Integrating genomic, transcriptomic and developmental approaches to investigate coloniality and life cycle evolution in the Hydractiniidae (Hydrozoa: Cnidaria)

BY

STEVEN M. SANDERS

Submitted to the graduate degree program in Ecology and Evolutionary Biology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Chairperson Paulyn Cartwright

_____

Justin Blumensteil

_____

Mark Holder

_____

Kirsten Jensen

_____

Erik Lundquist

Date Defended: December 16, 2014

The Dissertation Committee for STEVEN M. SANDERS certifies that this is the approved version of the following dissertation:

Integrating genomic, transcriptomic and developmental approaches to investigate coloniality and life cycle evolution in the Hydractiniidae (Hydrozoa: Cnidaria)

_____

Chairperson Paulyn Cartwright

Date approved: December 16, 2014

ABSTRACT

Integrative approaches to evolutionary biology yield rich data through which we can truly begin to understand the marvels of life. This dissertation integrates genomic, transcriptomic, and developmental approaches to understand the evolution of prominent life history characters of the cnidarian class Hydrozoa, including the transition from solitary to colonial forms, an elaboration of coloniality known as polyp polymorphism, and medusae (jellyfish) evolution and loss. While these characters have been repeatedly explored phylogenetically, recognizing interesting and complex evolutionary patterns of character transitions, understanding of these complex patterns of character evolution will ultimately come from insight into their development. In this dissertation, I have developed workflows for analyzing RNA-Seq data in both an intra- and interspecific comparative context. Using next-generation sequencing I not only characterize entire transcriptomic expression profiles in various tissues of two hydractiniid hydrozoans, *Hydractinia symbiolongicarpus* and *Podocoryna carnea*, but also assess and accurately characterize intra- and interspecific changes in gene expression. Using these unbiased differential expression analyses, I identify correlated changes in expression and propose candidate genes and gene pathways that are potentially involved in these key transitions. Furthermore, using whole mount *in situ* hybridization to characterize the spatial expression of various candidates genes, I validated each approach showing expression consistent with their role in the development of a particular tissue or life cycle stage. Results presented in this dissertation suggest that the differential regulation of gene expression, as well as novel gene gain and loss appear to have played an important role in hydrozoan life cycle transitions. Moreover, these results reveal the power of these unbiased genomic/transcriptomic methods over traditional comparative candidate gene approaches to address longstanding questions of hydrozoan morphology and evolution.

ACKNOWLEDGEMENTS

There are many people that I need to thank for their help and support throughout this process. I'd first and foremost like to thank my wife, Jessica. Jessica you have been my rock throughout these last four and a half years. None of this would have happened without your constant love and support. Thank you for believing in me enough to move to Kansas, leaving behind all of our friends and family (and the ocean), so that I could follow my dreams. Even more, you have given me two wonderful boys, Austin and Owen, who also deserve special thanks for teaching me the true meaning of time management. I'd also like to thank my parents, who have always been there for me. From a very early age, you instilled in me a sense determination and drive. You constantly pushed me to strive for more and chase after my goals. You all have made my life so full and have been my motivation to work hard and finish this degree.

I'd also like to thank current/past members of the Cartwright lab: Annalise Nawrocki, Bastian Bentlage, Steve Davis, Sally Chang, and Mariya Shcheglovitova. All of you have contributed in different ways to completion of this dissertation and made the Cartwright lab a fun and productive environment to work in. Similarly, I'd like to thank my doctoral committee: Justin Blumentsteil, Mark Holder, Kirsten Jensen, and Erik Lundquist. Thank you all for your guidance and insight. There are also a number of collaborators and researchers, here at KU and elsewhere, that deserve recognition for their aid in the completion of this dissertation: Clark Bloomer, Uri Frank, Christine Schnitzler, Andy Baxevanis, Jim Coffman, Peter Schuchert, Casey Dunn, Felipe Zapata, Matt Nicotra, Leo Buss, Neil Blackstone, Lina Hileman, Jaime Walters, and Ed Wiley.

Lastly, I'd like to thank my advisor, Paulyn. While it was maybe painfully obvious how unqualified/under-skilled I was for a position in the lab during our first interview, you were willing to take a chance on me. Throughout this whole process you have constantly challenged and pushed me to be a better scientist/researcher, never accepting less than my best. Thank you for your continued guidance and the knowledge you have passed onto me. It has been a privilege to work with you over these years and I look forward to many years of future collaborations.

TABLE OF CONTENTS

INTRODUCTION

Integrative approaches to evolutionary biology yield rich data through which we can truly begin to understand the marvels of life. This dissertation integrates genomic, transcriptomic, and developmental approaches to understand the evolution of prominent life history characters of the cnidarian class Hydrozoa. These include the transition from solitary to colonial forms, an elaboration of coloniality known as polyp polymorphism, and medusae (jellyfish) evolution and loss. Evolutionary patterns of these characters have been repeatedly explored phylogenetically [1-10]. While phylogenies are important for recognizing evolutionary patterns of character transitions, understanding the genetic processes underlying these complex patterns of character evolution will ultimately come from insight into their development. As a means to assess and accurately characterize changes in gene expression correlated with the evolution and development of the above-mentioned features, I use next-generation sequencing to characterize entire transcriptomic expression profiles in tissues of two hydractiniid hydrozoans, *Hydractinia symbiolongicarpus* and *Podocoryna carnea*.

Chapter 1 - Polyp polymorphism

In **Chapter 1** of this dissertation, I use short read Illumina data to identify key genes and gene pathways potentially involved in the evolution and development of an elaboration of hydrozoan coloniality, known as polyp polymorphism. Colonial hydrozoans are composed of individual polyps connected through continuous epithelia and a shared gastrovascular cavity. Despite their simple epithelial construction, many hydrozoan species evolved complex colonies through functional specialization of genetically identical yet morphologically distinct polyp types, conferring a division of labor within the colony [11-13]. The main labor divisions are

between feeding, reproduction, and defense, where specialized polyp types are morphologically distinct, reflecting their particular functions. *Hydractinia symbiolongicarpus* has four different polyp types: feeding polyps (called gastrozooids), reproductive polyps (gonozooids), defensive and food gathering polyps (dactylozooids), and less common defensive polyps (tentaculozooids). It has long been hypothesized that these specialized polyps arose through evolutionary alterations in oral-aboral patterning in the ancestral gastrozooid [14-16]. While previous studies using candidate gene approaches have identified patterning genes specific to different polyp types [17,18], in this chapter I use RNA-Seq in an unbiased survey of genes differentially expressed (DE) between three of the four polyp types of *H. symbiolongicarpus*. This chapter has been published in the journal *BMC Genomics* (coauthors Mariya Shcheglovitova and Paulyn Cartwright)[19] and is formatted according to journal guidelines.

Chapters 2 & 3 - Medusae evolution

Hydrozoans exhibit complex life cycles, alternating between a benthic asexually reproducing polyp stage and a pelagic sexually reproducing medusa stage. In most hydrozoan species however, the medusa life cycle stage is reduced [20] and sexual maturity is reached in a gonophore that resembles an earlier ontogenetic stage of medusae development. Gonophore development can range from completely reduced structures that lack any resemblance to medusa called a sporosac, to more developed forms called medusoids, that may or may not detach and swim, but lack the ability to feed, to the fully developed medusa stage that detaches from the hydroid polyp and can feed, swim, and sexually reproduce in the water column. Several phylogenetic studies have recovered strong support for repeated independent losses of medusae [1,6-8,10], and possible re-gain [7,8,10]. Previous gene expression studies have reported

expression patterns of canonical Wnt pathway components consistent with their role in medusae axial patterning [21-24]. Expression patterns from these studies suggest that down regulation of Wnt pathway elements may be involved in the arrest of the medusae developmental program in those species that lack medusae [23], although little is known of Wnt signaling in species with a fully developed pelagic medusa stage.

In **Chapter 2** of this dissertation, I use an intraspecific approach to RNA-Seq (similar to **Chapter 1**) to investigate the evolution and development of medusae. In this chapter I assemble, annotate, and assess differential expression of the transcriptome of *Podocoryna carnea* using RNA-Seq data collected from three different life cycle stages. These results, in conjunction with previously published studies in other hydrozoans, suggest that changes in the regulation of the canonical Wnt signaling pathway may be involved in the evolution and development of hydrozoan medusae and their reduced forms. This chapter is formatted for publication in the journal *EvoDevo* where it is currently under review.

Chapter 3 of this dissertation presents an unbiased, interspecific RNA-Seq workflow for assessing differential expression of orthologous genes between *H. symbiolongicarpus* and *P. carnea*. Similar to **Chapter 2**, the goal of this chapter is to reveal genes and gene pathways involved in medusae evolution. Where these two chapters differ is in the scope of the analyses. While **Chapter 2** use RNA-Seq to quickly capture and characterize the expression of a single signaling pathway within life cycles stages of *P. carnea*, **Chapter 3** takes an unbiased approach to characterize interspecific differential expression of orthologous genes and identify potential gene gain and loss associated with phenotypic differences, namely between the sporosac and medusa in *H. symbiolongicarpus* and *P. carnea*, respectively. This chapter is formatted for publication in the journal *Molecular Biology and Evolution* where it has been submitted.

Chapter 4 - Coloniality

Chapter 4 of this dissertation revisits hydrozoan coloniality. In Hydrozoa, species are either solitary or colonial. These two life history strategies characterize the two major lineages of Hydrozoa: Trachylina (mostly solitary) and Hydroidolina (mostly colonial) [8]. Colonial hydrozoan species are composed of individual polyps connected by a tube-like structure called a stolon. In the typical colonial hydrozoan life cycle, stolons will elongate, branch, and give rise to new polyps asexually, forming a reticulate network of polyps connected to one another by continuous epithelia and a shared gastrovascular cavity. Previous phylogenetic studies have recovered topologies with colonial lineages as paraphyletic [25,26] and character evolution studies since have recovered coloniality evolving at the base of the Hydroidolina followed by several independent loses [8]. A recent phylogenomic study recovered all colonial hydrozoans as monophyletic [27]. Moreover, analyses in **Chapter 2** revealed that the membrane bound Wnt receptor, *frizzled3*, is specific to colonial hydrozoans, suggesting a potential role in signaling colonial patterning. In this chapter, I report a new molecular phylogeny of membrane bound *frizzled* genes from 22 cnidarian transcriptomes and genomes. Furthermore, I report spatial expression of *frizzled3* in colony-specific tissues in *H. symbiolongicarpus* and *P. carnea*.

REFERENCES

1. Cunningham CW, Buss LW: **Molecular evidence for multiple episodes of paedomorphosis in the family Hydractiniidae.** *Biochem. Sys. Ecol.* 1993, **21**:57-69.
2. Marques AC, Migotto AE: **Cladistic analysis and new classification of the family Tubulariidae (Hydrozoa, Anthomedusae).** *Pap. Avuls. Zool.* 2001, **41**:465–88.
3. Collins AG: **Phylogeny of Medusozoa and the evolution of cnidarian life cycles.** *J. Evol. Biol.* 2002, **15**:418-432.
4. Marques AC, Collins AG: **Cladistic analysis and new classification of the family Tubulariidae (Hydrozoa, Anthomedusae).** *Invert. Biol.* 2004, **123**:23-42.
5. Collins AG, Schuchert P, Marques AC, Jankowski T, Medina M, Schierwater B: **Medusozoan phylogeny and character evolution clarified by new large and small subunit rDNA data and an assessment of the utility of phylogenetic mixture models.** *Sys. Biol.* 2006, **55**:97-115.
6. Leclère L, Schuchert P, Manuel M: **Phylogeny of the Plumularioidea (Hydrozoa, Leptothecata): evolution of colonial organisation and life cycle.** *Zool. Scripta* 2007, **36**:371-394.
7. Leclère L, Schuchert P, Cruaud C, Couloux A, Manuel M: **Molecular phylogenetics of Thecata (Hydrozoa, Cnidaria) reveals long-term maintenance of life history traits despite high frequency of recent character changes.** *Sys. Biol.* 2009, **58**:509-526.
8. Cartwright P, Nawrocki AM: **Character evolution in Hydrozoa (phylum Cnidaria).** *J. Integ. Comp. Biol.* 2010, **50**:456-472.
9. Nawrocki AM, Cartwright P: **A novel mode of colony formation in a hydrozoan through fusion of sexually generated individuals.** *Curr. Biol.* 2012, **22**:1-5.

10. Miglietta MP, Cunningham CW: **Evolution of life cycle, colony morphology, and host-specificity in the Family Hydractiniidae (Hydrozoa, Cnidaria).** *Evolution* 2012, **doi**:10.1111/j.1558-5646.2012.01717.x.

