays a role in sting potency during human envenomation. Many local and systematic symptoms manifested in victims of venomous box jellyfish stings includes severe back and muscle pain, difficulty breathing, hypertension and even death (for a review of symptoms of box jellyfish stings see (Brinkman & Burnell 2009)). A study showed that most nematocyst tubules from the tentacles of "harmful" box jellyfish were longer than 200 μm in length after being discharged from the capsule (Kitatani *et al.* 2015). This suggests that longer nematocyst tubules are important in capturing larger prey, or prey with a thicker (or hard) epidermis, and therefore human envenomation by a longer tubule might be more severe because of the greater potential for direct intravenous venom injection.

The complexity and severity of symptoms resulting from envenomation by the tentacles of *A. alata* (Chung *et al.* 2001; Yoshimoto & Yanagihara 2002) suggests that other biological factors are involved that enhance nematocyst deployment methods or augment sting potency. When considering potential sources of indirect envenomation it is important to consider the role of the clusters of nematocyst warts covering the exterior of the bell of cubozoans species (Bentlage & Lewis 2012). The exterior bell warts of *Carukia* contain penetrant mastigophores, and have been documented in inflicting Irukandji syndrome as extensive as that caused by direct contact with the tentacles of that same species (Barnes 1964). Conversely, the bell warts of *A. alata* consist of non-penetrant isorhizas (Lewis *et al.* 2013; Mariscal 1974), and therefore contact with the bell does not typically pose a risk of envenomation. However, it has been reported that *A. alata* releases an abundance of intact nematocysts into the surrounding seawater during massive spawning aggregations (Lewis et al. In Review); this species (as *C. alata*) has also been documented releasing profuse amounts of mucous containing unfired nematocysts when

inadvertently handled (Graham 1998). It is not clear how representative this phenomenon of "nematocyst ejection" is among cubozoans, but in the case of *A. alata* the presence of intact nematocysts in the water could represent a possible irritant causing the characteristic rash and welts presenting in *A. alata* sting victims (Burnett 2009; Figueroa Rosa 2015).

Anatomy of the stomach and gastrovascular system in Carybdea and Alatina alata

Our current understanding of the morphology and histology of the cubozoan digestive tract (manubrium, stomach and gastric phacellae) is based on studies of the internal anatomy of *Carybdea* species by Conant (1898), Ishida (1936), and Larson (1976), and on this study on *A. alata*. The digestive tract of *Carybdea* species (e.g., *C. xamachana*, *C. marsupialis*, *C. rastonii*) and *A. alata* share many morphological characters: At each corner of the stomach both species possess four interradial gastric phacellae composed of multiple tufts of gastric cirri attached to one (*Carybdea*) or multiple (*A. alata*) stalks, like the hairs of a brush, formed as outgrowths of the lower stomach wall (Fig 4.2b, c). The floor of the stomach of *C. xamachana* is composed mainly of densely crowded high columnar cells whose granular appearance is due to a high concentration of vacuoles (or granules); interspersed among the columnar cells are mucus secreting goblet cells (Conant). In the absence of histological studies for *A. alata*, we were unable to verify that the cellular composition of *A. alata* matches that of the stomach floor of *C. xamachana*. However, membrane bound vacuoles were observed within gastric cirri in live *A. alata* medusae (Fig 4.2f), which are contiguous with the floor of the stomach. The vesicles were seen moving towards the distal tips of each of the gastric cirri and being secreted into the stomach cavity via duct-like channels. In both *C. xamachana* and *A. alata* (and possibly all cubozoans) the highly contractile gastric cirri are covered in thick mucus, likely due to the stomach floor being composed of glandular cells and mucus-containing goblet cells. One major

difference between the anatomy of the digestive tract of *C. xamachana* and *A. alata* is the absence of nematocysts along the mouth lips of the manubrium or within the gastric cirri of *A. alata*.

Table 4.1 Review of cubozoan species prey items and geographical location.

Cubozoan Family	Cubozoan species	Prey item(s)	Geographical location	Reference(s)
Alatinidae	*Alatina alata*	plankton (?), larval fish, caridean shrimps, polychaetes	Puerto Rico, USA (Atlantic)	Arneson and Cutress 1976; Pers comm., R. N. Larson (Fig 2a); Lewis et al. 2013
Carukiidae	*Malo filipina*	Rabbitfish	Aurora, The Philippines (Pacific)	Pers comm., S. Tuason
Carukiidae	*Carukia barnesi*	plankton, larval fish (*Acanthochromis* sp.)	Double Island, North Queensland, Australia (Pacific)	Courtney et al. 2005
Tamoyidae	*Tamoya haplonema*	fish (*Teleostei*)	Shangrilá & Paraná, Brazil (Atlantic)	Junior & Hadadd 2008
Carybdeidae	*Carybdea xamachana* (as *C. marsupialis*)	polychaetes (*Ceratonereis*), crustaceans (*Acartia*), fish (*Jenkinsia*), and eel larvae	Puerto Rico, USA (Atlantic)	Larson 1976
Carybdeidae	*Carybdea marsupialis*	plankton	Alicante, Spain (Adriatic Sea)	Acvedo et al. 2013
Carybdeidae	*Carybdea brevipedalia* (as *C. rastonii*)	fish	Shizuoka, Japan (Pacific)	Ishida 1936
Carybdeidae	*Carybdea rastonii*	mysids, larval fish	Hawaii, USA (Pacific)	Matsumoto 1996
Tripedaliidae	*Tripedalia cystophora*	copepods (*Oithano nana*)	Puerto Rico, USA (Atlantic)	Stewart 1996
Tripedaliidae	*Copula sivickisi*	polychaetes heteronereids, gammarid amphipods, cumaceans	Magnetic Island, Townsville, Australia (Pacific)	Hartwick 1991
Chiropsalmidae	*Chiropsalmus quadrumanus*	crustaceans: sergestid shrimp (*Peisos petrunkevitchi*), peneoidean shrimps, brachyuran larvae, isopods, crabs; nematodes, fish, fish eggs	Shangrilá & Paraná, Brazil (Atlantic)	Júnior & Hadadd 2008
Chiropsalmidae	*Chiropsalmus sp.*	shrimp	Australia (Pacific)	Carrette et al. 2002
Chirodropidae	*Chironex fleckeri*	shrimp (*Actes australia*), fish	Australia (Pacific)	Barnes 1966; Hamner and Hamner 1994; Carrette et al. 2002

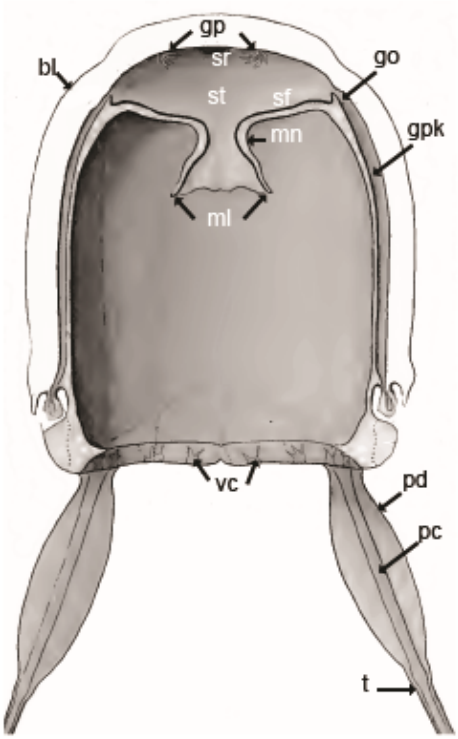

Figure 4.3 Schematic of the internal anatomy of the box jellyfish *Carybdea*.
A *Carybdea* bell cut in half longitudinally at the perradius (along the vertical axis at the mid-point between opposing tentacles). Dark shading indicates the gastrovascular system in which the fluids in the stomach are shuttled to the gastric pockets (gastro vascular cavity) via four gastric ostia (only two shown here) located on each side of the bell. A connection between the gastric pockets and the tentacles is facilitated via a circular channel around the perimeter of the base of the bell (not shown here) giving rise to four pedalial canals (only two shown here) and multiple marginal velarial canals (eight per quadrant). Abbreviations b=bell, go=gastric ostium, gp=gastric phacellae, gpk=gastric pocket, ml=mouth lips, mn=manubrium (feeding tube), pc=pedalial canal, pd=pedalium, sf=stomach floor, sr=stomach roof, st=stomach, t=tentacle, vc=velarial canals. *Carybdea xaymacana* box jellyfish depicted here ~30 mm in bell height. Modified from (Conant 1898).

Gastric cirri morphology

Morphological examination of the four gastric phacellae in the stomach of *A. alata* revealed that each gastric phacella is comprised of ~ 20 tufts, each of which originates as a single basal gelatin-like trunk connected proximally to the stomach floor by a central lumen, and diverges distally into multiple branches comprising up to 100 terminal gastric cirri all sharing a

central space within each gastric phacella (Fig 4.2a-f). Consistent with previous studies on *A. alata* morphology, no nematocysts were identified in the gastric cirri , and no prey items were associated with the digestion tract (Gershwin 2005; Lewis et al. 2013; Lewis Ames et al. In Press). When a single tuft comprising five gastric cirri was dissected from a gastric phacella in the stomach of a live *A. alata* medusa (Fig 4.2 b, d, f), all cirri continuously writhed and constricted independent of one another (Supplemental File 4.I). A careful examination of the gastric cirri tuft revealed a continuous membrane outlining the periphery of each gastric cirrus connecting them at the base (the stalk) and creating a conjoint shared space within the tuft (Fig 4.2d). Within the central space and each gastric cirrus numerous non-nucleated membrane bound granules (10 – 15 μm in diameter) were seen being transported from the base of the tuft up into individual gastric cirri towards the distal tips (Fig 4.2f). Horizontal striations were observed within the central axis of supportive "gelatin" in a somewhat regular pattern along the length of each cirrus corresponding to duct-like channels (Fig 4.2d, f). Granules were witnessed being transported through the duct-like channels across the membrane along the length of the gastric cirri and secreted into what would be the stomach in vivo. The contents of the granules were not examined in this study. However, macerated gastric cirri tissue dissected from a preserved *A. alata* museum specimen was found to be composed almost entirely of granules (Fig 4.2e) identical in size and morphology to those seen moving about within the gastric cirri of a live *A. alata* box jellyfish (Fig 4.2f) and match the description of granules located within the glandular cells of *C. xamachana* (Conant 1898).

Prey digestion

Within the stomach of *C. xaymacana*, prey are held in place by the gastric cirri, where nematocysts likely aid in prey stabilization and incapacitation (Larson 1976). Prey items within

the stomach are covered by mucus secreted by the gastric phacellae and stomach wall that exhibits extracellular protease activity (Larson 1976). Functional assays utilizing mucus recovered from the manubrium, gastric fluid (from the stomach cavity) and phacellae of *C. brevipedalia* (as *C. rastonii*) revealed that tryptase activity was elevated in the gastric phacellae but not in the rest of the digestive tract, i.e., manubrium and stomach wall, in the absence of any prey items (fish) (Ishida 1936). However, after prey items were introduced, tryptase activity became elevated in the gastric fluid from the stomach and within hours was equal to levels initially measured in the gastric phacellae (Ishida 1936). Tryptase detected in the gastric fluid and manubrium was presumed to be secreted from the gastric phacellae, and gradually filled the entire digestive tract to facilitate extracellular digestion (Ishida 1936). Digestion in *Carybdea* species, and cubozoans in general, has been described as a relatively rapid process, taking only a few hours (Acevedo *et al.* 2013; Hamner *et al.* 1995; Ishida 1936; Larson 1976). During digestion, food particles are seen circulating directionally through the transparent gastrovascular system of the box jellyfish (Larson 1976). Contant (1897) noted that "the digestive juices left the nervous system of the fish intact so that from the stomach of *C. xamachana* could be obtained beautiful dissections, or rather macerations, of the brain, cord, and lateral nerves of a small fish." Digestion was considered finished when debris (fish scales and eyes) was regurgitated from the manubrium, and *C. xamachana* ingested new prey items (Larson 1976).