11. Muller W: **Experimentelle Untersuchungen Über Stockentwicklung, Polypendifferenzierung, und sexualchimären bei *Hydractinia echinata*.** *Wilhelm Roux' Arch Entwicklungsmech. Org.* 1964, **155**:181-268.

12. Berrill NJ: **Growth and form in gymnoblastic hydroids; polymorphism within the Hydractiniidae.** *J. Morph.* 1953, **92**:241-272.

13. Burnett AL, Sindelar W, Diehl N: **An examination of polymorphism in the hydroid *Hydractinia echinata*.** *J. Mar. Biol. Assoc. U.K.* 1967, **47**:645-658.

14. Agassiz A: **North American Acalephae.** (University Press: Welsh, Bigelow, Cambridge, U.K.). 1865.

15. Haeckel E: **Report on the Siphonophorae Collected by H. M. S. Challenger During the Years 1873–76.** (The Challenger Reports, London). 1888.

16. Huxley TH: **The Oceanic Hydrozoa: A Description of the Calycophoridae and Physophoridae Observed During the Voyage of the H. M. S. "Rattlesnake" in the years 1846–1850: With a General Introduction.** (Ray Society, London). 1859.

17. Cartwright P, Bowsher J, Buss LW: **Expression of the HOX type gene, Cnox-2, and the division of labor in a colonial hydroid.** *PNAS* 1999, **96**:2183-2186.

18. Mokady O, Dick MH, Lackschewitz D, Schierwater B, Buss LW: **Over one-half billion years of head conservation? Expression of an *ems* class gene in *Hydractinia symbiolongicarpus* (Cnidaria: Hydrozoa).** *PNAS* 1998, **95**:3673-3678.

19. Sanders SM, Shcheglovitova M, Cartwright P: **Differential gene expression between functionally specialized polyps of the colonial hydrozoan *Hydractinia symbiolongicarpus* (Phylum Cnidaria).** *BMC Genomics* 2014, **15**:406.

20. Gibbons MJ, Janson LA, Ismail A, Samaai T: **Life cycle strategy, species richness and distribution in marine Hydrozoa (Cnidaria: Medusozoa).** *J. Biogeogr.* 2010, **37**:441-448.

21. Müller W, Frank U, Teo R, Mokady O, Guette C, Plickert G: **Wnt signaling in hydroid development: ectopic heads and giant buds induced by GSK-3*ß* inhibitors.** *Int. J. Dev. Biol.* 2007, **51**:211-220.

22. Duffy DJ, Plickert G, Kuenzel T, Tilmann W, Frank U: **Wnt signaling promotes oral but suppresses aboral structures in *Hydractinia* metamorphosis and regeneration.** *Dev.* 2010, **137**:3057-3066.

23. Nawrocki AM, Cartwright P: **Expression of Wnt pathway genes in polyps and medusa-like structures of *Ectopleura larynx* (Cnidaria: Hydrozoa).** *Evol. & Dev.* 2013, **15**:373-384.

24. Momose T, Houliston E: **Two oppositely localized frizzled RNAs as axis determinants in cnidarian embryo.** *PLoS Biol.* 2007, **5**:e70.

25. Collins AG, Winkelmann S, Hadrys H, Schierwater B: **Phylogeny of Capitata and Corynidae (Cnidaria, Hydrozoa) in light of mitochondrial 16S rDNA data.** *Zool. Scripta* 2005, **34**:91-99.

26. Cartwright P, Evans NM, Dunn CW, Marques AC, Miglietta MP, Schuchert P, Collins AG: **Phylogenetics of Hydroidolina (Hydrozoa: Cnidaria).** *J. Mar. Biol. Assoc.* 2008, **88**:1163-1172.

27. Zapata F, Goetz F, Smith S, Church S, Sanders SM, Ames CL, McFadden CS, France S, Daly M, Collins AG, Haddock SHD, Dunn CW, Cartwright P: **Phylogenomic analyses of Cnidaria.** *In preparation*.

CHAPTER 1:

Differential gene expression between functionally specialized polyps of the colonial hydrozoan *Hydractinia symbiolongicarpus* (Phylum Cnidaria)

Abstract

Background

A colony of the hydrozoan *Hydractinia symbiolongicarpus* comprises genetically identical yet morphologically distinct and functionally specialized polyp types. The main labor divisions are between feeding, reproduction, and defense. In *H. symbiolongicarpus*, the feeding polyp (called a gastrozooid) has elongated tentacles and a mouth, which are absent in the reproductive polyp (gonozooid) and defensive polyp (dactylozooid). Instead, the dactylozooid has an extended body column with an abundance of stinging cells (nematocysts) and the gonozooid bears gonophores on its body column. Morphological differences between polyp types can be attributed to simple changes in their axial patterning during development, and it has long been hypothesized that these specialized polyps arose through evolutionary alterations in oral-aboral patterning of the ancestral gastrozooid.

Results

An assembly of 66,508 transcripts (>200bp) were generated using short-read Illumina RNA-Seq libraries constructed from feeding, reproductive, and defensive polyps of *H. symbiolongicarpus*. Using several different annotation methods, approximately 54% of the transcripts were annotated. Differential expression analyses were conducted between these three polyp types to isolate genes that may be involved in functional, histological, and patterning differences between polyp types. Nearly 7K transcripts were differentially expressed in a polyp-significant manner, including members of the homeodomain, myosin, toxin, and BMP gene families. We report the spatial expression of a subset of these polyp-significant transcripts to validate our differential expression analyses.

Conclusions

While potentially originating through simple changes in patterning, polymorphic polyps in *Hydractinia* are the result of differentially expressed functional, structural, and patterning genes. The differentially expressed genes identified in our study provide a starting point for future investigations of the developmental patterning and functional differences that are displayed in the different polyp types that confer a division of labor within a colony of *H. symbiolongicarpus*.

Keywords

Hydractinia symbiolongicarpus, RNA-Seq, polymorphism, differential expression, transcriptome assembly, annotation

Background

Colonial hydrozoans are composed of individual polyps connected through continuous epithelia and a shared gastrovascular cavity. Hydrozoans are members of the phylum Cnidaria, which are characterized by their diploblastic construction, comprising only two epithelial layers, the epidermis and gastrodermis. Despite their simple epithelial construction, many hydrozoan species evolved complex colonies through functional specialization of genetically identical yet morphologically distinct polyp types, conferring a division of labor within the colony [1-3]. This division of labor is known as polyp polymorphism [1-3].

The main labor divisions are between feeding, reproduction, and defense, where specialized polyp types are morphologically distinct, reflecting their particular functions. *Hydractinia symbiolongicarpus* has four different polyp types (Fig. 1). The feeding polyp (called a gastrozooid) has a mouth and tentacles, which are absent in the reproductive polyp (gonozooid), defensive and food gathering polyp (dactylozooid), and the less common defensive polyp (tentaculozooid, not shown). The dactylozooid has an elongated body column with an abundance of epithelial muscular cells and nematocytes (stinging cells). The gonozooid bears gonophores, which house the gametes. The gonozooid and dactylozooid are similar in their distal ends, with clusters of nematocysts and lacking a functional mouth and elongate tentacles. The tentaculozooid resembles a single tentacle of the gastrozooid, but is the size of an individual polyp.

It has long been hypothesized that these specialized polyps arose through evolutionary alterations in oral-aboral patterning in the ancestral gastrozooid [4-6]. Previous studies using candidate gene approaches have identified patterning genes specific to different polyp types. Cartwright et al. [7] focused on the involvement of *Cnox-2*, a parahox gene, in patterning these

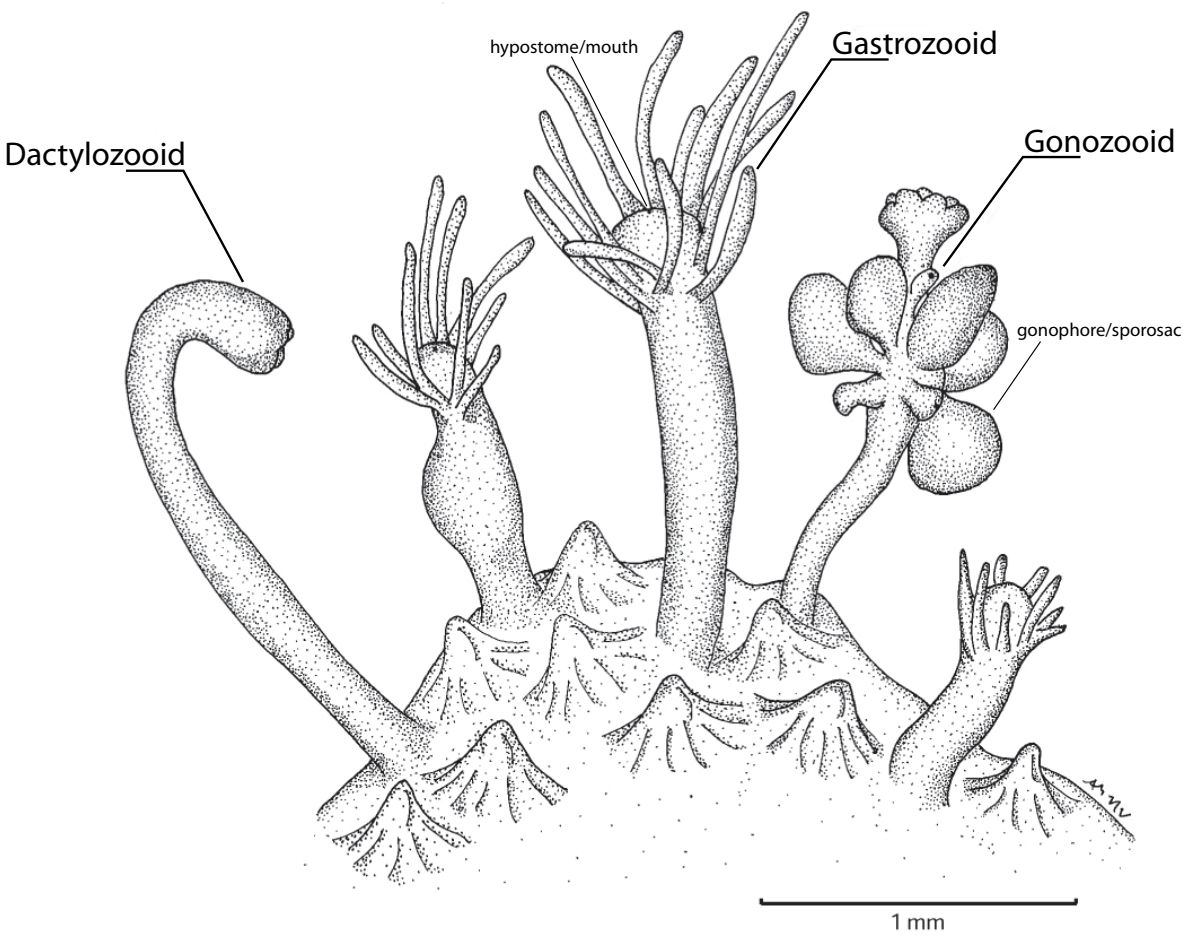


Figure 1 – Colony of *Hydractinia symbiolongicarpus.*

Illustration showing the different polymorphic polyps that comprise *H. symbiolongicarpus* colonies (tentaculozooids not shown)**.** Modified from Cartwright & Nawrocki [79].

different polyp types of *H. symbiolongicarpus*. Immunolocalization of the *Cnox-2* protein showed expression in body column tissue and down-regulation in oral structures of the gastrozooid. Mokady et al. [8] compared expression of *Cn-ems* (*empty spiracles* homolog) between gastrozooids and gonozooids of *H. symbiolongicarpus*. Whole mount *in situ* hybridization revealed no expression of *Cn-ems* in the gonozooid, while mRNAs were detected in the gastrodermal epithelia ("digestive cells") of the gastrozooid.

More recently, Siebert et al. [9] used an RNA-seq approach to examine differential expression between several polyp types of another hydrozoan, the siphonophore *Nanomia bijuga*. Although the focus of their study was to evaluate next generation sequencing (NGS) platforms for differential expression (DE), they confirmed, through whole mount *in situ* hybridization, that at least one gene identified through their DE analyses (*isogroup03256*) was expressed in a polyp-specific manner.

With the advent of NGS technologies, an unbiased approach to identify genes involved in the differentiation of different tissues (e.g. [10]) and developmental stages (e.g. [11]), or those that are differentially expressed between species (e.g. [12]) can be made without reference to particular candidate genes. We report a transcriptome assembly, annotations, and DE analyses between three different polyp types in *H. symbiolongicarpus.* Our results, validated with whole mount *in situ* hybridization, confirmed that DE analyses using RNA-Seq is a powerful approach for identifying genes and pathways involved in conferring a division of labor within this colonial organism.

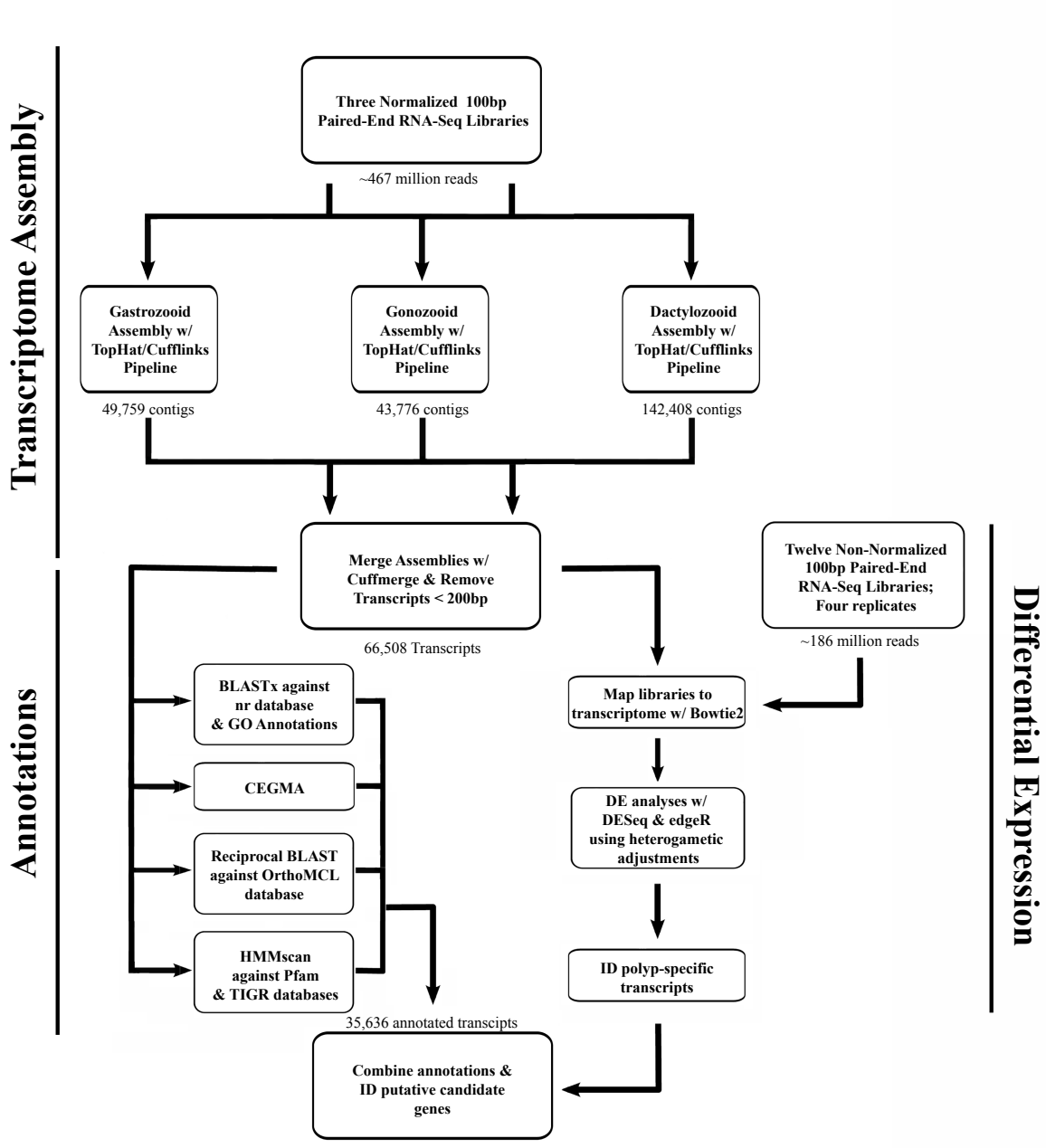

Transcriptome Assembly
Three Normalized 100bp Paired-End RNA-Seq Libraries
~467 million reads
Gastrozooid Assembly w/ TopHat/Cufflinks Pipeline
49,759 contigs
Gonozooid Assembly w/ TopHat/Cufflinks Pipeline
43,776 contigs
Dactylozooid Assembly w/ TopHat/Cufflinks Pipeline
142,408 contigs
Merge Assemblies w/ Cuffmerge & Remove Transcripts < 200bp
66,508 Transcripts
Annotations
BLASTx against nr database & GO Annotations
CEGMA
Reciprocal BLAST against OrthoMCL database
HMMscan against Pfam & TIGR databases
35,636 annotated transcipts
Combine annotations & ID putative candidate genes
Twelve Non-Normalized 100bp Paired-End RNA-Seq Libraries; Four replicates
~186 million reads
Map libraries to transcriptome w/ Bowtie2
DE analyses w/ DESeq & edgeR using heterogametic adjustments
ID polyp-specific transcripts
Differential Expression

Figure 2 – Workflow of transcriptome assembly through annotation and differential expression analyses.

Raw reads from three normalized libraries were filtered based on quality score and separately mapped to unpublished genomic scaffolds of *H. symbiolongicarpus* using TopHat 2.0.6, assembled using Cufflinks 2.1.1, merged into a single assembly using cuffmerge, and filtered by transcript size, removing assembled transcripts less than 200bp in length. Blast2GO, CEGMA, HMMscan, and orthoMCL were used to annotate the transcriptome. Differential expression began with mapping 12 non-normalized libraries to the final transcriptome assembly with Bowtie2. DE was then assessed with DESeq and edgeR, and polyp-significant DEs were compared to the annotated transcriptome.

Results and Discussion

Transcriptome Assembly and Annotation

From the three normalized libraries, 49,759, 43,776, and 142,408 contigs were assembled for the gastrozooid, gonozooid, and dactylozooid, respectively. Individual transcriptomes were merged into a single assembly of 101,518 unique transcripts using cuffmerge (Fig. 2). Cuffmerge merges novel and common transcripts into a single assembly and removes artifact constructions, improving the overall quality of the assembly. This step allows for easy annotation and differential expression analyses of a single assembly, without concerns regarding orthology assignments between multiple assemblies. After filtering for transcripts less than 200bp in length, our final assembly consisted of 66,508 transcripts, with an N50 of 1,451bp (Additional file 1).

Approximately 54% of the transcriptome (35,636 transcripts) was annotated using Blast2GO, CEGMA, orthoMCL, and HMMscan (Additional file 2), with these transcripts showing significant similarity to sequences in at least one database in our annotation pipeline (Fig. 2,3). These include 416 (91%) of the "core" and 238 (96%) of the "ultra-conserved" eukaryotic genes identified using CEGMA (Additional file 3). Figure 3 shows the number of transcripts annotated by one or more of the annotation methods.

Table 1. Number of DE transcripts in different pairwise comparisons of the libraries.

	Full Dataset		Adjusted Dataset	
	DESeq	**edgeR**	**DESeq**	**edgeR**
Dact vs Gast	662	2,498	2,062	4,230
Dact vs Gono	2,312	16,879	10,341	18,899
Gast vs Gono	4,245	16,889	11,908	18,744
Male vs Female	11,798	12,886		

Full Dataset corresponds to the number of transcripts recovered from DE analyses. Adjusted dataset refers to counts following heterogametic adjustments.

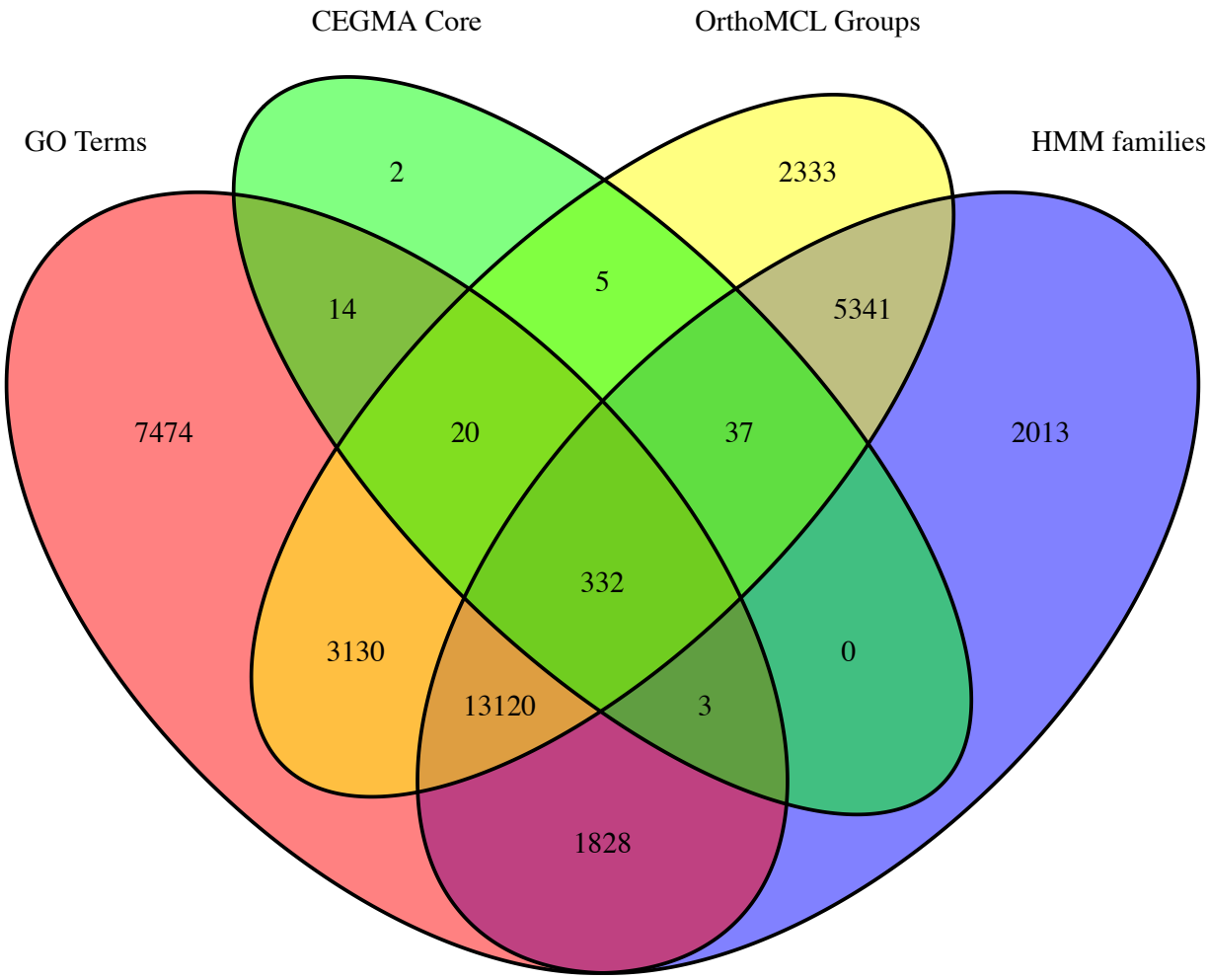


Figure 3. Venn diagram displaying the number of transcripts annotated by each method. Gene Ontology terms were added with Blast2GO using the BLASTX algorithm against NCBI's nr protein database and a threshold of 1 x 10^{-03}. A set of conserved eukaryotic genes was identified with CEGMA. HMM protein families from the PFAM and TIGR databases were assigned to the amino acid translation of the most likely reading frame for each transcript (identified using an open reading frame prediction tool) using HMMscan under default settings. HMMscan annotations were constrained to a significance threshold of 0.01. Orthogroups were assigned to the same amino acid translations using the orthoMCL web server.

Differential Expression Analyses

Statistically significant differences in expression between different polyp types were detected using two DE packages, DESeq and edgeR. Figure 4C, D shows the effect of heterogametic adjustments on the euclidean distances (sum of the pairwise distance across all transcripts) between libraries. Both DESeq and edgeR reveal that dactylozooids and gastrozooids share the fewest number of DE transcripts and the smallest change in the number of DEs recovered after the heterogametic adjustments (Table 1), while DE analyses including gonozooids show a much larger increase in the number of DE transcripts after those adjustments. This large increase can be explained by the huge amount of variability found when ignoring the sexual differences between gametic tissue in gonozooid samples. The DE analysis between the male and female gonozooid libraries identified 11,798 (DESeq) and 12,886 (edgeR) transcripts significantly up- or down-regulated (Table 1, Fig. 4B). Removal of all male/female DE transcripts clusters gonozooid samples by polyp type rather than sex and increases the distance between gastrozooids and dactylozooids (Fig. 4D), while treating male and female gonozooid libraries as different conditions reduces the average dispersion estimate for each transcript, essentially increasing the power of the DE analyses (Fig. 4E and 4F). Yet, even after the heterogametic adjustments, gonozooids still have the largest number of polyp-significant transcripts (transcripts that are strictly up- or down-regulated in a particular polyp type when compared to other polyps; Fig. 5, Table 2, Additional file 4).