Circulation of gastric fluids and nutrients

In *Carybdea* and *A. alata* the four gastric ostia open perradially at a point between adjacent gastric phacellae, on each of the four walls of the stomach (Fig 4.3). Each ostium is contiguous with one of the four gastric pockets lying peripherally within each of the four vertical sides of the bell (Fig 4.3), separated from one another by thin strips of interradial vascular

lamella. Gastric ostia serve as channels connecting the stomach to the peripheral part of the gastrovascular system. Movement of fluids (nutrients) into the gastrovascular system and of waste or reproductive material out is regulated by the flexible lower margin of each gastric ostium that forms a valve (Fig 4.3). Each gastric pocket extends the vertical length of the bell, where it branches into velarial canals at the base of the bell at the velarial turnover. A circular canal runs the perimeter of the bell near the base; interradially (at each of the four corners of the bell) the circular canal gives rise to each tentacle canal which passes through the center of the pedalium (the wide wing-like structure at the base of the tentacle), extending into the lumen of the tentacle which is hollow and encompassed by stacked rings of nematocysts extending the length of the tentacle. Within the tentacles lumen, nematocysts developing are in direct communication with the peripheral gastrovascular system by way of fluids circulating in and out of the tentacle canal.

Venom and digestive secretory products

In many venomous animals paralogous proteins with roles in venom and digestion are synthesized in separate organs (Junqueira-de-Azevedo *et al.* 2015). However, in box jellyfish, neither a venom gland nor a digestive gland has been documented. However, the hypothesized presence of a putative gland cells in the stomach of *A. alata*, as was speculated by Lewis Ames et al. (In Review), suggests that secretion of granules containing pre-packaged toxins and toxin-like enzymes synthesized and secreted centrally may provide an alternative mode for venom secretion via a concentration of gland cells in the stomach floor, for incapacitation and digestion of ingested prey. Furthermore, these putative pre-packed venom granules might also be transported to nematogenic regions in the tentacles, via the connection between the stomach and peripheral gastrovascular system, where they are possibly incorporated into developing

nematocysts. The putative secretory role played by the gastric cirri of *A. alata* and *C. xamachana* (Conant 1897, 1898; Larson 1976), and the secretory nature of the granules appears functionally similar to gland cells recently identified in a sea anemone that secretes toxins independent of nematocytes (Moran *et al.* 2012, 2013). If this putative central region of gland cell concentration is in fact a point of secretion of toxic-like enzymes and other toxins with roles in digestion and envenomation, this would represent the first documented gland-like structure with a dual role in digestion and venom in cnidarians. Testing this hypothesis in cubozoans must await additional studies documenting prey capture and feeding, histological examination of the ultrastructure of the digestive tract, and proteome characterization and candidate gene localization in body parts where venom synthesis has been documented.

Evaluation of de novo transcriptome assembly

In total we evaluated 154 candidate toxin transcripts to validate whether the *A. alata* transcriptome assembly might have inadvertently misincorporated certain raw reads as a single transcript that may in fact belong to multiple toxin gene copies but with low sequence variation. This method of identifying potential "missing transcripts" was demonstrated by (see Macrader et al. 2015). By mapping the raw reads back onto the transcripts we determined that the majority of the original transcripts had properly been recovered in the original transcriptome, and for 52 transcripts potential heterozygous alleles were recovered. Beyond potential heterozygotes, there were no additional "hidden" toxin gene copies or alleles recovered from the original *A. alata* transcriptome assembly. Given the stringent assembly parameters employed by the authors during de novo assembly of the *A. alata* transcriptome (Lewis Ames et al. In Press) beyond just the default settings, it appears that potentially problematic transcripts were appropriately

removed during *de novo* transcriptome assembly, resulting in a final product that serves as an excellent tool for an array of downstream transcriptome profiling studies.

Tissue specific expression of candidate toxins

Within the tentacles, the most highly upregulated transcripts identified from queried sequences were those with roles in tentacle structure (e.g., collagens) and documented roles in nematogenesis (e.g., minicollagens) (Fig 4.4) – the development of nematocysts within specialize nematocytes via post-Golgi secretions (Hessinger & Lenhoff 1988). Conversely, transcripts resembling venom associated candidate genes were expressed at much higher levels in the gastric cirri, with serine proteases, metalloproteases, and GM2 ganglioside collectively being expressed at in higher proportions than any other group of venom associated transcripts in the tentacles (Fig 4.4). Among the sequences queried in this analysis the most highly upregulated venom associated transcripts in both the tentacle and gastric cirri corresponded to homologs of the zinc metalloprotease family (Fig 4.4). Zinc metalloproteases make up a major component of venom in many species by activating toxins or acting as toxins (Calvete *et al.* 2009), and are reported to exhibit a dual role in initial prey immobilization during envenomation and in prey digestion (Bottrall *et al.* 2010).

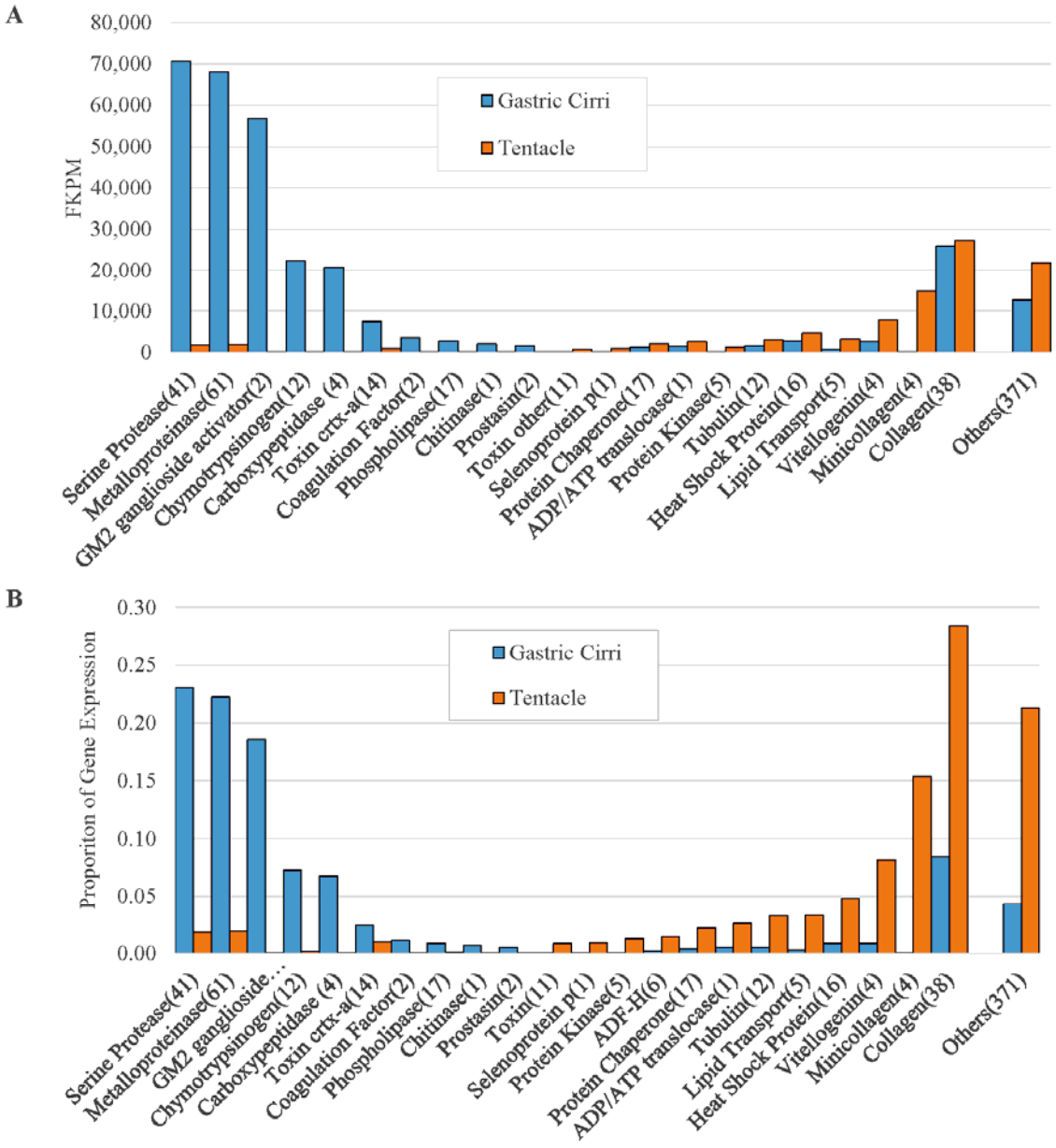

Figure 4.4 a,b Expression level (FPKM) comparisons of box jellyfish toxins and nematocysts associated proteins in the gastric cirri and tentacles.
Expression values for each group of genes are represented as either (a) FPKM or (b) percent of focal candidate gene expression. Numbers in parenthesis indicate how many transcripts were identified for each of the toxin or nematocyst associated proteins.

In box jellyfish zinc metalloproteases are reportedly deployed from nematocysts, in combination with other venom components, to disable homeostatic processes in the target organism (Casewell *et al.* 2012; Chera *et al.* 2006; Pan *et al.* 1998). Although zinc metalloproteases have been identified in the tentacles of box jellyfish (Lewis Ames et al. In Press; Avila Soria 2009; Brinkman et al. 2015) and other cnidarians (Balasubramanian *et al.* 2012; Gacesa *et al.* 2015; Ponce *et al.* 2015, 2016; Weston *et al.* 2013), the recruitment of these toxic-like enzymes as potential cnidarian venom components has not been evaluated in an evolutionary context. To elucidate evolutionary relationships among the genes in this large gene family we reconstructed a zinc metalloprotease gene tree including 44 candidate zinc metalloprotease transcripts from *A. alata*, and previously characterized zinc metalloproteases publically available for cnidarians and other metazoans (NCBI Genbank) for a total of 316 zinc metalloprotease sequences represented in the gene tree. The *A. alata* candidate zinc metalloprotease transcripts were found throughout the gene tree, with many individual gene clusters having high bootstrap support, but low support between gene clusters (Additional Figure 4.I). When evaluating individual transcript expression in combination with the gene tree analyses, we found that that there were no regions of the gene tree that corresponded to candidate zinc metalloprotease being consistently expressed at higher levels in either the tentacle or gastric cirri (Additional Figure 4.I). These results indicate that in a comparative context (gene tree reconstruction and gene expression levels) the zinc metalloproteases identified in our study do not appear to have an exclusive venom-related role in the tentacles of *A. alata*.