Our DE analyses revealed several polyp-significant genes that are consistent with previous studies using candidate gene approaches in cnidarians as discussed below. Furthermore, our analyses revealed additional genes that were not previously considered to play specific developmental, functional and/or structural roles in cnidarians. Below we summarize of few of

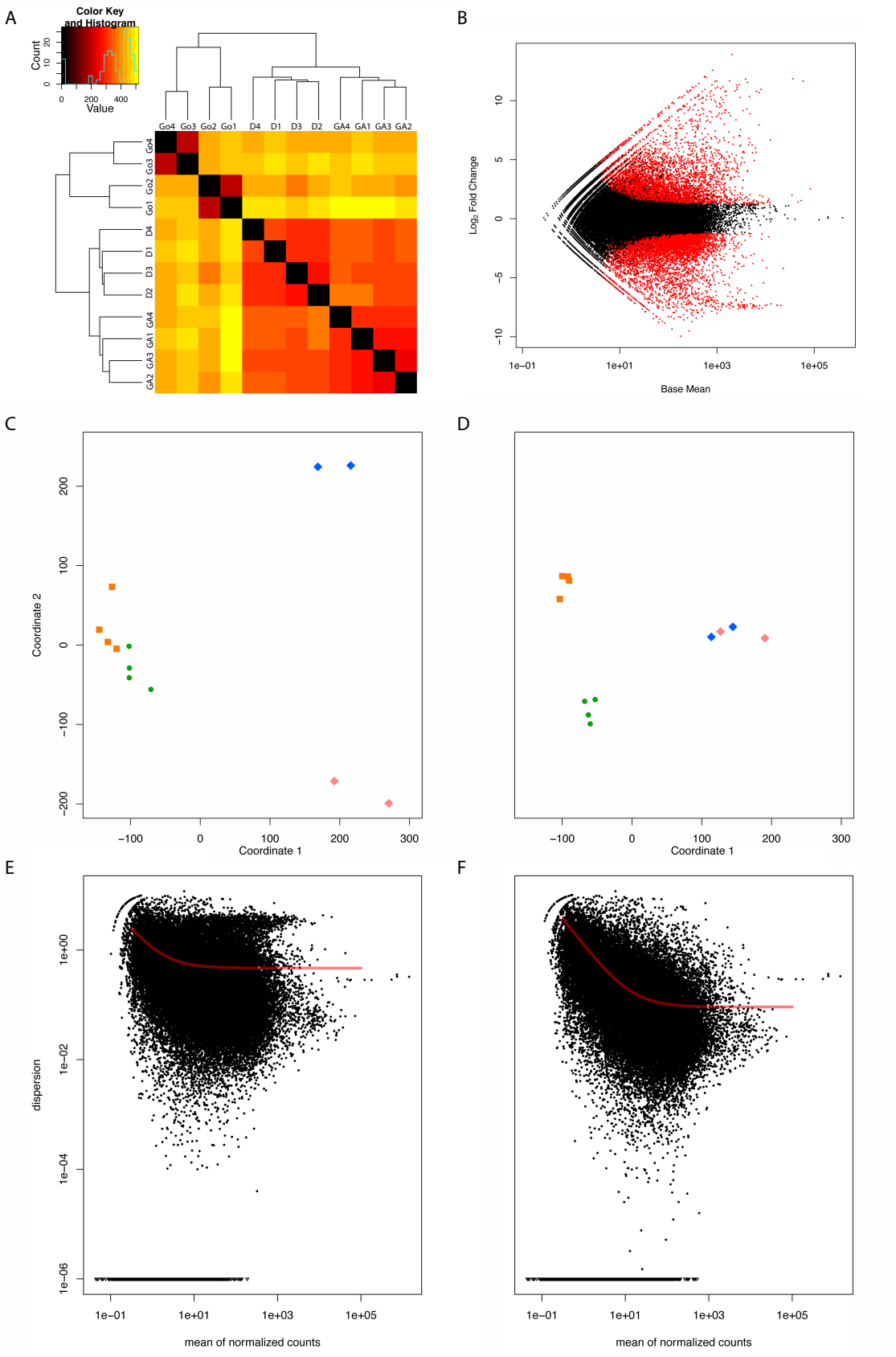

A
Color Key and Histogram
Count
Value
Go4 Go3 Go2 Go1 D4 D1 D3 D2 GA4 GA1 GA3 GA2
B
Log2 Fold Change
Base Mean
C
Coordinate 2
Coordinate 1
D
Coordinate 1
E
dispersion
mean of normalized counts
F
mean of normalized counts

Figure 4. Effects of heterogametic expression on library distances.

A. Heatmap of the Euclidean distances between all twelve libraries prior to heterogametic adjustments. Samples Go1 and Go2 correspond to female gonozooid libraries, while Go3 and Go4 correspond to male gonozooid libraries. B. MA plot of the DE analysis between male and female gonozooid libraries in DESeq. Red dots indicate statistically significant DE transcripts. $Log_2FoldChange > 0$ corresponds to expression levels higher in the male gonozooid libraries, and $Log_2FoldChange < 0$ corresponds to expression levels higher in the female gonozooid libraries. C. Euclidean distances plotted in two dimensions prior to heterogametic adjustments. (Legend: ■-gastrozooids; ●-dactylozooids; ◆-male gonozooids; ◆-female gonozooids). D. Euclidean distances plotted in two dimensions after all statistically significant heterogametic transcripts are removed. E. Plot of the estimated dispersion values against the mean of normalized counts of each transcript when binning both male and female gonozooid libraries in a single condition. Fitted dispersion values indicated by the red line. F. Plot of the estimated dispersion values against the mean of normalized counts of each transcript when male and female gonozooid libraries treated as separate conditions.

these results and suggest areas of interest for further study.

Table 2. Number of transcripts always up- or down-regulated in a specific polyp.

padj < .05

	DESeq		edgeR		BOTH	
	UP	**DOWN**	**UP**	**DOWN**	**UP**	**DOWN**
Gastrozooid	1,067	40	1,934	148	955	31
Gonozooid	3,505	3,405	11,304	1,851	3,491	1,562
Dactlyozooid	444	29	999	180	332	20

Transcripts that have support as polyp-significant (must be significant in only two of the three pairwise comparisons) by both DESeq and edgeR are what we refer to as polyp-significant in the text.

Gametogenic expression. While adjusting for differences in gene expression between males and females greatly reduced the effect of heterogametic expression on the DE analyses, genes likely involved in non-sex specific gametogenesis were found up-regulated in the gonozooids. Of the 76 polyp-significant DE transcripts annotated with functional terms that include mitosis, cell cycle, and germline maintenance (Additional file 5), 69 are up-regulated in the gonozooids, including four DE transcripts annotated as known hydrozoan stem cell markers including *nanos* [13]*, vasa* [14], and *piwi* [15]. This is consistent with expression studies of *nanos* and *vasa* genes in a closely related species, *Hydractinia echinata* [13,14].

Homeobox genes. Several homeobox transcripts are differentially expressed between different polyp types (Additional file 5). Homeobox genes up-regulated in the gastrozooid include members of the LIM (*lhx*), sine oculus (*six*), empty spiracles (*ems*), and PRD classes, confirmed by molecular phylogenetic analysis of cnidarian homeodomains (Additional files 5,6). The up-regulation of the empty spiracles homolog (100% bootstrap [BS] support; Additional files 5,6), *Cn-ems*, is consistent with the findings of Mokady et al. [8] discussed previously. Up-regulated gastrozooid expression of two *lhx*-like transcripts, one *six*-like and one *orthopedia* (PRD class) transcript (100%, 99%, and 98% BS support, respectively; Additional files 5,6) is

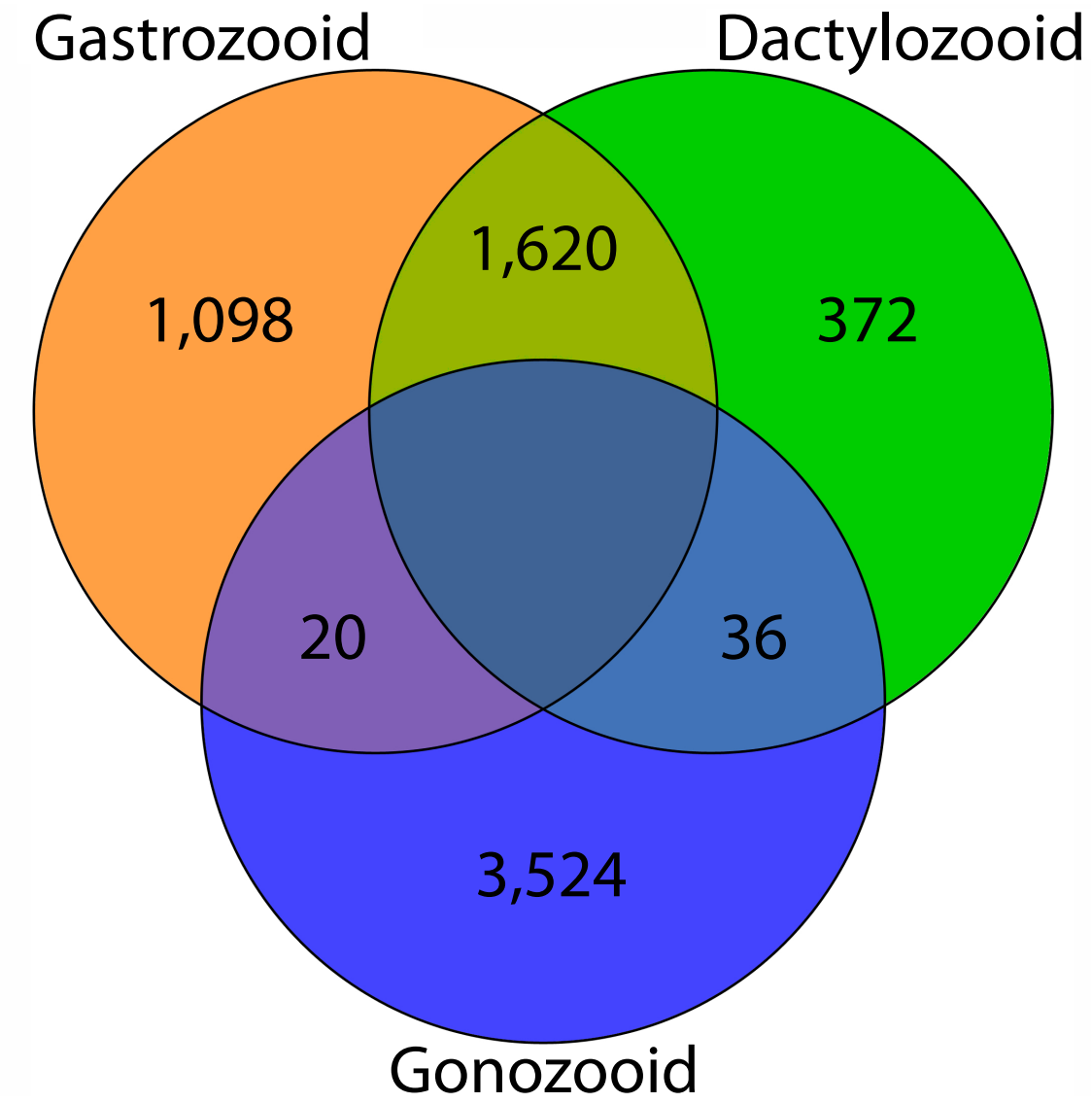


Figure 5 – Venn diagram showing numbers of polyp-significant transcripts.

Transcripts significantly up- or down-regulated ($p_{adj} < 0.05$) in a particular polyp when compared to either of the other two polyp types from both edgeR and DESeq were considered polyp-significant. The intersection of each circle is the number of transcripts down-regulated in the polyp type excluded from that intersection (e.g. 20 transcripts are down-regulated in the dactylozooids). Down-regulation of a transcript in a particular polyp type equates to equivalent up-regulated expression in the other two polyp types.

also consistent with expression studies in other cnidarians, including the scyphozoan *Aurelia* [16], the anthozoan *Nematostella* [17,18, 19], and the hydrozoans *Craspedacusta* [20], *Cladonema* [21], and *Podocoryna* [21], where their expression was found in regions specific to feeding and/or digestion, including tentacles and gastric tissue.

One of the homeodomain-containing transcripts up-regulated in the gonozooid belongs to the POU class (Additional files 5,6). Expression of POU homeodomain transcription factors has also been categorized in other cnidarians, including *Aurelia* [22] and *H. echinata* [23]. In *H. echinata*, the POU gene, *pln*, is expressed around interstitial stem cell (i-cells) [23]. The *H. symbiolongicarpus* ortholog to *pln* (100% BS support; Additional files 5,6) is up-regulated in the gonozooid, which is consistent with that of the other stem cell markers mentioned previously.