Gene ontologies (GO)

An EdgeR analysis in this study of the candidate venom and nematogenesis genes identified 1871 transcripts that were differentially expressed between *A. alata* gastric cirri and

tentacles. Of these 1260 had significant BLAST hits against the UNIPROT database, of which 1223 were functionally categorized in a variety of gene ontology groups. Among these, the most highly expressed unique gene ontology groups were visualized in the program REVIGO (Fig 4.5). Overall in our analysis we identified a significantly higher ration of GO groups in the gastric cirri than in the tentacle sample. Within the GO Biological Process (BP) domain, transcripts associated with digestion and proteolysis were in highest abundance. In the Molecular Function (MF) domain, transcripts involved with peptidase activity and binding activity (small molecules and ions) were the most abundant (Fig 4.5); these functional groups were consistently most highly expressed in the gastric cirri. Conversely, within the tentacle the GO groups recovered were expressed at substantially lower levels, with the most highly expressed transcripts in the BP and MF domain were associated with translation and structural molecules, respectively (Fig 4.5). A full list of GO groups and associated transcripts is provided in Supplemental File 4.III.

Additional toxin candidates

Previous transcriptome profiling of the tentacles of *A.alata* (Lewis Ames et al. In Press) did not conclusively identify likely bioactive candidates upregulated exclusively in the tentacles that might be responsible for causing the symptoms presented in *A. alata* sting victims (e.g., Figueroa Rosa 2015). In total in this study we identified 1787 toxin-like transcripts that shared sequence similarity to known toxin genes, but that had not previously been identified in the *A. alata* transcriptome. Of these identified putative toxin genes, 578 transcripts contained a signaling peptide region. Similarly to the initial differential gene expression analyses we conducted, our findings here also show that the more highly expressed candidate toxins are found within the gastric cirri rather than the tentacle (Table 4.2a). Using the program SignalP we

identified 3517 transcripts containing a signaling region, many of which did not resemble known toxin genes. However, like our previous findings, these results showed that transcripts containing a signaling region were consistently more abundant in the gastric cirri than in the tentacle (Table 4.2b).

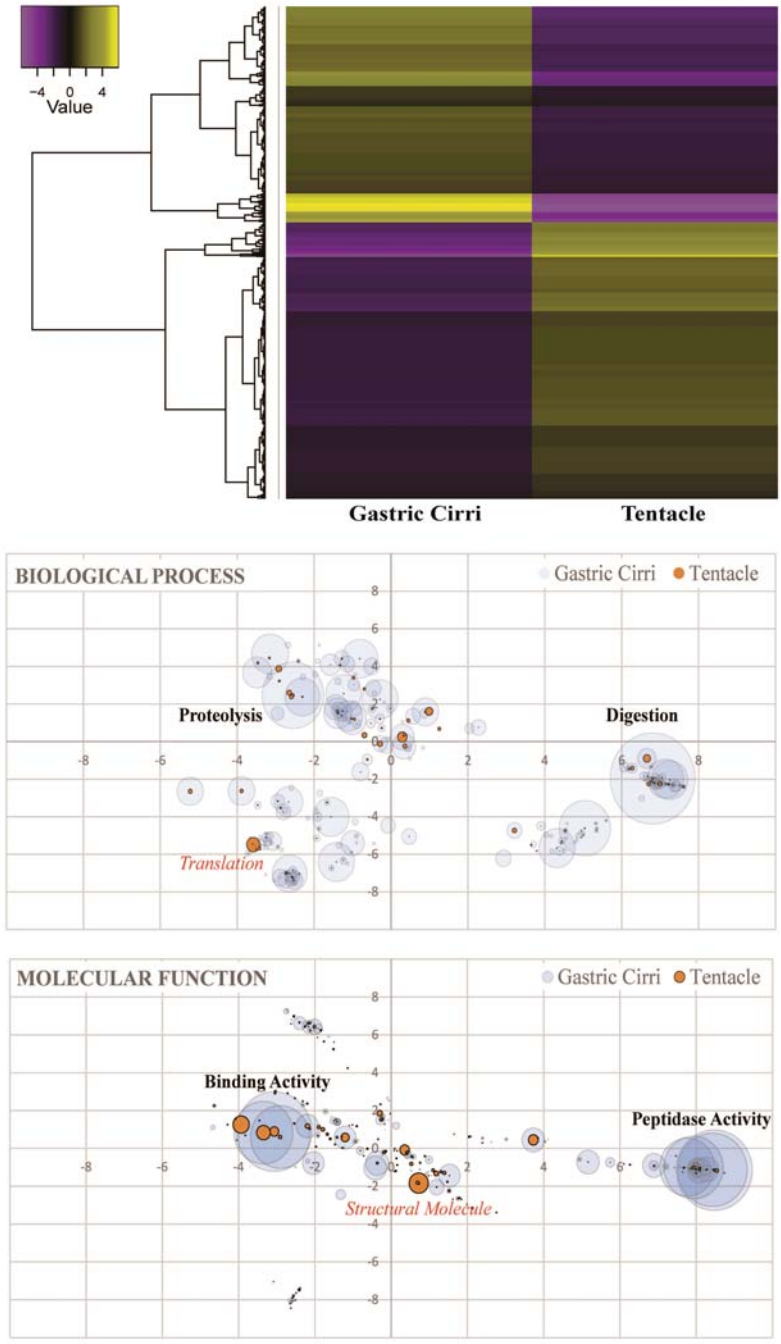

Figure 4.5 Differential gene expression and gene ontology (GO).
(a) Heat map for gastric cirri (light blue) and tentacles (orange) in *Alatina alata* with REVIGO plots depicting overall expression of identified gene ontology (GO) groups in the (b) biological process and (c) molecular function domains. b) Within the biological process domain GO associated with proteolysis and digestion were expressed at significantly higher levels in the gastric cirri, while the most notable GO group from the tentacles were associated with

translation. C) In the molecular function domain GO groups closely associated with peptidase and binding activity were expressed at much higher levels, while GO groups associated with structural molecules were the most abundant in the tentacles.

Table 4.2 a, b. Ten most highly expressed candidate toxic-like sequences and transcripts with signaling region for gastric cirri (GC FPKM) and tentacle (T FPKM). Duplicate hits within each search are noted in bold. * indicates transcript identified using both approaches.

(a) ToxProt Top BLAST hits

Transcript	GC FPKM	E-value	ID	Protein names	Organism
comp63246_c0_seq1*	65796.36	5.7	Q07932	Pilosulin-1	Myrmecia pilosula (Jack jumper ant)
comp66861_c0_seq1*	15029.07	9.00E-06	Q7LZF5	Thrombin-like enzyme catroxobin-1	Crotalus atrox (Western diamondback rattlesnake)
comp63353_c0_seq1*	13650.81	2.2	P0DMA1	Conotoxin pc3a	Conus pictus (Cone snail)
comp20403_c0_seq1	5282.15	5.00E-24	Q9YGJ9	Snake venom serine protease Haly-2	Gloydius brevicaudus (Korean slamosa snake)
comp62074_c0_seq3	2558.95	7.3	P0C176	Peptide TsPep3	Tityus serrulatus (Brazilian scorpion)
comp52195_c0_seq1	1794.84	0.67	B1P1H1	Kappa-theraphotoxin-Cg1d	Chilobrachys guangxiensis (Chinese earth tiger tarantula)
comp53334_c0_seq1*	**1783.89**	**4.4**	**P0C2B0**	**Conotoxin ViVA**	**Conus virgo (Virgin cone)**
comp35599_c0_seq1	**1710.94**	**1.4**	**Q95WD1**	**Toxin CsE8**	Centruroides sculpturatus (Bark scorpion)
comp59348_c0_seq1	1703.65	4.00E-06	P0DJG3	Thrombin-like enzyme TLBan	Bothrocophias andianus (Andean lancehead)
comp54362_c0_seq1	**1523.38**	**1.00E-19**	**G3LU44**	**Translationally-controlled tumor protein homolog**	Loxosceles intermedia (Brown spider)

Transcript	T FPKM	E-value	ID	Protein names	Organism
comp68091_c0_seq1*	8457.37	0.086	B6DD38	U14-lycotoxin-Ls1a	Lycosa singoriensis (Wolf spider) (Aranea singoriensis)
comp76287_c0_seq3*	4081.53	7.00E-04	A7X4K7	Waprin-Phi2	Philodryas olfersii (Green snake)
comp75135_c0_seq1*	2777.31	4.2	Q9BJV8	Omega-hexatoxin-Asp2a	Atrax sp. (strain Illawarra) (Funnel-web spider)
comp53334_c0_seq1*	**2667.8**	**4.4**	**P0C2B0**	**Conotoxin ViVA**	**Conus virgo (Virgin cone)**
comp35599_c0_seq1	**2478.75**	**1.4**	**Q95WD1**	**Toxin CsE8**	**Centruroides sculpturatus (Bark**

					scorpion)
comp54362_c0_seq1	**2247.73**	**1.00E-19**	**G3LU44**	**Translationally-controlled tumor protein homolog**	**Loxosceles intermedia (Brown spider)**
comp35571_c0_seq1	2145.5	1.7	G1AS77	Conotoxin Vc6.11	Conus victoriae (Queen Victoria cone)
comp77096_c0_seq1	2004.16	0.42	Q5Y4V3	U3-agatoxin-Ao1f	Agelena orientalis (Spider)
comp78185_c0_seq1	1775.11	7	P86990	Non-toxic venom protein	Rhopalurus junceus (Caribbean blue scorpion)
comp57678_c0_seq1	1691.18	1.9	B3FIU2	U12-theraphotoxin-Hs1a	Haplopelma schmidti (Chinese bird spider)

(b) SignalP Top BLAST hits

Transcript	GC FPKM	E-value	ID	Protein names	Organism
comp63246_c0_seq1*	65796.36	0.37	Q0TXH0	Eukaryotic translation initiation factor 3 subunit K	Phaeosphaeria nodorum (Glume blotch fungus)
comp78135_c0_seq1	56621.03	3.84E-11	Q60648	Ganglioside GM2 activator	Mus musculus (Mouse)
comp20389_c0_seq1	28885.47	1.52E-53	P40313	Chymotrypsin-like protease CTRL-1	Homo sapiens (Human)
comp76457_c0_seq1	24676.96	8.5	Q5FCW3	Putative elongation factor Tu-like protein	Ehrlichia ruminantium (strain Welgevonden)
comp61419_c0_seq1	19420.18	1.86E-84	P48052	Carboxypeptidase A2	Homo sapiens (Human)
comp53328_c1_seq1	16975.74	-	-	-	-
comp53348_c0_seq1	16780.85	1.3	Q7XVM8	Transport inhibitor response 1-like protein Os04g0395600	Oryza sativa subsp. japonica (Rice)
comp66861_c0_seq1*	15029.07	7.67E-51	P08419	Chymotrypsin-like elastase family member 2A	Sus scrofa (Pig)
comp62035_c0_seq1	14625.55	2.28E-46	B8V7S0	CUB and peptidase domain-containing protein 1	Acropora millepora (Staghorn coral)
comp63353_c0_seq1*	13650.81	8.29E-	P04813	Chymotrypsinogen 2	Canis lupus familiaris (Dog)

		57			
Transcript	**T FPKM**	**E-value**	**ID**	**Protein names**	**Organism**
comp68091_c0_seq1*	8457.37	2.40E-58	Q6P4Z2	Collagen alpha-1	Xenopus tropicalis (Western clawed frog)
comp35596_c0_seq1	8436.25	1.22E-62	P02466	Collagen alpha-2	Rattus norvegicus (Rat)
comp67666_c0_seq1	7038.22	3.3	Q9C9L5	Wall-associated receptor kinase-like 9	Arabidopsis thaliana (Mouse-ear cress)
comp76287_c0_seq3*	4081.53	2.44E-51	B8V7R6	Collagen alpha chain	Acropora millepora (Staghorn coral)
comp78195_c0_seq1	3708.28	-	-	-	-
comp63341_c0_seq1	3257.39	7.08E-128	Q26636	Cathepsin L	Sarcophaga peregrina (Flesh fly)
comp75135_c0_seq1*	2777.31	1.10E-62	Q9U943	Apolipophorins [Cleaved into: Apolipophorin-2	Locusta migratoria (Migratory locust)
comp53334_c0_seq1*	2667.8	0	P63018	Heat shock cognate 71 kDa protein	Rattus norvegicus (Rat)
comp63348_c0_seq1	2591.46	1.37E-22	Q9VQ62	Protein NPC2 homolog	Drosophila melanogaster (Fruit fly)
comp20388_c0_seq1	2388.4	4.17E-91	Q7ZWJ4	60S ribosomal protein L18a	Danio rerio (Zebrafish)

Ultimately, we found that the tentacle did not exhibit particularly high expression levels of any candidate toxins that might potentially be correlated with envenomation symptoms, as all were expressed at proportionately lower levels compared to the gastric cirri. Unexpectedly when screening the tentacle transcriptome against the ToxProt data set, the four most highly expressed toxin-like transcripts had higher sequence similarity to proteins that are non-toxic from the SignalP analysis (Table 4.2b). This may be indicative of structural gene products that share similar protein motifs with venom from other taxa rather than the identified transcripts being actual toxins. Despite using several approaches to analyze these candidates herein, all structural genes with known roles in nematogenesis were consistently and exclusively expressed at high levels in the tentacle, while toxin-like transcripts were consistently expressed at proportionately higher levels in the gastric cirri (Fig 4.4; Fig 4.5; Table 4.2a). These results suggest comparatively minimal secretion of toxin-like genes in the tentacles, or that putative candidate toxin genes responsible for envenomation syndrome are expressed at such low levels in the tentacles that they evaded detection in our analysis.

Conclusion

In this study, our combined molecular and morphological analysis revealed that not only are the transcripts encoding putative toxins and toxic-like enzymes almost exclusively expressed in the gastric cirri, their expression is not particularly localized in a region where venom deployment is known to occur in cubozoans i.e., within nematocysts in the tentacle. Given our observations of a putative central gland secreting granules via duct-like channels in the gastric cirri, we have speculated that numerous granules secreted by the gastric cirri of *A. alata* contain a concoction of digestive enzymes and venom components as is reflected in the findings of the

transcriptome profiling of the gastric cirri in this study. This study serves as a preliminary investigation of the morphology and molecular characters of the digestive tract of a box jellyfish. In the future, extending this type of study to include other cubozoan taxa, in particular those lacking gastric cirri, can help elucidate whether gland cells with a dual role in secretion of venom and digestion proteins were present in the last common ancestor of Cubozoa. Our findings, in conjunction with the fact that gastric cirri, unlike nematocysts, are not a synapomorphy of the Cnidaria, suggests that the presence of a putative central "venom and digestion gland-like structure" possibly predates the emergence of gastric cirri, and maybe even nematocysts, and that protein components corresponding to digestive enzymes in the cubozoan digestive tract may have been recruited for novel venom-related roles in the tentacles and nematocyst warts where nematocysts are in highest abundance.

Supplemental Figures

Supplemental Figure 4.I Zinc metalloprotease gene tree. *A. alata* candidate zinc metalloprotease amino acid sequences (as estimated ORFs) were aligned with MAFFT (Katoh & Standley 2013) using the L-INS-i algorithm, BLOSUM62 scoring matrix, 1.53 gap open penalty, and 0.123 offset value, alongside zinc metalloproteases from several venomous and non-venomous animal species (obtained from NCBI Genbank). A zinc metalloprotease gene tree was constructed using the program FastTree (Price *et al.* 2010) using default settings and 1000 bootstrap replicates using seqboot in the PHYLIP package (Felsenstein 2009).

Supplemental Files

Supplemental File 4.I *A. alata* gastric phacella video (.mp4). Video was taken using light microscopy of the gastric phacella (composed of cirri) excised from a live *A. alata* box jellyfish (bell height ~70 mm) (a still frame is reproduced herein as Fig 4.2).

Supplemental File 4.II Custom python script for gene ontology (GO) analyses. This script was used to average the expression values derived from the EdgeR matrix across all GO groups.

Supplemental File 4.III Full list of gene ontology (GO) groups and transcripts for *A. alata*. This comprehensive list of all transcripts with blast hits corresponding to SignalP and/or ToxProt expression. Results are provided for individual REVIGO analyses on the gastric cirri tentacle samples.

Acknowledgements

I acknowledge my coauthor Jason Macrander. We thank Scott Tuason and Ron N. Larson for providing consent to use their photos, and to Allen G. Collins for help with photos and videos, who acknowledges the Mary & Robert Pew Public Education Fund and Austin Lin. Data were analyzed on the Smithsonian High Performance Computing Cluster Hydra and the Ohio Biodiversity Conservation Partnership (OBCP) cluster. We would like to thank Marymegan Daly and Lisle Gibbs for inviting us to the SICB venom symposium where an earlier version of this work was presented thanks to US National Science Foundation DEB-1257796.

Appendices

Bibliography

Acevedo, M.J., Fuentes, V.L., Olariaga, A., Canepa, A., Belmar, M.B., Bordehore, C. & Calbet, A. (2013) Maintenance, feeding and growth of *Carybdea marsupialis* (Cnidaria: Cubozoa) in the laboratory. *Journal of Experimental Marine Biology and Ecology* 439, 84–91.

Adonin, L.S., Shaposhnikova, T.G. & Podgornaya, O. (2012) *Aurelia aurita* (Cnidaria) Oocytes' Contact Plate Structure and Development. *PLoS ONE* 7, 1–10.

Agassiz, L. (1862) *Contributions to the natural history of the United States of America. Vol IV*. Little Brown, Boston. Available from: http://www.biodiversitylibrary.org/item/54510#page/9/mode/1up (accessed 2 Oct 2013).

Andrews, S. (2014) FastQC. *http://www.bioinformatics.babraham.ac.uk/projects/fastqc/*.

Appeltans, W., Ahyong, S., Anderson, G., Angel, M., Artois, T., Bailly, N. & et al. (2012) The Magnitude of Global Marine Species Diversity. *Current Biology* 22, 2189–2202.

Arendt, D., Hausen, H. & Purschke, G. (2009) The 'division of labour' model of eye evolution. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364, 2809–17.

Arneson, A.C. (1976) Life History of *Carybdea alata* Reynaud, 1830 (Cubomedusae). M.S. Thesis, University of Puerto Rico, Mayaguez, Commonwealth of Puerto Rico, USA.

Arneson, A.C. & Cutress, C.E. (1976) Coelenterate Ecology and Behavior. Life History of *Carybdea alata Reynaud, 1830 (Cubomedusae)*. G. O. Mackie (Ed). Plenum Publishing Corporation.

Aungsuchawan, S., Browdy, C.L. & Withyachumnarnkul, B. (2011) Sperm capacitation of the shrimp *Litopenaeus vannamei.*. *Aquaculture Research* 42, 188–195.

Avila Soria, G. (2009) Molecular characterization of *Carukia barnesi* and *Malo kingi*, Cnidaria; Cubozoa; Carybdeidae. PhD Thesis. James Cook University

Badré, S. (2014) Bioactive toxins from stinging jellyfish. *Toxicon : Official journal of the International Society on Toxinology* 91, 114–25.

Balasubramanian, P.G., Beckmann, A., Warnken, U., Schnölzer, M., Schüler, A., Bornberg-Bauer, E., Holstein, T.W. & Özbek, S. (2012) Proteome of *Hydra* nematocyst. *Journal of Biological Chemistry* 287, 9672–9681.

Barnes, J.H. (1964) Cause and efects of Irukandji stingings. *The Medical journal of Australia*, 34–37.

Barnes, J.H. (1966) The Cnidarian and their evolution. *Studies on Three Venomous Cubomedusae.* Academic Press, London and New York.

Bauchot, M.-L., Daget, J. & Bauchot, R. (1990) L'ichtyologie en France au début du XIXe siecle. L'histoire naturelle despoissons de Cuvier et Valenciennes. *(Bulletin du Muséum national d'Histoire naturelle 4e ser. 12 sect A, no 1, supplément: 3–142). Paris:* Museum national d'histoire naturelle, Paris.

Beckmann, A. & Özbek, S. (2012) The nematocyst: A molecular map of the cnidarian stinging organelle. *International Journal of Developmental Biology* 56, 577–582.

Bentlage, B. (2010) *Carybdea alata* auct. (Cubozoa): rediscovery of the *Alatina grandis*

type. *Zootaxa* 2713, 52–54.
Bentlage, B., Cartwright, P., Yanagihara, A.A., Lewis, C., Richards, G.S. & Collins, A.G. (2010) Evolution of box jellyfish (Cnidaria: Cubozoa), a group of highly toxic invertebrates. *Proceedings of the Royal Society B: Biological Sciences* 277, 493–501.
Bentlage, B. & Lewis, C. (2012) An illustrated key and synopsis of the families and genera of carybdeid box jellyfishes (Cnidaria: Cubozoa: Carybdeida), with emphasis on the "Irukandji family" (Carukiidae). *Journal of Natural History* 46, 2595–2620.
Bielecki, J., Zaharoff, A.K., Leung, N.Y., Garm, A. & Oakley, T.H. (2014) Ocular and extraocular expression of opsins in the rhopalium of *Tripedalia cystophora* (Cnidaria: Cubozoa). *PloS one* 9, e98870.
Bigelow, H.B. (1918) Some Medusae and Siphonophora from the western Atlantic. Intergovernmental Panel on Climate Change (Ed). *Bulletin of the Museum of Comparative Zoology of Harvard College* 62, 363–442.
Bigelow, H.B. (1938) Plankton of Bermuda Oceanographic Expeditions. VIII. Medusae Taken During the Years 1929 and 1930. *Zoologica. Scientific Contributions of the New York Zoological Society* 23, 99–189.
Bottrall, J.L., Madaras, F., Biven, C.D., Venning, M.G. & Mirtschin, P.J. (2010) Proteolytic activity of Elapid and Viperid Snake venoms and its implication to digestion. *Journal of venom research* 1, 18–28.
Brekhman, V., Malik, A., Haas, B., Sher, N. & Lotan, T. (2015) Transcriptome profiling of the dynamic life cycle of the scypohozoan jellyfish *Aurelia aurita*. *BMC Genomics* 16.
Bridge, D., Cunningham, C.W., Schierwater, B., Desalle, R.O.B. & Buss, L.E.O.W. (1992) Class-level relationships in the phylum Cnidaria : Evidence from mitochondrial genome structure. *Proceedings of the National Academy of Sciences of the United States of America* 89, 8750–8753.
Brinkman, D.L., Aziz, A., Loukas, A., Potriquet, J., Seymour, J. & Mulvenna, J. (2012) Venom proteome of the box jellyfish *Chironex fleckeri*. *PloS one* 7, e47866.
Brinkman, D.L. & Burnell, J.N. (2009) Biochemical and molecular characterisation of cubozoan protein toxins. *Toxicon* 54, 1162–1173.
Brinkman, D.L., Jia, X., Potriquet, J., Kumar, D., Dash, D., Kvaskoff, D. & Mulvenna, J. (2015) Transcriptome and venom proteome of the box jellyfish *Chironex fleckeri*. *BMC Genomics* 16, 407.
Burnett, J.W. (2009) Treatment of Atlantic cnidarian envenomations. *Toxicon* 54, 1201–1205.
Calder, D.R. (2009) Cubozoan and scyphozoan jellyfishes of the Carolinian Biogeographic Province, southeastern USA. *Royal Ontario Museum Contributions in Science 3*, 1–58.
Calvete, J.J., Sanz, L., Angulo, Y., Lomonte, B. & Gutiérrez, J.M. (2009) Venoms, venomics, antivenomics. *FEBS Letters* 583, 1736–1743.
Carrette, T., Alderslade, P. & Seymour, J. (2002) Nematocyst ratio and prey in two Australian cubomedusans, *Chironex fleckeri* and *Chiropsalmus* sp. *Toxicon* 40, 1547–1551.
Carrette, T., Straehler-Pohl, I. & Seymour, J. (2014) Early Life History of *Alatina* cf. *moseri* Populations from Australia and Hawaii with Implications for Taxonomy