Myosins. Myosin genes are a superfamily of molecular motor proteins, primarily associated with muscular contraction and cell movement. Here we find a complex pattern of differential expression of several different myosin transcripts up-regulated in each polyp type (four, six, and four unique transcripts in the gastrozooids, gonozooids, and dactylozooids, respectively), spanning several myosin classes (Additional file 5,7). Of particular note is the up-regulation of a tropomyosin transcript in the gonozooids. In the hydrozoan *Podocoryna carnea*, a tropomyosin, *tpm2*, is expressed solely in the striated muscle of the developing and adult medusa life cycle stages and not in the polyp [24], as opposed to *tpm1*, which is ubiquitously expressed in both polyp and medusae stages [25]. In *Hydractinia*, gonophore development is greatly truncated and never reaches the medusae stage. Instead they form sporosacs, which are believed to lack all medusae like features, including striated muscle necessary for medusae to swim [26-28]. Phylogenetic analysis of cnidarian myosins did not recover any well-supported orthologous relationship between this polyp-significant tropomyosin and other known cnidarian

tropomyosins, although orthology assignments of several other polyp-significant myosins were revealed (Additional file 7). Further discovery of tropomyosin genes in additional cnidarian taxa are necessary to determine if different tropomyosin orthologs are specific to certain medusae features and/or reduced developmental forms.

Toxins. While research into the characterization and properties of cnidarian toxins is on the rise, very little is known of their function and location of endogenous expression [29]. We identified 13 DE transcripts annotated as some type of toxin (three up-regulated in the gastrozooids, seven in dactylozooids, and one in gonozooids; two down-regulated in gonozooids; Additional file 5). Phylogenetic analysis of cnidarian toxins recovered a monophyletic cluster of six *H. symbiolongicarpus* 'echotoxin' transcripts as sister to a group of anthozoan toxin genes (60% BS support; Additional file 8), and a strongly supported (92% BS; Additional file 8) sister relationship between a four *H. symbiolongicarpus* toxins and two scyphozoan toxins from *Aurelia* (TX1 and TX2; Additional file 8). The remaining three polyp-significant toxins were not placed in any well-supported orthologous groups. Further study is warranted to determine if these toxins each play a unique role in different functions, such as prey capture, defense, and/or digestion.

Astacins. A large number of transcripts belonging to the astacin subfamily are up-regulated in the gastrozooid (44 total, Additional file 5; Additional file 9), consistent with one of their roles as digestive enzymes in other metazoans [30-33]. Expression studies of several astacin genes in hydrozoans also suggest a role in digestion. In *P. carnea, pmp1* is expressed in both the mouth of the polyp and the manubrium of the medusa stage [34]. Immunolocalization of the HMP1 protein found it expressed in the head and tentacle regions of *Hydra* [35], while Kumpfmüller et al. [36] found *farm1* expressed in both the epi and gastrodermal layers of gastric

region of *Hydra*.

Its important to note that digestion is just one function of the astacin subfamily. Another function is in regeneration, as shown in *H. echinata*, where Möhrlen et al. [37] found astacins *hea1* and *hea2* expressed throughout development and soon after the gastrozooid is subjected to tissue injury (expression in other polyp types not mentioned). HMP1 was also up-regulated during head regeneration in *Hydra* [35]. Orthologs of *hea1* and *hea2* were among the 44 gastrozooid-significant astacins in our study (94% and 99% BS support, respectively; see Additional files 5,9). Up-regulation of these transcripts may be a result of tissue damage response during dissections prior to RNA extractions. However, it is interesting that they are specific to the gastrozooid, suggesting that gonozooids and dactylozooids may have different regenerative properties than gastrozooids [38,1,3].

In situ hybridization

Figure 6 shows whole mount *in situ* hybridization (ISH) results of several polyp-significant transcripts identified through the DE analyses (listed in Table 3). DE analyses reported several different toxin transcripts to be differentially expressed between the different polyps. Polyp specificity of one of the three toxins identified as gastrozooid-significant by DE analyses, referred to here as *toxin_5320,* was confirmed by ISH. This transcript was expressed solely in specific gastrodermal cells around the base of the hypostome/tentacle margin of the gastrozooids. Three distinct cell types populate the gastrodermis of the hypostome in *Hydractinia*: gastrodermal epithelia (including digestive cells) and two glandular cell types (spumeous and spherulous cells) [39-41]. *Toxin_5320* expression appears to be limited to the spherulous cells of the hypostome (Fig. 6, Additional file 10). DE analyses found *toxin_3875* to

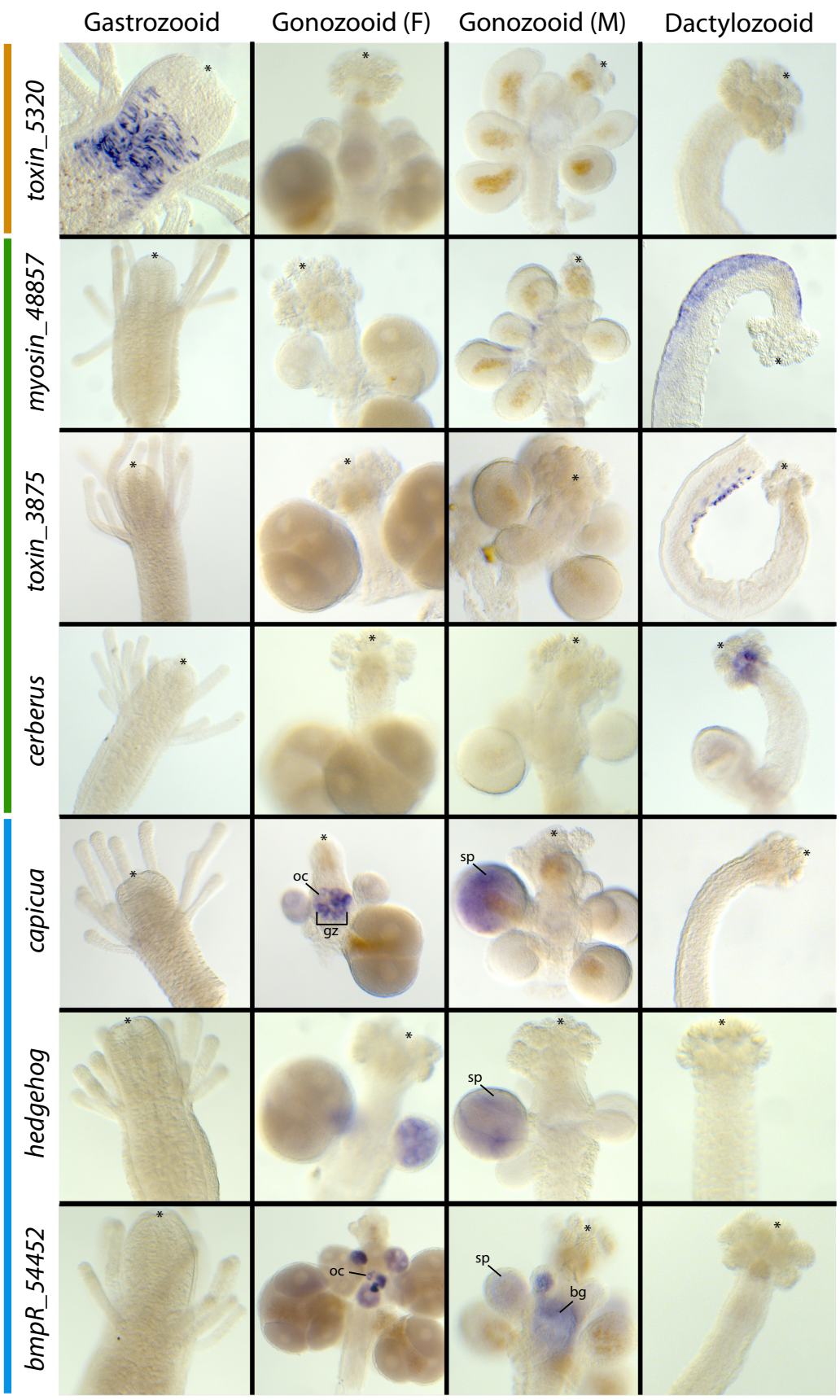

Gastrozooid
Gonozooid (F)
Gonozooid (M)
Dactylozooid
toxin_5320
myosin_48857
toxin_3875
cerberus
capicua
hedgehog
bmpR_54452
oc
gz
sp
oc
sp
sp
bg

Figure 6 – Images of whole mount *in situ* hybridization of polyp-significant transcripts.

***** = distal end of polyp; oc = oocytes; gz = germinal zone; sp = sperm; bg = non-specific staining.

be dactylozooid-significant and ISH found expression to be limited to nematocytes primarily found in the proximal portion of the body column of the dactylozooid (Fig. 6, Additional file 10).

Table 3. Polyp-significant DE transcripts analyzed with whole mount *in situ* hybridization.

Transcript ID	Name	Top blast hit	HMM family	Polyp Type
Hs_transcript_5320	*toxin_5320*	echotoxin a	Sea anemone cytotoxic protein	gastrozooid
Hs_transcript_48857	*myosin_48857*	myosin heavy chain isoform a	Myosin tail	dactylozooid
Hs_transcript_3875	*toxin_3875*	echotoxin a	Sea anemone cytotoxic protein	dactylozooid
Hs_transcript_44185	*cerberus*	cerberus 1	DAN domain	dactylozooid
Hs_transcript_16185	*capicua*	transcription factor capicua	HMG (high mobility group)	gonozooid
Hs_transcript_1524	*hedgehog*	indian hedgehog b	Hint module	gonozooid
Hs_transcript_54452	*bmpR_54452*	BMP receptor	Protein kinase domain	gonozooid

A myosin gene, referred to here as *myosin_48857*, was identified as a dactylozooid-significant gene by DE analyses. ISH confirmed this, recovering expression limited to the ectoderm of the body column of the extended side when the dactylozooid is curled in on itself (Fig. 6). Minor expression is also detected around the base of the gonophores and on the body column of some gonozooids (not shown).

The gene *cerberus* is also found to be dactylozooid-significant by DE analysis. This gene is only expressed in the gastrodermis beneath the clusters of nematocysts at the distal end of the dactylozooid (Fig. 6). Expression studies of *cerberus* in other metazoans have shown it to act as an antagonist of TGF-ß and Wnt signaling [42,43]. Here, expression in the dactylozooids is consistent with its antagonist role in Wnt signaling. *H. symbiolongicarpus*' canonical wnt, *HsWnt3* [GenBank:KF745052], is expressed at the distal tip of the dactylozooid (not shown, unpublished). *Cerberus* is expressed at the proximal boundary of *HsWnt3* expression, potentially acting to maintain *HsWnt3*'s expression boundary. This however appears to be specific to the

dactylozooids, as *H. symbiolongicarpus* feeding polyps express Wnt3 (not shown, unpublished), similar to other hydrozoan feeding polyps including *H. echinata* [44-46], *P. carnea* (unpublished), and *Hydra* [47-51]), but do not express *cerberus* (Fig. 6).

ISH also confirmed the specificity of several gonozooid-significant DE transcripts. Expression of the *hedgehog* homolog is restricted to the gastrodermis of both male and female gonophores (Fig. 6). Expression of a bmp receptor gene, tentatively called *bmpR_54452,* and *capicua* are primarily limited to developing oocytes in females and the gastrodermis of male gonophores (Fig. 6). ISH expression patterns of these transcripts suggest their involvement in some stage of meiotic/mitotic division during gametogenesis. In *Hydractinia*, oogenesis begins in the germinal zone (body column) of female gonozooids and oocyte differentiation continues after moving into the gonophores [52,1,2], while spermatogenesis takes place entirely in the gastrodermis of the male gonophores [52,2].

For several of these transcripts, expression in the females might not only be associated with germline proliferation, but with maternal transcript generation as well. Maternal expression of *capicua* and BMP receptors in early embryonic development has been reported in other metazoans [53,54]. Expression corresponding to maternal transcript generation is consistent with strong expression around developing oocytes in the germinal zone. By contrast, ISH of *hedgehog* in *Hydractinia* recovered no expression in the germinal zone of female gonophores (Fig. 6). Instead, its expression was limited to the gastrodermal tissues surrounding maturing oocytes in female gonophores. This is consistent with *hedgehog* genes implicated in germline proliferation and differentiation in other metazoans [55,56] and in *Nematostella,* where one *hedgehog* appears to be involved in germline proliferation, but lacks maternal expression [57].

Conclusion

Our non-biased approach of characterizing differential expression in different polyp types enabled us to identify key genes potentially involved in the morphological and functional differences between these different polyps. However, in interpreting results from a DE analysis, it is important to understand the distinction between biological relevance and statistical significance. We do not propose that every transcript in our list of putative polyp-significant genes is involved in the patterning or function of these different polyps, nor do we report to have captured all polyp-significant genes. One type of information not captured in this method would be those genes whose spatial or temporal expression (but not abundance) confers differences between polyp types. For example, the parahox gene *Cnox-2,* which was shown to be expressed in all polyps uniformly except for the oral region of the gastrozooid [7] was not recovered in the DE analysis. This is likely due to the fact that *Cnox-2* has different patterns of expression but not distinct differences in abundance between polyp types.

Even given the potential limitations, this unbiased approach of RNA-Seq DE analysis, selectively validated through *in situ* hybridization, identified many potential patterning and functional/structural genes without limiting our investigations to particular candidate genes. While potentially originating through simple changes in patterning, polymorphic polyps in *Hydractinia* are the result of differentially expressed functional, histological, and patterning genes. The DE genes identified in our study provide a starting point for future investigations of the developmental patterning and functional differences that are displayed in the different polyp types that confer a division of labor within a colony of *H. symbiolongicarpus*.

Materials and Methods

Animal Care

Colonies of *H. symbiolongicarpus* encrusting on gastropod shells occupied by the hermit crab *Pagurus longicarpus* were purchased from Marine Biological Laboratories (Woods Hole MA). Some colonies of *H. symbiolongicarpus* were surgically explanted onto microscope slides, placed in slide racks kept in seawater (REEF CRYSTALS, Aquarium Systems) aquaria, maintained at 21°C, and fed 2-3-day-old nauplii of *Artemia* three times a week. *Pagurus longicarpus* were maintained in similar conditions and fed frozen shrimp three times a week.

Tissue Collection and RNA Isolation

Tissue and RNA preps were divided into two categories based on the ultimate use of the samples (transcriptome assembly or DE analyses). Gastrozooids, gonozooids, and dactylozooids were individually dissected and collected from colonies encrusting the gastropod shells inhabited by *P. longicarpus*. The fourth polyp type (tentaculozooid) was not collected due to its rare occurrence in a colony. Excised polyps were immediately flash-frozen and stored at -80°C until RNA extractions were performed. Care was taken to only include polyp tissue and to exclude tissue from the stolons and stolonal mat of the colony. In order to obtain sufficient quantities of tissue, polyps from multiple colonies were often pooled together.

RNA extractions were carried out on pooled samples of approximately 100 individuals of a single polyp type. Total RNA was isolated using the TriReagent isolation protocol (Invitrogen) followed by a DNase treatment using the TURBO DNase kit (Ambion) or performed at the University of Kansas Medical Genome Sequencing Facility (KUMC-GSF) according to standard Illumina protocols. In samples collected for transcriptome assembly, gonozooid samples were

from both male and female colonies and were pooled together during RNA extraction whereas, for the gonozooid samples collected for the downstream DE analyses, males and females were kept separate from tissue collection through sequencing.

Library Construction and Sequencing

RNA libraries were constructed according to the TruSeq RNA Sample Preparation Guide (Illumina) using the TruSeq RNA Sample Preparation Kit (Box A). To increase transcript discovery, libraries used for transcriptome assembly were normalized using the Evrogen duplex-specific thermostable nuclease (DSN) kit following the Illumina DSN Normalization protocol. DNA fragments with adapters ligated on both ends were PCR-enriched after DSN normalization. Three normalized libraries were constructed with an average insert size of 160bp and subsequently barcoded, pooled, and multiplexed across three lanes of an Illumina HiSeq2000 flowcell.

For DE analyses, a total of twelve other libraries (four for each polyp type, including two male and two female gonozooid libraries) were constructed similarly, but without DSN normalization at KUMC-GSF. These samples were barcoded, pooled, and multiplexed on a single lane of an Illumina HiSeq2500 flowcell. All libraries were 100bp paired-end and sequenced at KUMC-GSF.

Transcriptome Assembly and Annotation

The workflow from sequencing through transcript annotation and differential expression analyses is shown in Figure 2. Raw reads from all three normalized libraries were filtered based on quality score and separately mapped to a set of unpublished genomic scaffolds of *H.*

symbiolongicarpus using TopHat 2.0.6 [58]. TopHat alignments were assembled into transcripts using Cufflinks 2.1.1 [59], generating three separate assemblies, one for each library. These assemblies were then merged into a single assembly using the cuffmerge function from Cufflinks [59]. This assembly was then filtered by transcript size, removing assembled transcripts less than 200bp in length. This assembly has been submitted to the NCBI Transcriptome Shotgun Assembly (TSA) database (Accession Number GAWH00000000 [60]). The raw reads have been submitted to the NCBI Sequence Read Archive (SRA; Project Number: SRX474462).

Transcripts were annotated using several different methods. Gene Ontology (GO) terms were added with Blast2GO [61,62], using the BLASTX algorithm and a significance threshold of 1×10^{-03} to search against NCBI's non-redundant (NR) protein database. Annotation names from the GO analysis represent the top BLAST hit (Additional file 2). A set of conserved eukaryotic genes was identified with CEGMA v2.4 [63] (Additional file 3). HMM (hidden markov model) protein families from the PFAM [64] and TIGR [65] databases were assigned to the amino acid translation of the most likely reading frame (identified using an open reading frame prediction tool [66]) of each transcript using HMMscan [67] under default settings. HMMscan annotations were constrained to a significance threshold of 0.01 (Additional file 2). Orthogroups were assigned to the same amino acid translations using the orthoMCL web server [68] (Additional file 2).

Differential Expression Analyses

Reads from the 12 non-normalized RNASeq libraries were mapped to the transcriptome assembly using Bowtie2 2.0.2 [69]. The raw reads from these libraries have been submitted to the NCBI Sequence Read Archive (SRA; Project Number: SRX474878). Counts for transcripts

for each library were extracted from the bowtie output (.sam files) using a python script that only counts reads in which both paired reads mapped to the same transcript (Additional file 11). The count data for each library was then fed through the DESeq [70] and edgeR [71] packages to assess statistically significant DE between all pairwise combinations of polyp types, including a comparison between male and female gonozooids. Both methods were used because they often give distinctly different results, with DESeq generally being more conservative in its assessment [72-74].

Given that the goal of this study was to identify differential gene expression between somatic tissues in the different polyp types, it was necessary to reduce the effect of gametogenic expression for the DE analyses. In *Hydractinia*, there are no discernable morphological differences between male and female gonozooids aside from the type of gametes present. Thus it can be assumed that any differences in expression between male and female gonozooids can be attributed to differences in gametogenesis (heterogametic expression) and need to be accounted for prior to DE analyses between polyp types.

In an effort to distinguish between gametogenic-specific expression and expression specific to gonozooid polyp identity, several preliminary DE analyses were conducted to adjust for gametagenic expression (Fig. 7). First, a DE analysis was conducted between male and female gonozooid libraries (step 1, Fig. 7), identifying significantly ($p_{adj} < 0.05$) up- or down-regulated transcripts (step 2, Fig. 7). Second, transcripts found to be significant were excluded from the template pool; a DE analysis was then performed on non-significant transcripts that included counts from both male and female libraries (step 3 shown in black, Fig. 7). However, this analysis excluded maternal transcripts that could also play a role in somatic morphogenesis. To include these maternal transcripts, a second DE analysis was conducted on transcripts up-

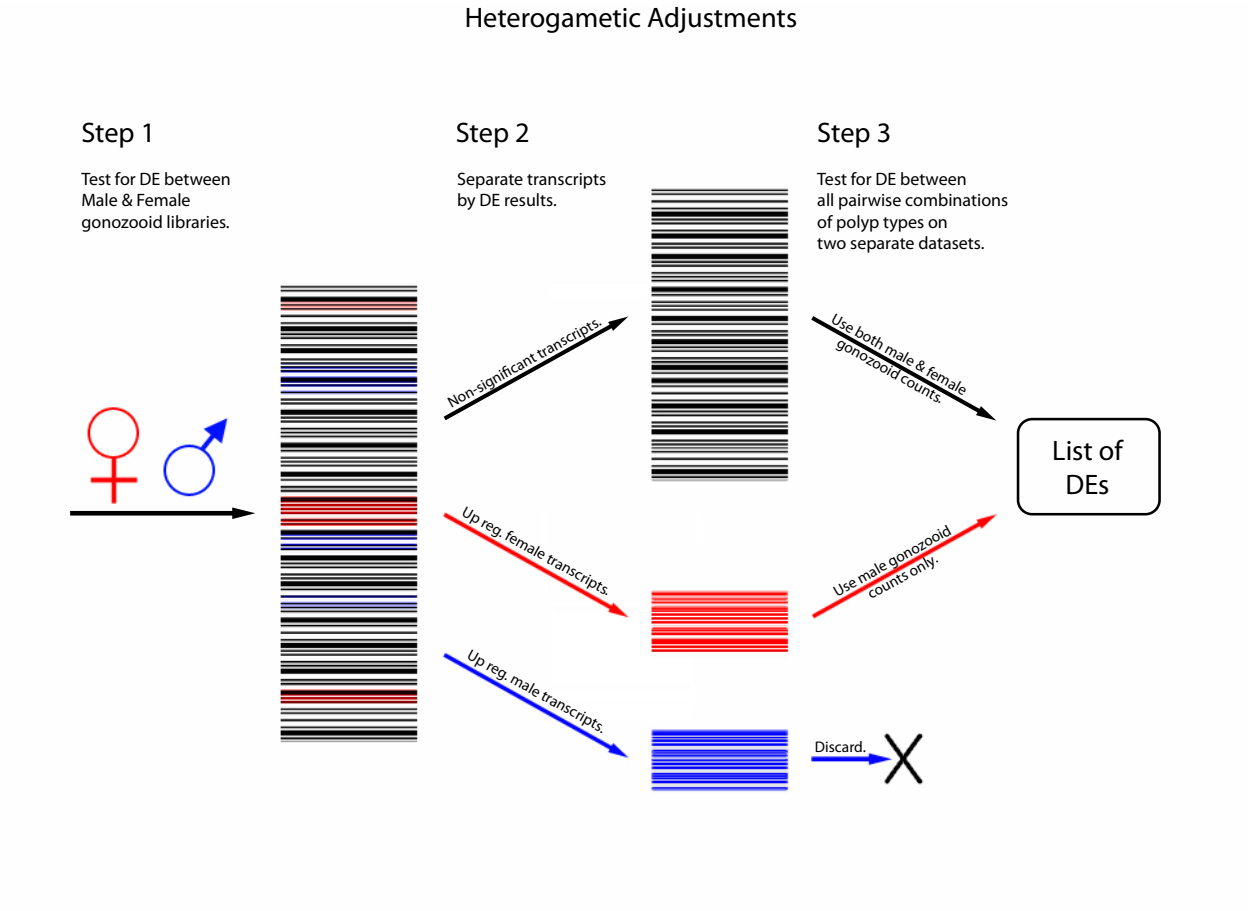


Figure 7. Diagram illustrating approach taken for adjusting for gametic expression.

DE analysis was conducted between male and female gonozooid libraries (step 1), identifying significantly up- or down-regulated transcripts (step 2). Those transcripts found to be significant were excluded from the template pool and a DE analysis on those non-significant transcripts included counts from both male and female libraries (step 3 shown in black). Then, a second DE analysis was conducted on transcripts up-regulated in female gonozooids. For these comparisons, only the expression counts from male gonozooids libraries were used (step 3 shown in red). Results from the both analyses (step 3) were combined.

regulated in female gonozooids (putative maternal transcripts). For these comparisons, only the expression counts from the male gonozooid libraries were used (step 3 shown in red, Fig. 7). In this approach, expression patterns consistent with developmental patterning and functional specialization were less likely obscured by the expression of genes specific to gametogenesis in the DE analyses. Results from both analyses (step 3, Fig. 7) were combined.

These DE analyses produced a list of DE transcripts specific to one or more of the three pairwise comparisons made between the different libraries, but not ones truly specific to a certain polyp type. In order to identify these polyp-significant transcripts, only transcripts significantly up- or down-regulated ($p_{adj} < 0.05$) in a particular polyp when compared to either of the other two polyp types (must be significant in only two of the three pairwise comparisons) from both edgeR and DESeq were considered polyp-significant. Figure 5 is a Venn diagram that lists the number of transcripts that meet these requirements (Additional file 4).

Probe Synthesis and *in situ* hybridization

Several polyp-significant transcripts identified during the DE analyses were selected for confirmation and further investigation with whole mount *in situ* hybridization (ISH) experiments (Table 3). Sequences for these transcripts were identified in the assembly, amplified from cDNA, cloned using the Invitrogen TOPO-TA Cloning Kit, and DIG labeled riboprobes were synthesized from clones using the Invitrogen T7/T3 Megascript kit. ISH of these transcripts were performed following methods from Nawrocki & Cartwright [75].