(Cubozoa: Carybdeida, Alatinidae). *PloS one* 9, e84377.
Cartwright, P. & Collins, A. (2007) Fossils and phylogenies: Integrating multiple lines of evidence to investigate the origin of early major metazoan lineages. *Integrative and Comparative Biology* 47, 744–751.
Cartwright, P., Halgedahl, S.L., Hendricks, J.R., Jarrard, R.D., Marques, A.C., Collins, A.G. & Lieberman, B.S. (2007) Exceptionally preserved jellyfishes from the Middle Cambria. *PLoS ONE* 2, 1–7.
Casewell, N.R., Wüster, W., Vonk, F.J., Harrison, R. a & Fry, B.G. (2012) Complex cocktails: the evolutionary novelty of venoms. *Trends in ecology & evolution* 28, 219–29.
Chang, E.S., Neuhof, M., Rubinstein, N.D., Diamant, A., Philippe, H., Huchon, D. & Cartwright, P. (2015) Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. *Proceedings of the National Academy of Sciences* 112, 201511468.
Chen, Y., Mccarthy, D., Ritchie, M., Robinson, M. & Smyth, G.K. (2016) *edgeR : differential expression analysis of digital gene expression data User ' s Guide*.
Chera, S., de Rosa, R., Miljkovic-Licina, M., Dobretz, K., Ghila, L., Kaloulis, K. & Galliot, B. (2006) Silencing of the *Hydra* serine protease inhibitor Kazal1 gene mimics the human SPINK1 pancreatic phenotype. *Journal of cell science* 119, 846–857.
Chiaverano, L.M., Holland, B.S., Crow, G.L., Blair, L. & Yanagihara, A.A. (2013) Long-term fluctuations in circalunar Beach aggregations of the box jellyfish *Alatina moseri* in Hawaii, with links to environmental variability. *PloS one* 8, e77039.
Chung, J.J., Ratnapala, L.A., Cooke, I.M. & Yanagihara, A.A. (2001) Partial purification and characterization of a hemolysin (CAH1) from Hawaiian box jellyfish (*Carybdea alata*) venom. *Toxicon* 39, 981–990.
Collins, A.G. (2009) Recent Insights into Cnidarian Phylogeny. 000.
Collins, A.G., Bentlage, B., Gillan, W., Lynn, T.H., Morandini, A.C. & Marques, A.C. (2011) Naming the Bonaire banded box jelly, *Tamoya ohboya*, n. sp. (Cnidaria: Cubozoa: Carybdeida: Tamoyidae). *Zootaxa* 68, 53–68.
Collins, A.G., Schuchert, P., Marques, A.C., Jankowski, T., Medina, M. & Schierwater, B. (2006) 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.
Conant, F.S. (1897) Notes on the Cubomedusae. *Johns Hopkins Univeristy Circulars* 132, 8–10.
Conant, F.S. (1898) *The Cubo*medusae. By Franklin Story Conant. A memorial volume. Memoirs from the biological laboratory of the Johns Hopkins University, IV, 1. *The Johns Hopkins Press, Baltimore.* Available from: http://www.biodiversitylibrary.org/item/16812#page/9/mode/1up.
Crow, G.L. (2015) Box Jellyfish (Cubozoa : Carybdeida) in Hawaiian Waters, and the First Record of *Tripedalia cystophora* in Hawai'i. 108, 93–108.
Cutress, C.E. (1970) Investigation of the biology and control of noxious coelenterates occurring in the coastal waters of Peurto Rico. First Annual Report (January 1 to December 31, 1969). *Jellyfish Project JF 2-6*. Commonwealth of Puerto Rico, Mayaguez
Cutress, C.E. (1971) Investigation of the biology and control of noxious coelenterates

occurring in the coastal waters of Peurto Rico. Second Annual Report (January 1 to December 31, 1970). *Jellyfish Project JF 2-6*. Commonwealth of Puerto Rico, Mayaguez

Cutress, C.E. (1972) Investigation of the biology and contorl of noxious coelenterates occurring in the coastal water of Puerto Rico. Third annual and summary report (January 1 to December 31, 1971). *Jellyfish Project JF 2-6 (Jellyfish Act PL-89-720)*. Commonwealth of Puerto Rico

Daly, M., Brugler, M.R., Cartwright, P., Collins, A.G., Dawson, M.N., Fautin, D.G., France, S.C., Mcfadden, C.S., Opresko, D.M., Rodriguez, E., Romano, S.L. & Stake, J.L. (2007) The phylum Cnidaria : A review of phylogenetic patterns and diversity 300 years after Linnaeus. 182, 127–182.

Danneels, E., Van Vaerenbergh, M., Debyser, G., Devreese, B. & de Graaf, D. (2015) Honeybee Venom Proteome Profile of Queens and Winter Bees as Determined by a Mass Spectrometric Approach. *Toxins* 7, 4468–4483.

David, C.N. & Challoner, D. (1974) Distribution of Interstitial Cells and Differentiating Nematocytes in Nests in *Hydra attenuata*. *American Zoologist* 14, 537–542.

David, C.N., Ozbek, S., Adamczyk, P., Meier, S., Pauly, B., Chapman, J., Hwang, J.S., Gojobori, T. & Holstein, T.W. (2008) Evolution of complex structures: minicollagens shape the cnidarian nematocyst. *Trends in genetics : TIG* 24, 431–8.

Desmond Ramirez, M., Speiser, D.I., Sabrina Pankey, M. & Oakley, T.H. (2011) Understanding the dermal light sense in the context of integrative photoreceptor cell biology. *Visual neuroscience* 28, 265–279.

Dunn, C.W., Giribet, G., Edgecombe, G.D. & Hejnol, A. (2014) Animal Phylogeny and Its Evolutionary Implications. *Annual Review of Ecology, Evolution, and Systematics* 45, 371–395.

Eckert, T. (2012) In silico Study on Sulfated and Non-Sulfated Carbohydrate Chains from Proteoglycans in Cnidaria and Interaction with Collagen. *Open Journal of Physical Chemistry* 02, 123–133.

Engel, U., Ozbek, S., Streitwolf-Engel, R., Petri, B., Lottspeich, F. & Holstein, T.W. (2002) Nowa, a novel protein with minicollagen Cys-rich domains, is involved in nematocyst formation in *Hydra*. *Journal of cell science* 115, 3923–3934.

Felsenstein, J. (2009) PHYLIP (Phylogeny Inference Package).

Feuda, R., Hamilton, S.C., Mcinerney, J.O. & Pisani, D. (2012) Metazoan opsin evolution reveals a simple route to animal vision. 109, 18868–18872.

Feuda, R., Rota-Stabelli, O., Oakley, T.H. & Pisani, D. (2014) The comb jelly opsins and the origins of animal phototransduction. *Genome biology and evolution* 6, 1964–71.

Figueroa Rosa, B.J. (2015) Mucho ojo con las aguavivas. *Primerahora*.

Forêt, S., Knack, B., Houliston, E., Momose, T., Manuel, M., Quéinnec, E., Hayward, D.C., Ball, E.E. & Miller, D.J. (2010) New tricks with old genes: The genetic bases of novel cnidarian traits. *Trends in Genetics* 26, 154–158.

Fry, B.G., Roelants, K., Champagne, D.E., Scheib, H., Tyndall, J.D. a, King, G.F., Nevalainen, T.J., Norman, J. a, Lewis, R.J., Norton, R.S., Renjifo, C. & de la Vega, R.C.R. (2009) The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annual review of genomics and human genetics* 10, 483–511.

Fuchs, B., Wang, W., Graspeuntner, S., Li, Y., Insua, S., Herbst, E.M., Dirksen, P., B??hm, A.M., Hemmrich, G., Sommer, F., Domazet-Loo, T., Klostermeier, U.C.,

Anton-Erxleben, F., Rosenstiel, P., Bosch, T.C.G. & Khalturin, K. (2014) Regulation of polyp-to-jellyfish transition in *Aurelia aurita*. *Current Biology* 24, 263–273.

Furukawa, T., Kozak, C.A. & Cepko, C.L. (1997) Rax, a novel paired-type homeobox gene, shows expression in the anterior neural fold and developing retina. *Proceedings of the National Academy of Sciences* 94, 3088–3093.

Gacesa, R., Chung, R., Dunn, S.R., Weston, A.J., Jaimes-Becerra, A., Marques, A.C., Morandini, A.C., Hranueli, D., Starcevic, A., Ward, M. & Long, P.F. (2015) Gene duplications are extensive and contribute significantly to the toxic proteome of nematocysts isolated from *Acropora digitifera* (Cnidaria: Anthozoa: Scleractinia). *BMC Genomics* 16, 774.

Galliot, B. (2012) *Hydra*, a fruitful model system for 270 years. *International Journal of Developmental Biology* 56, 411–423.

Garm, A., Coates, M.M., Gad, R., Seymour, J. & Nilsson, D.-E. (2007)a) The lens eyes of the box jellyfish *Tripedalia cystophora* and *Chiropsalmus* sp. are slow and color-blind. *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology* 193, 547–57.

Garm, A., Lebouvier, M. & Tolunay, D. (2015) Mating in the box jellyfish *Copula sivickisi*-Novel function of cnidocytes. *Journal of morphology*.

Garm, A., O'Connor, M., Parkefelt, L. & Nilsson, D.-E. (2007)b) Visually guided obstacle avoidance in the box jellyfish *Tripedalia cystophora* and *Chiropsella bronzie*. *The Journal of Experimental Biology* 210, 3616–23.

Gegenbaur, C. (1857) Versuch eines Systemes der Medusen, mit Beschreibung neuer oder wenig gekannter Formen; zugleich ein Beitrag zur Kenntniss der Fauna des Mittelmeeres. *Zeitschrift für Wissenschaftliche Zoologie*, 202–273, plates vii–x.