Molecular Phylogenetic Analyses

Several gene trees were constructed of select gene families, including homeodomains,

myosins, toxins, and astacins (Additional files 6-9). Cnidarian sequences belonging to families of interest were mined from the nr NCBI database and aligned using Mafft [76]. Depending on the family, either the L-insi or E-insi alignment algorithm was used. Only polyp-significant *H. symbiolongicarpus* sequences annotated with these families were included in the alignments. Maximum likelihood estimates of the molecular phylogenies of these gene families were then produced using RAxML [77] on the CIRPES portal [78] using the rapid bootstrapping (-f a) algorithm with 1,000 bootstrap replicates under the PROTGAMMA+WAG model.

Competing interests

The authors declare they have no competing interests.

Abbreviations

DE: differential expression or differentially expressed; NGS: next generation sequencing; bp: base pair; ISH: whole mount *in situ* hybridization; DSN: Duplex-Specific thermostable nuclease; GO: gene ontology; HMM: hidden markov model; ORF: open reading frame.

Authors' contributions

SMS performed tissue collections for the 12 non-normalized libraries, assembled and annotated transcriptome, designed the heterogametic adjustment algorithm, performed DE analyses and ISH, and drafted the manuscript. MS performed tissue collection, RNA isolation, normalizations, and library construction of the 3 normalized libraries. MS also wrote the python script that extracts counts from the .sam files. PC oversaw the design and implementation of the study and helped draft the manuscript. All authors have read and approve of the final manuscript.

Acknowledgments

We would like to thank Uri Frank for help with the ISH protocol. We thank Clark Bloomer at KUMC-GSF for help in RNA isolation, library construction, and sequencing. We are grateful for the draft genomic scaffolds provided by Christine Schnitzler and Andy Baxevanis and for their comments on an earlier draft of this manuscript. This works was supported by NSF grant DEB 0̶95357 and MDIBL Visiting Scientist Fellowship to PC, and the MDIBL Towle

Fellowship to SMS.

References

1. Muller W: **Experimentelle Untersuchungen Über Stockentwicklung, Polypendifferenzierung, und sexualchimären bei *Hydractinia echinata*.** *Wilhelm Roux' Arch Entwicklungsmech. Org.* 1964, **155**:181-268.
2. Berrill NJ: **Growth and form in gymnoblastic hydroids; polymorphism within the Hydractiniidae.** *J. Morph.* 1953, **92**:241-272.
3. Burnett AL, Sindelar W, Diehl N: **An examination of polymorphism in the hydroid *Hydractinia echinata*.** *J. Mar. Biol. Assoc. U.K.* 1967, **47**:645-658.
4. Agassiz A: **North American Acalephae.** (University Press: Welsh, Bigelow, Cambridge, U.K.). 1865.
5. Haeckel E: **Report on the Siphonophorae Collected by H. M. S. Challenger During the Years 1873–76.** (The Challenger Reports, London). 1888.
6. Huxley TH: **The Oceanic Hydrozoa: A Description of the Calycophoridae and Physophoridae Observed During the Voyage of the H. M. S. "Rattlesnake" in the years 1846–1850: With a General Introduction.** (Ray Society, London). 1859.
7. Cartwright P, Bowsher J, Buss LW: **Expression of the HOX type gene, Cnox-2, and the division of labor in a colonial hydroid.** *PNAS* 1999, **96**:2183-2186.
8. Mokady O, Dick MH, Lackschewitz D, Schierwater B, Buss LW: **Over one-half billion years of head conservation? Expression of an *ems* class gene in *Hydractinia symbiolongicarpus* (Cnidaria: Hydrozoa).** *PNAS* 1998, **95**:3673-3678.
9. Siebert S, Robinson MD, Tintori SC, Goetz F, Helm RR, Smith SA, Shaner N, Haddock SHD, Dunn CW: **Differential gene expression in the siphonophore *Nanomia bijuga* (Cnidaria) assessed with multiple next-generation sequencing workflows.** *PLoS One*

2011, **6**:e22953.

10. Hao DC, Ge G, Xiao P, Zhang Y, Yang L: **The first insight into the tissue specific *Taxus* transcriptome via Illumina second generation sequencing.** *PLoS One* 2011, **6**:e21220.
11. Helm RR, Siebert S, Tulin S, Smith J, Dunn CW: **Characterization of differential transcript abundance through time during *Nematostella vectensis* development.** *BMC Genomics* 2013, **14**:266.
12. Lui S, Lin L, Jiang P, Wang D, Xing Y: **A comparison of RNA-Seq and high-density exon array for detecting differential expression between closely related species.** *Nucl. Acids Res.* 2010, **39**:578-588.
13. Kanska J, Frank U: **Novels roles for *Nanos* in neural cell fate determination revealed by studies in a cnidarian.** *J. Cell Science* 2013, doi: 10.1242/ jcs.127233.
14. Rebscher N, Volk C, Teo R, Plickert G: **The germ plasm component *vasa* allows tracing of the interstitial stem cells in the Cnidarian *Hydractinia echinata*.** *Dev. Dynamics* 2008, **237**:1736-1745.
15. Seipel K, Yanze N, Schmid V: **The germ line and somatic stem cell gene *Cniwi* in the jellyfish *Podocoryne carnea*.** *J. Dev. Biol.* 2004, **48**: 1-7.
16. Bebenek HG, Gates RD, Morris J, Hartenstein V, Jacobs DK: ***sine oculis* in basal Metazoa.** *Dev. Genes Evol.* 2004, **214**:342-351.
17. Martindale MQ, Pang K, Finnerty JR: **Investigating the origins of triploblasty: 'mesodermal' gene expression in a diploblastic animal, the sea anemone *Nematostella vectensis* (phylum, Cnidaria; class, Anthozoa).** *Dev.* 2004, **131**:2463-2474.
18. Srivastava M, Larroux C, Lu DR, Mohanty K, Chapman J, Degnan BM, Rokhsar DS: **Early evolution of the LIM homeobox gene family.** *BMC Biol.* 2010, **8**:4.

19. Mazza ME, Pang K, Reitzel AM, Martindale MQ, Finnerty JR: **A conserved cluster of three PRD-class homeobox genes (*homeobrain, rx* and *orthopedia*) in the Cnidaria and Protostomia.** *EvoDevo* 2010, **1**:3.

20. Hroudova M, Vojta P, Strnad H, Krejcik Z, Ridl J, Paces J, Vlcek C, Paces V: **Diversity, phylogeny and expression patterns of pou and six homeodomaintranscription factors in hydrozoan jellyfish *Craspedacusta sowerbyi*.** *PLoS One* 2012, **7**:e36420.

21. Stierwald M, Yanze N, Bamert RP, Kammermeier L, Schmid V: **The *Sine oculis/Six* class family of homeobox genes in jellyfish with and without eyes: development and eye regeneration.** *Dev. Biol.* 2004, **274**:70-81.

22. Nakanishi N, Yuan D, Hartenstein V, Jacobs DK: **Evolutionary origin of rhopalia: insights from cellular-level analyses of Otx and POU expression patterns in the developing rhopalial nervous system.** *Evol. & Dev.* 2010, **12**:404-415.

23. Millane RC, Kanska J, Duffy DJ, Seoighe C., Cunningham S, Plickert G, Frank U: **Induced stem cell neoplasia in a cnidarian by ectopic expression of a POU domain transcription factor.** *Dev*. 2011, **138**:2429-2439.

24. Gröger H, Callaerts P, Gehring WJ, Schmid V: **Gene duplication and recruitment of a specific tropomyosin into striated muscle cells in the jellyfish *Podocoryne carnea*.** *J. Exp. Zoo.* 1999, **285**:378-386.

25. Baader CD, Schmid V, Schuchert P: **Characterization of a tropomyosin cDNA from the hydrozoan *Podocoryne carnea*.** *Fed. Eur. Biochem. Soc.* 1993, **328**:63-66.

26. Miglietta MP, Cunningham CW: **Evolution of life cycle, colony morphology, and host specificity in the family Hydractiniidae (Hydrozoa, Cnidaria).** *Evol*. 2012, doi:10.1111/j.1558-5646.2012.01717.x.

27. Bouillon J, Medel D, Cantero ALP: **The taxonomic status of the genus *Stylactaria* Stechow, 1921 (Hydroidomedusae, Anthomedusae, Hydractiniidae), with the description of a new species.** *Sci. Mar.* 1997, **61**:471-486.

28. Boero F, Sará M: **Motile sexual stages and evolution of Leptomedusae (Cnidaria).** *Boll. Zool.* 1987, **54**:131-139.

29. Frazão B, Vasconcelos V, Antunes A: **Sea anemone (Cnidaria, Anthozoa, Actiniaria) toxins: An overview.** *Mar. Drugs* 2012, **10**:1812-1851.

30. Yasumasu S, Yamada K, Akasaka K, Mitsunaga K, Iuchi I, Shimada H, Yamagami K: **Isolation of cDNAs for LCE and HCE, two constituent proteases of the hatching enzyme of *Oryzias latipes*, and concurrent expression of their mRNAs during development.** *Dev. Biol.* 1992, **153**:250-258.

31. Hishida R, Ishihara T, Kondo K, Katsura I: **hch-1, a gene required for normal hatching and normal migration of a neuroblast in *C. elegans*, encodes a protein related to TOLLOID and BMP-1.** *EMBO J.* 1996, **15**:411-4122.

32. Geier G, Zwilling R. **Cloning and characterization of a cDNA coding for *Astacus* embryonic astacin, a member of the astacin family of metalloproteases from the crayfish *Astacus astacus*.** *Eur. J. Biochem.* 1998, **253**:796-803.

33. Fan TJ, Katagiri C: **Properties of the hatching enzyme from *Xenopus laevis*.** *Eur. J. Biochem.* 2001, **268**:4892-4898.

34. Pan T, Gröger H, Schmid V, Spring J: **A toxin homology domain in an astacing-like metallopteinase of the jellyfish of *Podocoryne carnea* with a dual role in digestion and development.** *Dev. Genes Evol.* 1998, **208**:259-266.

35. Yan L, Pollock GH, Nagase H, Sarras Jr MP: **A $25.7x10^3$ M_r hydra metalloproteinase**

(HMP1), a member of the astacin family, localizes to the extracellular matrix of *Hydra vulgaris* in a head-specific manner and has a developmental function. *Dev*. 1995, **121**:1591-1602.

36. Kumpfmüller G, Rybakine V, Takahashi T, Fujisawa T, Bosch TCG: **Identification of an astacin matrix metalloprotease as target gene for *Hydra* foot activator peptides.** *Dev. Genes Evol.* 1995, **209**:601-607.
37. Möhrlen F, Maniura M, Plickert G, Frohme M, Frank U: **Evolution of astacin-like metalloproteases in animals and their function in development.** *Evol. & Dev.* 2006, **8**:223-231.
38. Hazen AP: **Regeneration in *Hydractinia* and *Podocoryne*.** *Amer. Nat.* 1902, **36**:193-200.
39. Campbell RD: **Cell proliferation and morphological patterns in the hydroids *Tubularia* and *Hydractinia*.** *J. Embryol. Exp. Morph.* 1967, **17**:607-616.
40. Braverman M: **Studies on hydroid differentiation; the replacement of hypostomal gland cells of *Podocoryne carnea*.** *J. Morph.* 1968, **126**:95-106.
41. Bouillon J: **Les cellules glandulaires des hydroïdes et des hydroméduses; leur structure et la mature de leurs sécrétions.** *Cahiers Biologie Marine* 1966, **7**:157-205.
42. Hsu DR, Economides AN, Wang X, Eimon PM, Harland RM: **The *Xenpus* dorsalizing factor Gremlin indentifies a novel family of secreted proteins that antagonize BMP activities.** *Mol. Cell* 1998, **1**:673-683.
43. Piccolo S, Agius E, Leyns L, Bhattacharyya B, Grunz H, Bouwmeester T, De Robertis EM: **The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals.** *Nature* 1999, **397**:707-710.
44. Duffy DJ, Plickert G, Kuenzel T, Tilmann W, Frank U: **Wnt signaling promotes oral but**

suppresses aboral structures in *Hydractinia* metamorphosis and regeneration. *Dev*. 2010, **137**:3057-3066.