Gershwin, L., Richardson, A.J., Winkel, K.D., Fenner, P.J., Lippmann, J., Hore, R., Avila-Soria, G., Brewer, D., Kloser, R.J., Steven, A. & Condie, S. (2013) Biology and ecology of irukandji jellyfish (Cnidaria: Cubozoa). *Advances in Marine Biology*. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24182899.

Gershwin, L.-A. (2005) *Carybdea alata* auct. and *Manokia stiasnyi*, reclassification to a new family with description of a new genus and two new species. *Memoirs of the Queensland Museum* 51, 501–523.

Gershwin, L.-A. (2006) Zootaxa, Nematocysts of the Cubozoa. 57, 1–57.

Gershwin, L.-A., De Nardi, M., Winkel, K.D. & Fenner, P.J. (2010) Marine Stingers: Review of an Under-Recognized Global Coastal Management Issue. *Coastal Management* 38, 22–41.

Gnerre, S., Maccallum, I., Przybylski, D., Ribeiro, F.J., Burton, J.N., Walker, B.J., Sharpe, T., Hall, G., Shea, T.P., Sykes, S., Berlin, A.M., Aird, D., Costello, M., Daza, R., Williams, L., Nicol, R., Gnirke, A., Nusbaum, C., Lander, E.S. & Jaffe, D.B. (2011) High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proceedings of the National Academy of Sciences of the United States of America* 108, 1513–8.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. & Regev, A. (2011) Full-length transcriptome assembly from

RNA-Seq data without a reference genome. *Nature Biotechnology* 29, 644–652.
Graham, M. (1998) Short Papers and Notes: First Report of *Carybdea alata* var. *grandis* (Reynaud 1830) (Cnidaria: Cubozoa) from the Gulf of Mexico. *Gulf of Mexico Science* 1, 28–30.
Gruhl, A. & Okamura, B. (2012) Development and myogenesis of the vermiform *Buddenbrockia* (Myxozoa) and implications for cnidarian body plan evolution. *EvoDevo* 3, 10.
Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N. & Regev, A. (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* 8, 1494–1512.
Haeckel, E. (1880) *System der Acraspeden – Zweite Hälfte des Systems der Medusen*. Denkschriften der Medizinisch– Naturwissenschaftlichen Gesellschaft zu Jena, Germany.
Hamner, W., Jones, M. & Hamner, P. (1995) Swimming, feeding, circulation and vision in the Australian box jellyfish, *Chironex fleckeri* (Cnidaria: Cubozoa). *Marine and Freshwater Research* 46, 985–990.
Hausman, R.E. & Burnett, A.L (1971) The mesoglea of *Hydra*. IV. A qualitative radioautographic study of the protein component. *Journal of Experimental Zoology* 177, 435–446.
Hayakawa, H., Andoh, T. & Watanabe, T. (2007) Identification of a novel yolk protein in the hermatypic coral *Galaxea fascicularis*. *Zoological Science* 24, 249–255.
Hayward, A., Takahashi, T., Bendena, W.G., Tobe, S.S. & Hui, J.H.L. (2010) Comparative genomic and phylogenetic analysis of vitellogenin and other large lipid transfer proteins in metazoans. *FEBS Letters* 584, 1273–1278.
Hessinger, D.A. & Lenhoff, H.M. (1988) *Nematocysts The Biology of*. Academic Press Inc.
Hofmann, D.K., Neumann, R. & Henne, K. (1978) Strobilation, budding and initiation of scyphistoma morphogenesis in the rhizostome *Cassiopea andromeda* (Cnidaria: Scyphozoa). *Marine Biology* 47, 161–176.
Horibata, Y., Higashi, H. & Ito, M. (2001) Transglycosylation and Reverse Hydrolysis Reactions of Endoglycoceramidase from the Jellyfish. *Journal of Biochemistry* 130, 263–268.
Houliston, E., Momose, T. & Manuel, M. (2010) *Clytia hemisphaerica*: a jellyfish cousin joins the laboratory. *Trends in genetics : TIG* 26, 159–67.
Humann, P. & Deloach, N. (2002) *Reef creature Identification: Florida, Caribbean, Bahamas*. 2nd ed. New World Publications Inc.
Hwang, J.S., Takaku, Y., Momose, T., Adamczyk, P., Özbek, S., Ikeo, K., Khalturin, K., Hemmrich, G., Bosch, T.C.G., Holstein, T.W., David, C.N. & Gojobori, T. (2010) Nematogalectin, a nematocyst protein with GlyXY and galectin domains, demonstrates nematocyte-specific alternative splicing in *Hydra*. *Proceedings of the National Academy of Sciences of the United States of America* 107, 18539–18544.
Ishida, J. (1936) Note on the digestion of *Charybdea rastonii*. *Annotations Zoological of Japan* 15, 449–452.

Iyengar, M.R., Iyengar, C.W., Chen, H.Y., Brinster, R.L., Bornslaeger, E. & Schultz, R.M. (1983) Expression of creatine kinase isoenzyme during oogenesis and embryogenesis in the mouse. *Developmental biology* 96, 263–268.

Jouiaei, M., Casewell, N.R., Yanagihara, A.A., Nouwens, A., Cribb, B.W., Whitehead, D., Jackson, T.N.W., Ali, S. a, Wagstaff, S.C., Koludarov, I., Alewood, P., Hansen, J. & Fry, B.G. (2015)a) Firing the sting: chemically induced discharge of cnidae reveals novel proteins and peptides from box jellyfish (*Chironex fleckeri*) venom. *Toxins* 7, 936–50.

Jouiaei, M., Yanagihara, A.A., Madio, B., Nevalainen, T., Alewood, P. & Fry, B. (2015)b) Ancient Venom Systems: A Review on Cnidaria Toxins. *Toxins* 7, 2251–2271.

Júnior, M.N. & Haddad, M.A. (2008) The diet of cubomedusae (cnidaria, cubozoa) in southern Brazil. 56, 157–164.

Junqueira-de-Azevedo, I.L.M., Bastos, C.M.V., Ho, P.L., Luna, M.S., Yamanouye, N. & Casewell, N.R. (2015) Venom-related transcripts from *Bothrops jararaca* tissues provide novel molecular insights into the production and evolution of snake venom. *Molecular biology and evolution* 32, 754–66.

Kass-Simon, G. & Scappaticci, Jr. (2002) The behavioral and developmental physiology of nematocysts. *Canadian Journal of Zoology* 80, 1772–1794.

Katoh, K. & Standley, D.M. (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* 30, 772–780.

Kayal, E., Bentlage, B., Collins, A.G., Kayal, M., Pirro, S. & Lavrov, D. V (2012) Evolution of linear mitochondrial genomes in medusozoan cnidarians. *Genome biology and evolution* 4, 1–12.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics (Oxford, England)* 28, 1647–9.

Kitatani, R., Yamada, M., Kamio, M. & Nagai, H. (2015) Length Is Associated with Pain: Jellyfish with Painful Sting Have Longer Nematocyst Tubules than Harmless Jellyfish. *Plos One* 10, e0135015.

Koyanagi, M., Takano, K., Tsukamoto, H., Ohtsu, K., Tokunaga, F. & Terakita, A. (2008) Jellyfish vision starts with cAMP signaling mediated by opsin-G(s) cascade. *Proceedings of the National Academy of Sciences of the United States of America* 105, 15576–80.

Kozmik, Z., Ruzickova, J., Jonasova, K., Matsumoto, Y., Vopalensky, P., Kozmikova, I., Strnad, H., Kawamura, S., Piatigorsky, J., Paces, V. & Vlcek, C. (2008)a) Assembly of the cnidarian camera-type eye from vertebrate-like components. *Proceedings of the National Academy of Sciences of the United States of America* 105, 8989–93.

Kozmik, Z., Swamynathan, S.K., Ruzickova, J., Jonasova, K., Paces, V., Vlcek, C. & Piatigorsky, J. (2008)b) Cubozoan crystallins: Evidence for convergent evolution of pax regulatory sequences. *Evolution and Development* 10, 52–61.

Kramp, P.L. (1961) Synopsis of the Medusae of the World. *Journal of the Marine Biological Association of the UK* 40, 1–149.

Krueger, F. (2012) Trim Galore!
[http://www.bioinformatics.babraham.ac.uk/projects/trim galore/]
Kumar, J.P. (2009) The sine oculis homeobox (SIX) family of transcription factors as regulators of development and disease. *Cell Mol Life Sci.* 66, 565–583.
Kurz, E.M., Holstein, T.W., Petri, B.M., Engel, J. & David, C.N. (1991) Mini-collagens in *Hydra* nematocytes. *Journal of Cell Biology* 115, 1159–1169.
Land, M.F. & Nilsson, D.-E. (2002) *Animal Eyes (Oxford Animal Biology Series)*. Oxford University Press, USA.
Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* 10, R25.
Larson, R.J. (1976) Cubomedusae: Feeding functional morphology, behavior and phylogenetic position. In: G. O. Mackie (Ed), *Coelenterate Ecology and Behavior*. Springer US, Boston, MA, pp. 237–245.
Larson, R.J., Mills, C.E. & Harbison, G.R. (1991) Western Atlantic midwater hydrozoan and scyphozoan medusae: in situ studies using manned submersibles. *Hydrobiologia* 216-217, 311–317.
Lawley, J.W., Lewis Ames, C., Bentlage, B., Yanagihara, A.A., Goodwill, R., Kayal, E., Hurwitz, K. & Collins, A.G. (In press). The box jellyfish *Alatina alata* has a circumtropical distribution. *The Biological Bulletin*.
Lesson, R. (1837) Prodrome d'une monographie des Méduses. Extraite d'une Histoire manuscrite des méduses, en 3 volumes 4e avec 200 planches coloriées, ouvrage entièrement termini. Rochefort.
Lesson, R. (1843) Historie naturelle des zoophytes. Acalèphes. *Encyclopédique de Roret.* Roret, Paris.
Levitan, S., Sher, N., Brekhman, V., Ziv, T., Lubzens, E. & Lotan, T. (2015) The making of an embryo in a basal metazoan: Proteomic analysis in the sea anemone *Nematostella vectensis*. *Proteomics* 15, 4096–4104.
Levy, O., Appelbaum, L., Leggat, W., Gothlif, Y., Hayward, D.C., Miller, D.J. & Hoegh-Guldberg, O. (2007) Light-Responsive Cryptochromes from a Simple Multicellular Animal, the Coral Acropora millepora. *Science* 318, 467–470.
Lewis Ames, C., Collins, A.G., Ryan, J.F., Bely, A.E. & Cartwright, P. 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* In Press
Lewis, C. & Bentlage, B. (2009) Clarifying the identity of the Japanese Habu-kurage, *Chironex yamaguchii,* sp. nov. (Cnidaria: Cubozoa: Chirodropida). *Zootaxa* 65, 59 – 65.
Lewis, C., Bentlage, B., Yanagihara, A.A., Gillan, W., Blerk, J.V.A.N., Keil, D.P., Bely, A.E. & Collins, A.G. (2013) Redescription of *Alatina alata* (Reynaud, 1830) (Cnidaria: Cubozoa) from Bonaire, Dutch Caribbean. *Zootaxa* 3737, 473–487.
Lewis, C. & Long, T.A.F. (2005) Courtship and reproduction in *Carybdea sivickisi* (Cnidaria: Cubozoa). *Marine Biology* 147, 477–483.
Li, B. & Dewey, C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics* 12, 323.
Li, R., Yu, H., Xue, W., Yue, Y., Liu, S., Xing, R. & Li, P. (2014) Jellyfish venomics and

venom gland transcriptomics analysis of *Stomolophus meleagris* to reveal the toxins associated with sting. *Journal of Proteomics* 106, 17–29.
Liedvogel, M. & Mouritsen, H. (2010) Cryptochromes--a potential magnetoreceptor: what do we know and what do we want to know? *Journal of the Royal Society, Interface / the Royal Society* 7 Suppl 2, S147–S162.
Liegertová, M., Pergner, J., Kozmiková, I., Fabian, P., Pombinho, A.R., Strnad, H., Pačes, J., Vlček, Č., Bartůněk, P. & Kozmik, Z. (2015) Cubozoan genome illuminates functional diversification of opsins and photoreceptor evolution. *Scientific Reports* 5, 11885.
Linnaeus, C. (1758) i-iv Holmiae Systema Naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Tomus I. *Editio dec. Laurentii Salvii.*
Macmanes, M.D. & Eisen, M.B. (2013) Improving transcriptome assembly through error correction of high-throughput sequence reads. *PeerJ* 1, e113.
Macrander, J., Broe, M. & Daly, M. (2016) Tissue-Specific Venom Composition and Differential Gene Expression in Sea Anemones. *Genome biology and evolution*.
Macrander, J., Brugler, M.R. & Daly, M. (2015) A RNA-seq approach to identify putative toxins from acrorhagi in aggressive and non-aggressive *Anthopleura elegantissima* polyps. *BMC genomics* 16, 221.
Mali, B., Möhrlen, F., Frohme, M. & Frank, U. (2004) A putative double role of a chitinase in a cnidarian: pattern formation and immunity. *Developmental and comparative immunology* 28, 973–81.
Manuel, M. (2009) Early evolution of symmetry and polarity in metazoan body plans. *Comptes Rendus - Biologies* 332, 184–209.
Mariscal, R.N. (1974) Nematocysts. In: L. Muscatine and H. Lenoff (Eds), *Coelenterate biology: reviews and new perspectives*. Academic Press, New York, pp. 129 – 178.
Mariscal, R.N., Conklin, E.J. & Bigger, C.H. (1977) The Ptychocyst, a major new category of cnida used in tube construction by a cerianthid anemone. *Biological Bulletin* 152, 392–405.
Marques, A.C. & Collins, A.G. (2004)a) Cladistic analysis of Medusozoa and cnidarian evolution. *Invertebrate Biology* 123, 23–42.
Marques, A.C. & Collins, A.G. (2004)b) Cladistic analysis of Medusozoa and cnidarianevolution. *Invertebrate Biology* 123, 23–42.
Marques, A.C., García, J. & Lewis Ames, C. (2015) Internal fertilization and sperm storage in cnidarians: a response to Orr and Brennan. *Trends in Ecology & Evolution*, 1–2.
Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17.1, 10–12.
Mason, B., Schmale, M., Gibbs, P., Miller, M.W., Wang, Q., Levay, K., Shestopalov, V. & Slepak, V.Z. (2012) Evidence for Multiple Phototransduction Pathways in a Reef-Building Coral. *PLoS ONE* 7, 1–9.
Mayer, A. (1910) The Scyphomedusae. In: *Medusae of the World. Vol. III.* Carnegie Institution of Washington Publication, pp. 109(3), 499–735.
Mayer, A.G. (1906) Medusae of the Hawaiian Islands Collected by the Steamer albatross in 1902. *Bulletin of the US Fish Commission* 23, 1131–1143.
Mayer, A.G. (1915) *M*edusae of the Philippines and of Torres Straits. Being a report upon

the Scyphomedusae collected by the United States Fisheries Bureau Steamer "Albatross" in the Philippine islands and Malay archipelago, 1901-1910, and upon the Medusae collected by the expedition of the Carnegie institute of Washington D. C. to Torres Strait, Australia, in 1913. Carnegie Institution of Washington.
McClounan, S. & Seymour, J. (2012) Venom and cnidome ontogeny of the cubomedusae *Chironex fleckeri*. *Toxicon* 60, 1335–1341.
De Meyer, K. (1997) Bonaire, Netherlands Antilles. Coastal Region and Small Island (CSI). *Environment and development in coastal regions and in small islands.*, 1–9.
Milde, S., Hemmrich, G., Anton-Erxleben, F., Khalturin, K., Wittlieb, J. & Bosch, T.C.G. (2009) Characterization of taxonomically restricted genes in a phylum-restricted cell type. *Genome biology* 10, R8.
Miller, R.L. (1983) Cnidaria. In: K. G. Adiyodi and R. G. Adiyodi (Eds), *Reproductive Biology of Invertebrates, Volume II: Spermatogenesis and Sperm Function*. John Wiley & Sons Ltd., pp. pp. 23 – 73.
Möhrlen, F., Maniura, M., Plickert, G., Frohme, M. & Frank, U. (2006) Evolution of astacin-like metalloproteases in animals and their function in development. *Evolution and Development* 8, 223–231.
Moran, Y., Genikhovich, G., Gordon, D., Wienkoop, S., Zenkert, C., Ozbek, S., Technau, U. & Gurevitz, M. (2012) Neurotoxin localization to ectodermal gland cells uncovers an alternative mechanism of venom delivery in sea anemones. *Proceedings of the Royal Society B: Biological Sciences* 279, 1351–1358.
Moran, Y., Praher, D., Schlesinger, A., Ayalon, A., Tal, Y. & Technau, U. (2013) Analysis of Soluble Protein Contents from the Nematocysts of a Model Sea Anemone Sheds Light on Venom Evolution. *Marine Biotechnology* 15, 329–339.
Morandini, A.C. (2003) Deep-Sea medusae (Cnidaria : Cubozoa , Hydrozoa and Scyphozoa) from the coast of Bahia (western South Atlantic , Brazil). *Mitteilungen aus dem Museum für Naturkunde in Berlin. Zoologisches Museum und Institut für Spezielle Zoologie (Berlin)*, 13–25.
Nagai, H. (2003) Recent Progress in Jellyfish Toxin Study. *Journal of Health Science* 49, 337–340.
Nagai, H., Takuwa, K., Nakao, M., Ito, E., Miyake, M., Noda, M. & Nakajima, T. (2000)a) Novel Proteinaceous Toxins from the Box Jellyfish (Sea Wasp) *Carybdea rastoni*. *Biochemical and Biophysical Research Communications* 275, 582–588.
Nagai, H., Takuwa, K., Nakao, M., Sakamoto, B., Crow, G.L. & Nakajima, T. (2000)b) Isolation and Characterization of a Novel Protein Toxin from the Hawaiian Box Jellyfish (Sea Wasp) *Carybdea alata*. *Biochemical and Biophysical Research Communications* 275, 589–594.
Nevalainen, T.J., Peuravuori, H.J., Quinn, R.J., Llewellyn, L.E., Benzie, J. A H., Fenner, P.J. & Winkel, K.D. (2004) Phospholipase A2 in Cnidaria. *Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology* 139, 731–5.
Nilsson, D.-E., Gislén, L., Coates, M.M., Skogh, C. & Garm, A. (2005) Advanced optics in a jellyfish eye. *Nature* 435, 201–205.
Nordström, K., Wallén, Seymour, J. & Nilsson, D. (2003) A simple visual system without neurons in jellyfish larvae. *Proceedings of the Royal Society of London B: Biological Sciences* 270, 2349–2354.
O'Connor, M., Garm, A., Marshall, J.N., Hart, N.S., Ekström, P., Skogh, C. & Nilsson,

D.-E. (2010) Visual pigment in the lens eyes of the box jellyfish *Chiropsella bronzie*. *Proceedings. Biological sciences / The Royal Society* 277, 1843–8.
Okamura, B., Gruhl, A. & Bartholomew, J.L. (2015) Myxozoan evolution, ecology and development. *Myxozoan Evolution, Ecology and Development*, 1–441.
Oliveros, J.C. (2007) Venny. An interactive tool for comparing lists with Venn's diagrams.
Östman, C. (2000) A guideline to nematocyst nomenclature and classification , and some notes on the systematic value of nematocysts. *Scientia Marina* 64, 31–46.
Ozbek, S. (2011) The cnidarian nematocyst: a miniature extracellular matrix within a secretory vesicle. *Protoplasma* 248, 635–40.
Özbek, S., Balasubramanian, P.G. & Holstein, T.W. (2009) Cnidocyst structure and the biomechanics of discharge. *Toxicon* 54, 1038–1045.
Pan, T.L., Gröger, H., Schmid, V. & Spring, J. (1998) A toxin homology domain in an astacin-like metalloproteinase of the jellyfish *Podocoryne carnea* with a dual role in digestion and development. *Development Genes and Evolution* 208, 259–266.
Parra, G., Bradnam, K. & Korf, I. (2007) CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* 23, 1061–1067.
Petersen, K. (1979) Development of coloniality in Hydrozoa. In: G. Larwood and B. Rosen (Eds), *Biology and Systematics of Colonial Organisms*. Academic Press, New York, pp. 105–139.
Petersen, T.N., Brunak, S., von Heijne, G. & Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Meth* 8, 785–786.
Piatigorsky, J., Brient, W.E.O., Norman, B.L., Kalumuckt, K., Wistow, G.J., Borras, T., Nickerson, J.M. & Wawrousek, E.F. (1988) Gene sharing by δ-crystallin and argininosuccinate lyase. *Amino Acids* 85, 3479–3483.
Piatigorsky, J., Horwitz, J., Kuwabara, T. & Cutress, C.E. (1989) The cellular eye lens and crystallins of cubomedusan jellyfish. *Journal of Comparative Physiology A* 164, 577–587.
Piatigorsky, J., Horwitzs, J. & Norman, B.L. (1993) J1-crystallins of the Cubomedusan Jellyfish Lens Constitute a Novel Family Encoded in at Least Three Intronless Genes. *The Journal of Biological Chemistry* 2643, 11894–11901.
Piatigorsky, J., Norman, B., Dishaw, L.J., Kos, L., Horwitz, J., Steinbach, P.J. & Kozmik, Z. (2001) J3-crystallin of the jellyfish lens: similarity to saposins. *Proceedings of the National Academy of Sciences of the United States of America* 98, 12362–7.
Pineda, A.O. & Ellington, W.R. (1999) Structural and functional implications of the amino acid sequences of dimeric, cytoplasmic and octameric mitochondrial creatine kinases from a protostome invertebrate. *European Journal of Biochemistry* 264, 67–73.
Plachetzki, D.C., Degnan, B.M. & Oakley, T.H. (2007) The origins of novel protein interactions during animal opsin evolution. *PloS one* 2, e1054.
Plachetzki, D.C., Fong, C.R. & Oakley, T.H. (2010) The evolution of phototransduction from an ancestral cyclic nucleotide gated pathway. *Proceedings. Biological sciences / The Royal Society* 277, 1963–9.
Plachetzki, D.C., Fong, C.R. & Oakley, T.H. (2012)a) Cnidocyte discharge is regulated by light and opsin-mediated phototransduction. *BMC Biology* 10, 17.
Plachetzki, D.C., Fong, C.R. & Oakley, T.H. (2012)b) Cnidocyte discharge is regulated

by light and opsin-mediated phototransduction. *BMC biology* 10, 17.
Ponce, D., Brinkman, D.L., Luna-Ramirez, K., Wright, C.E. & Dorantes-Aranda, J.J. (2015) Comparative study of the toxic effects of *Chrysaora quinquecirrha* (Cnidaria: Scyphozoa) and *Chironex fleckeri* (Cnidaria: Cubozoa) venoms using cell-based assays. *Toxicon* 106, 57–67.
Ponce, D., Brinkman, D.L., Potriquet, J. & Mulvenna, J. (2016) Tentacle Transcriptome and Venom Proteome of the Pacific Sea Nettle, *Chrysaora fuscescens* (Cnidaria: Scyphozoa). *Toxins* 8, 102.
Porter, M.L., Blasic, J.R., Bok, M.J., Cameron, E.G., Pringle, T., Cronin, T.W. & Robinson, P.R. (2012) Shedding new light on opsin evolution. *Proceedings. Biological sciences / The Royal Society* 279, 3–14.
Price, M.N., Dehal, P.S. & Arkin, A.P. (2010) FastTree 2 Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* 5, e9490.
Rachamim, T., Morgenstern, D., Aharonovich, D., Brekhman, V., Lotan, T. & Sher, D. (2014) The Dynamically Evolving Nematocyst Content of an Anthozoan, a Scyphozoan, and a Hydrozoan. *Molecular Biology and Evolution* 32, 740–753.
Rahman, M.S., Kwon, W.-S. & Pang, M.-G. (2014) Calcium Influx and Male Fertility in the Context of the Sperm Proteome: An Update. *BioMed Research International* 2014, 1–13.
Reft, A.J. & Daly, M. (2012) Morphology, distribution, and evolution of apical structure of nematocysts in hexacorallia. *Journal of Morphology* 273, 121–136.
Reitzel, A.M., Tarrant, A.M. & Levy, O. (2013) Circadian clocks in the cnidaria: Environmental entrainment, molecular regulation, and organismal outputs. *Integrative and Comparative Biology* 53, 118–130.
Reynaud (1830) *Carybdea alata n. sp. In: Centurie Zoologique.* R. P. Lesson (Ed). Levrault, Paris France.
Rivera, A.S., Ozturk, N., Fahey, B., Plachetzki, D.C., Degnan, B.M., Sancar, A. & Oakley, T.H. (2012) Blue-light-receptive cryptochrome is expressed in a sponge eye lacking neurons and opsin. *Journal of Experimental Biology* 215, 1278–1286.
Rivera, A.S., Winters, I., Rued, A., Ding, S., Posfai, D., Cieniewicz, B., Cameron, K., Gentile, L. & Hill, A. (2013) The evolution and function of the Pax/Six regulatory network in sponges. *Evolution and Development* 15, 186–196.
Robinson, M.D., McCarthy, D.J. & Smyth, G.K. (2010) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.
Robinson, M.D. & Oshlack, A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* 11, R25.
Sanders, S.M., Shcheglovitova, M. & Cartwright, P. (2014) Differential gene expression between functionally specialized polyps of the colonial hydrozoan *Hydractinia symbiolongicarpus* (Phylum Cnidaria). *BMC genomics* 15, 406.
Schiariti, A., Christiansen, E., Morandini, A.C., da Silveira, F.L., Giberto, D.A. & Mianzan, H.W. (2012) (Cnidaria: Scyphozoa: Rhizostomeae): Individual traits related to sexual reproduction. *Marine Biology Research* 8, 255–264.
Schnitzler, C.E., Pang, K., Powers, M.L., Reitzel, A.M., Ryan, J.F., Simmons, D., Tada, T., Park, M., Gupta, J., Brooks, S.Y., Blakesley, R.W., Yokoyama, S., Haddock, S.H., Martindale, M.Q. & Baxevanis, A.D. (2012) Genomic organization, evolution,

and expression of photoprotein and opsin genes in *Mnemiopsis leidyi*: a new view of ctenophore photocytes. *BMC biology* 10, 107.
Shichida, Y. & Matsuyama, T. (2009) Evolution of opsins and phototransduction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364, 2881–95.
Shikina, S., Chen, C.J., Chung, Y.J., Shao, Z.F., Liou, J.Y., Tseng, H.P., Lee, Y.H. & Chang, C.F. (2013) Yolk formation in a stony coral *Euphyllia ancora* (Cnidaria, Anthozoa): Insight into the evolution of vitellogenesis in nonbilaterian animals. *Endocrinology* 154, 3447–3459.
Shimodaira, H. & Hasegawa, M. (1999) Letter to the Editor Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference. *Molecular Biology and Evolution* 16, 1114–1116.
Shinzato, C., Shoguchi, E., Kawashima, T., Hamada, M., Hisata, K., Tanaka, M., Fujie, M., Fujiwara, M., Koyanagi, R., Ikuta, T., Fujiyama, A., Miller, D.J. & Satoh, N. (2011) Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature* 476, 320–323.
Shpirer, E., Chang, E., Diamant, A., Rubinstein, N., Cartwright, P. & Huchon, D. (2014) Diversity and evolution of myxozoan minicollagens and nematogalectins. *BMC Evolutionary Biology* 14, 205.
Smith, D.R., Kayal, E., Yanagihara, A.A., Collins, A.G., Pirro, S. & Keeling, P.J. (2012) First complete mitochondrial genome sequence from a box jellyfish reveals a highly fragmented linear architecture and insights into telomere evolution. *Genome biology and evolution* 4, 52–8.
Southcott, R. V. (1967) Revision of some Carybdeidae (Scyphozoa: Cubomedusae), including a description of the jellyfish responsible for the "Irukandji syndrome." *Australian Journal of Zoology*, 651–671.
Speiser, D.I., Pankey, M., Zaharoff, A.K., Battelle, B. a, Bracken-Grissom, H.D., Breinholt, J.W., Bybee, S.M., Cronin, T.W., Garm, A., Lindgren, A.R., Patel, N.H., Porter, M.L., Protas, M.E., Rivera, A.S., Serb, J.M., Zigler, K.S., Crandall, K. a & Oakley, T.H. (2014) Using phylogenetically-informed annotation (PIA) to search for light-interacting genes in transcriptomes from non-model organisms. *BMC Bioinformatics* 15, 350.
Stewart, S.E. (1996) Field behavior of *Tripedalia cystophora* (class Cubozoa). *Marine and Freshwater Behaviour Physiology* 27, 175–188.
Straehler-Pohl, I., Garm, A. & Morandini, A.C. (2014) Sexual dimorphism in Tripedaliidae (Conant 1897) (Cnidaria, Cubozoa, Carybdeida). 3785, 533–549.
Suga, H., Schmid, V. & Gehring, W.J. (2008) Evolution and Functional Diversity of Jellyfish Opsins. *Current Biology* 18, 51–55.
Tamanaha, R.H. & Izumi, A.K. (1996) Persistent cutaneous hypersensitivity reaction after a Hawaiian box jellyfish sting (*Carybdea alata*). *Journal of the American Academy of Dermatology* 35, 991–993.
Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E. V, Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Smirnov, S., Sverdlov, A. V, Vasudevan, S., Wolf, Y.I., Yin, J.J. & Natale, D. a (2003) The COG database: an updated version includes eukaryotes. *BMC bioinformatics* 4, 41.

Terakita, A. (2005) Protein family review The opsins. *Genome Biology* 6, 1–9.
Trevisan-Silva, D., Gremski, L.H., Chaim, O.M., da Silveira, R.B., Meissner, G.O., Mangili, O.C., Barbaro, K.C., Gremski, W., Veiga, S.S. & Senff-Ribeiro, A. (2010) Astacin-like metalloproteases are a gene family of toxins present in the venom of different species of the brown spider (genus *Loxosceles*). *Biochimie* 92, 21–32.
Verrill, A.E. (1865) Synopsis of the Polyps and Corals of the North Pacific Exploring Expedition under Commodore C. Ringgold and Captain John Rodgers, U.S.N.). *Communications of the Essex Institute.* 4, 145–152.
Wallimann, T. & Hemmer, W. (1994) Creatine kinase in non-muscle tissues and cells. *Molecular and cellular biochemistry* 133-134, 193–220.
Weill, R. (1934) Contribution a l'etude des cnidaires et de leurs nematocystes II. Valeur taxonomique du cnidome. *Travaux de la Station Zoologique de Wimereux* 11, 351–355; 632–701.
Werner, B. (1973) Spermatozeugmen und Paarungsverhalten bei *Tripedalia cystophora* (Cubomedusae). *Marine Biology* 18, 212–217.
Weston, A.J., Chung, R., Dunlap, W.C., Morandini, A.C., Marques, A.C., Moura-da-Silva, A.M., Ward, M., Padilla, G., da Silva, L.F., Andreakis, N. & Long, P.F. (2013) Proteomic characterisation of toxins isolated from nematocysts of the South Atlantic jellyfish *Olindias sambaquiensis*. *Toxicon : Official journal of the International Society on Toxinology* 71, 11–7.
Wyss, A. (2004) Functions and Actions of Retinoids and Carotenoids : Building on the Vision of James Allen Olson Carotene Oxygenases : A New Family of Double Bond Cleavage Enzymes. *American Society for Nutritional Sciences*, 246–250.
Yamazaki, Y. & Morita, T. (2004) Structure and function of snake venom cysteine-rich secretory proteins. *Toxicon* 44, 227–231.
Yanagihara, A.A., Kuroiwa, J.M.Y., Oliver, L.M., Chung, J.J. & Kunkel, D.D. (2002) Ultrastructure of a novel eurytele nematocyst of *Carybdea alata* Reynaud (Cubozoa, Cnidaria). *Cell and tissue research* 308, 307–18.
Yanagihara, A.A. & Shohet, R. V (2012) Cubozoan venom-induced cardiovascular collapse is caused by hyperkalemia and prevented by zinc gluconate in mice. *PloS one* 7, e51368.
Yang, Y., Xiong, J., Zhou, Z., Huo, F., Miao, W., Ran, C., Liu, Y., Zhang, J., Feng, J., Wang, M., Wang, M., Wang, L. & Yao, B. (2014) The genome of the myxosporean *Thelohanellus kitauei* shows adaptation sto nutrient acquisition within its fish host. *Genome Biology and Evolution* 6, 3182–3198.
Yoshimoto, C.M. & Yanagihara, A.A. (2002) Cnidarian (coelenterate) application envenomations in Hawai ' i improve following heat. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 96, 300–303.
Zapata, F., Goetz, F.E., Smith, S. a., Howison, M., Siebert, S., Church, S.H., Sanders, S.M., Ames, C.L., McFadden, C.S., France, S.C., Daly, M., Collins, A.G., Haddock, S.H.D., Dunn, C.W. & Cartwright, P. (2015) Phylogenomic Analyses Support Traditional Relationships within Cnidaria. *Plos One* 10, e0139068.
Zenkert, C., Takahashi, T., Diesner, M.-O. & Özbek, S. (2011) Morphological and Molecular Analysis of the *Nematostella vectensis* Cnidom P. K. Dearden (Ed). *PLoS ONE* 6, e22725.
Zhang, M., Fishman, Y., Sher, D. & Zlotkin, E. (2003) Hydralysin, a novel animal group-

selective paralytic and cytolytic protein from a noncnidocystic origin in *Hydra*. *Biochemistry* 42, 8939–44.