45. Müller W, Frank U, Teo R, Mokady O, Guette C, Plickert G: **Wnt signaling in hydroid development: ectopic heads and giant buds induced by GSK-3*ß* inhibitors.** *Int. J. Dev. Biol.* 2007, **51**:211-220.
46. Plickert G, Jacoby V, Frank U, Müller WA, Mokady O: **Wnt signaling in hydroid development: Formation of the primary body axis in embryogenesis and its subsequent patterning.** *Dev. Biol.* 2006, **298**:368-378.
47. Broun M, Gee L, Reinhardt B, Bode HR: **Formation of the head organizer in hydra involves the canonical Wnt pathway.** *Dev*. 2005, **132**:2907-2916.
48. Gee L, Hartig J, Law L, Wittlieb J, Khalturin K, Bosch TCG, Bode HR: ***ß*-catenin plays a central role in setting up the head organizer in *Hydra*.** *Dev. Biol.* 2010, **340**:116-124.
49. Guder C, Pinho S, Nacak TG, Schmidt HA, Hobmayer B, Niehrs C, Holstein TW: **An ancient Wnt-Dickkopf anatagonism in *Hydra*.** *Dev*. 2005, **133**:901-911.
50. Lengfeld T, Watanabe H, Simakov O, Lindgens D, Gee L, Law L, Schmidt HA, Özbek S, Bode H, Holstein TW: **Multiple Wnts are involved in *Hydra* organizer formation and regeneration.** *Dev. Biol.* 2009, **330**:186-199.
51. Nakamura Y, Tsiairis CD, Özbek S, Holstein TW: **Autoregulatory and repressive inputs localize *Hydra Wnt3* to the head organizer.** *PNAS* 2011, **108**:9137-9142.
52. Bunting M: **The origin of the sex-cells in *Hydractinia* and *Podocoryne*; and the development of *Hydractinia*.** *J. Morph.* 1894, **9**:203-236.
53. Jiménez G, Guichet A, Ephrussi A, Casanova J: **Relief of gene repression by Torso RTK signaling: role of *capicua* in *Drosophila* terminal and dorsoventral patterning.** *Genes*

Dev. 1894, **14**: 224-231.

54. Monteiro R, van Dinther M, Bakkers J, Wilkinson R, Patient R, ten Diijke P, Mummery C: **Two novel type II receptors mediate BMP signaling and are required to establish left-right asymmetry in zebrafish.** *Dev. Biol.* 2008, **315**: 55-71.
55. Deshpande G, Swanhart L, Chiang P, Schedl P: **Hedgehog signaling in germ cell migration.** *Cell* 2001, **106**:759-769.
56. Bitgood MJ, Shen L, McMahon AP: **Sertoli cell signaling by Desert hedgehog regulates the male germline.** *Current Biol.* 1996, **6**:298-304.
57. Matus DQ, Magie CR, Pang K, Martindale MQ, Thomsen GH: **The Hedgehog gene family of the cnidarian, *Nematostella vectensis*, and implications for understanding metazoan Hedgehog pathway evolution.** *Dev. Biol.* 2008, **313**: 501-518.
58. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL: **TopHat2: accurate alignment of transcriptomes in the presence of insertation, deletions and gene fusions.** *Genome Biol.* 2013, **14**:R36.
59. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L: **Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation.** *Nature Biotech.* 2010, **28**: 511-515.
60. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GAWH00000000. The version described in this paper is the first version, GAWH01000000.

61. Conesa A, Götz S, Garcia-Gomez JM, Terol J, Talon M, Robles M: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.** *Bioinformatics* 2005, **21**:3674-3676.

62. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A: **High-throughput functional annotation and data mining with the Blast2GO suite.** *Nucl. Acids Res.* 2008, **36**:3420-3435.

63. Parra G, Bradnam K, Korf I: **CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes.** *Bioinformatics* 2007, **23**:1061-1067.

64. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD: **The Pfam protein families database.** *Nucl. Acid Res.* 2012, **Database Issue**:doi:10.1093/nar/gkr1065.

65. **TIGR Database** [http://blast.jcvi.org/web-hmm/]

66. Min XJ, Butler G, Storms R, Tsang A: **OrfPredictor: predicting protein-coding regions in EST-derived sequences.** *Nucleic Acids Res.* 2005, **Web Server Issue**:W677-W680. [http://proteomics.ysu.edu/tools/OrfPredictor.html]

67. HMMER [http://hmmer.org/]

68. Li L, Stoeckert CJ, Roos DS: **OrthoMCL: Identification of orthologous groups for Eukaryotic genomes.** *Genome Res.* 2003, **13**:2178-2189.

69. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nature Methods* 2012, **9**:35-359.

70. Anders S, Huber W: **Differential expression analysis for sequence count data.** *Genome Biol.* 2010, **11**:R106.

71. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductorpackage for differential expression analysis of digital gene expression data.** *Bioinformatics* 2010, **26**: 139-140.

72. Rapaport F, Khanin R, Liang Y, Pirun M, Krek A, Zumbo P, Mason CE, Socci ND, Betel D: **Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data.** *Genome Biol*. 2013, **14**:R95.

73. Anders S, McCarthy DJ, Chen Y, Okoniewski M, Smyth GK, Huber W, Robinson MD: **Count-based differential expression analysis of RNA sequencing data using R and Bioconductor.** *Nature Prot.* 2013, **8**: 1765-1786.

74. Soneson C, Delorenzi M: **A comparison of methods for differential expression analysis of RNA-seq data.** *BMC Bioinformatics* 2013, **14**:91.

75. Nawrocki AM, Cartwright P: **Expression of Wnt pathway genes in polyps and medusa-like structures of *Ectopleura larynx* (Cnidaria: Hydrozoa).** *Evol. & Dev.* 2013, **15**:373-384.

76. Katoh K, Kuma K, Toh H, Miyata T: **MAFFT version 5: improvement in the accuracy of multiple sequence alignment**. *Nucl. Acids Res.* 2005, **33**:511-518.

77. Stamatakis A, Hoover P, Rougemont J: **A fast bootstrapping alogorithm for the RAxML web-servers.** *Sys. Biol.* 2008, **57**:758-771.

78. Miller MA, Pfeiffer W, Schwartz T: **Creating the CIPRES Science Gateway for inference of large phylogenetic trees.** *GCE* 2010, 1-8.

79. Cartwright P, Nawrocki AM: **Character evolution in Hydrozoa (phylum Cnidaria).** *J. Integ. Comp. Biol.* 2010, **50**:456-472.

Additional files

Additional file 1. Histogram of the size distribution of assembled transcripts.

This does not include transcripts that were removed because they were < 200bp in length. Inset table displays assembly numbers and size statistics before and after filtering out the <200bp transcripts.

Additional file 2. Blast2GO, HMMscan, and orthoMCL annotations of all transcripts.

Additional file 3. CEGMA output.

Additional file 4. All polyp-significant transcripts.

This list includes the assembly sequence ID, top BLASTX hit, number of gene ontology IDs, top HMM protein domain, polyp specificity, top significance threshold, and transcript sequence. Polyp specificity is defined in two separate columns, 'Polyp' and 'Direction' (example: 'Gono, DOWN, .05,' would be a transcript that is significantly down regulated in the gonozooid when compared to the other two polyps at a significance level between .01 and .05).

Additional file 5. List of polyp-significant DEs discussed in Results and Discussion section.

This table is a subset of Additional file 4 and contains all the same information, but with BS support and the accession number of the closest cnidarian ortholog, if molecular phylogenetic analyses performed. In cases where phylogenetic analyses recovered well supported sister

relationships between two monophyloetic clades with no clear one to one orthologous relationships, a single accession number was selected from the non-*H. symbiolongicarpus* monophyly. List is divided up by color according to the subheading in the Results and Discussion section they are discussed in: Blue – 'Gametogenic expression'; Yellow – 'Homeodomains'; Pink – 'Myosins'; Green – 'Toxins'; Orange – 'Astacins'. (*) Marks transcripts with whole mount *in situ* hybridizations data in this study.

Additional file 6. Cnidarian homeodomain gene tree.

Molecular phylogeny of cnidarian homeodomains sampled from GenBank's nr database. Accession numbers are appended to the ends of the tip labels. Only polyp-significant homeodomains from *H. symbiolongicarpus* (highlighted in red) were included in the analysis. Fasta and alignment file available upon request.

Additional file 7. Cnidarian myosin gene tree.

Molecular phylogeny of cnidarian myosins sampled from GenBank's nr database. Accession numbers are appended to the ends of the tip labels. Only polyp-significant myosins from *H. symbiolongicarpus* (highlighted in red) were included in the analysis. Fasta and alignment file available upon request.

Additional file 8. Cnidarian toxin gene tree.

Molecular phylogeny of cnidarian toxins sampled from GenBank's nr database. Accession numbers are appended to the ends of the tip labels. Only polyp-significant toxins from *H.*

symbiolongicarpus (highlighted in red) were included in the analysis. Fasta and alignment file available upon request.

Additional file 9. Cnidarian astacin gene tree.

Molecular phylogeny of cnidarian astacins sampled from GenBank's nr database. Accession numbers are appended to the ends of the tip labels. Only polyp-significant astacins from *H. symbiolongicarpus* (highlighted in red) were included in the analysis. Fasta and alignment file available upon request.

Additional file 10. *In situ* hybridization (higher magnification) of toxins.

A. *toxin_5320.* B. *toxin_3875***.** sp = spumeous cells; ns = nematocyst; nc = nematocyte.

Additional file 11. Python script for extracting count data from .sam files.

Use: python counts-paired.py infile.sam

CHAPTER 2:

Patterns of Wnt signaling in the life cycle of *Podocoryna carnea* and its implications for medusae evolution in Hydrozoa (Cnidaria)

Abstract

Background

Hydrozoans are known for their complex life cycles, alternating between benthic, asexually reproducing polyps and pelagic, sexually reproducing medusae. In most hydrozoans however, the medusa life cycle stage is reduced and sexual maturity is reached in a gonophore that resembles an earlier ontogenetic stage of medusae development. Reduced forms are characterized by the degree of developmental truncation of the structures developing along the proximal-distal axes of the gonophore. In hydrozoans, canonical Wnt signaling has been implicated in defining and maintaining the oral-aboral axis of the polyp. Although the role of Wnt signaling in medusa development is largely unknown, Wnt gene expression in taxa with developmentally reduced gonophores suggests correlations between decreased Wnt pathway signaling and reduction of gonophore development. Here we use RNA-Seq data collected from three discrete life cycle stages of the medusae-bearing hydrozoan species *Podocoryna carnea,* to assemble, annotate, and assess differential expression (DE) of the transcriptome of *P. carnea*. Through DE analysis and *in situ* hybridization (ISH), we identify key components of canonical Wnt signaling up-regulated in a spatially restricted pattern during medusa development.

Results

Non-canonical and canonical Wnt signaling genes are significantly enriched in the pool of transcripts that are differentially expressed between life cycle stages of *P. carnea*. Spatial expression analyses revealed co-expression of the ligand *Wnt3* and receptor *frizzled3* at the distal/oral ends of the developmental axes of medusae and polyps in *P. carnea*.

Conclusions

DE and ISH results presented here reveal expression of canonical Wnt signaling consistent with it playing a role in the development of medusae, similar to what has been previously characterized in hydrozoan polyps. These findings, in conjunction with previous Wnt expression studies in taxa with reduced gonophores, suggests that down regulation of the Wnt pathway may play a key role in the loss of the medusa life cycle stage in hydrozoan evolution.

Keywords

Podocoryna carnea, RNA-Seq, transcriptomes, Wnt signaling, differential expression

Background

Hydrozoans exhibit complex life cycles, alternating between a benthic, asexually reproducing polyp stage and a pelagic, sexually reproducing medusa stage (Figure 1A). During medusa development, two axes are being patterned: the proximal-distal axis of the bell and the oral-aboral axis of the manubrium, the structure containing the mouth at it oral end and in some taxa, bearing gonads (Figure 1B). In most hydrozoan species however, the medusa life cycle stage is reduced [1] and sexual maturity is reached in a gonophore that resembles an ontogenetic stage of medusae development. These reduced forms are characterized by the level of developmental truncation of the structures developing from these axes. Sporosacs lack nearly all medusae-like characteristics and remain attached to the colony, whereas medusoids have remnants of medusa structures, do not feed, and may or may not detach from the colony [2]. Several phylogenetic studies have recovered strong support for repeated independent losses of medusae [3-7], and possible re-gain [5-7] in hydrozoan evolution.

In Hydrozoa, the canonical Wnt pathway has been implicated in defining and maintaining the anterior-posterior axis of embryos and planula larvae [8-12] as well as defining and maintaining the oral-aboral axis of polyps [13-25]. In the hydrozoans, *Hydra* and *Hydractinia*, over-expression of Wnt results in development of ectopic oral structures (such as tentacles) along their polyp body column [11,13,14,15,20,21,23]. During regeneration, misexpression of this pathway through knockdowns results in aboralized polyps (no head) [14,15], while over-expression results in oralized polyps (multiple heads) [13-15]. Similarly, over-expression during metamorphosis leads to an oralized primary polyp that lacks an aboral end [15,26].

While expression of the Wnt pathway is well characterized in hydrozoan larval and polyp forms, little is known about its role in medusa development, although previous studies have

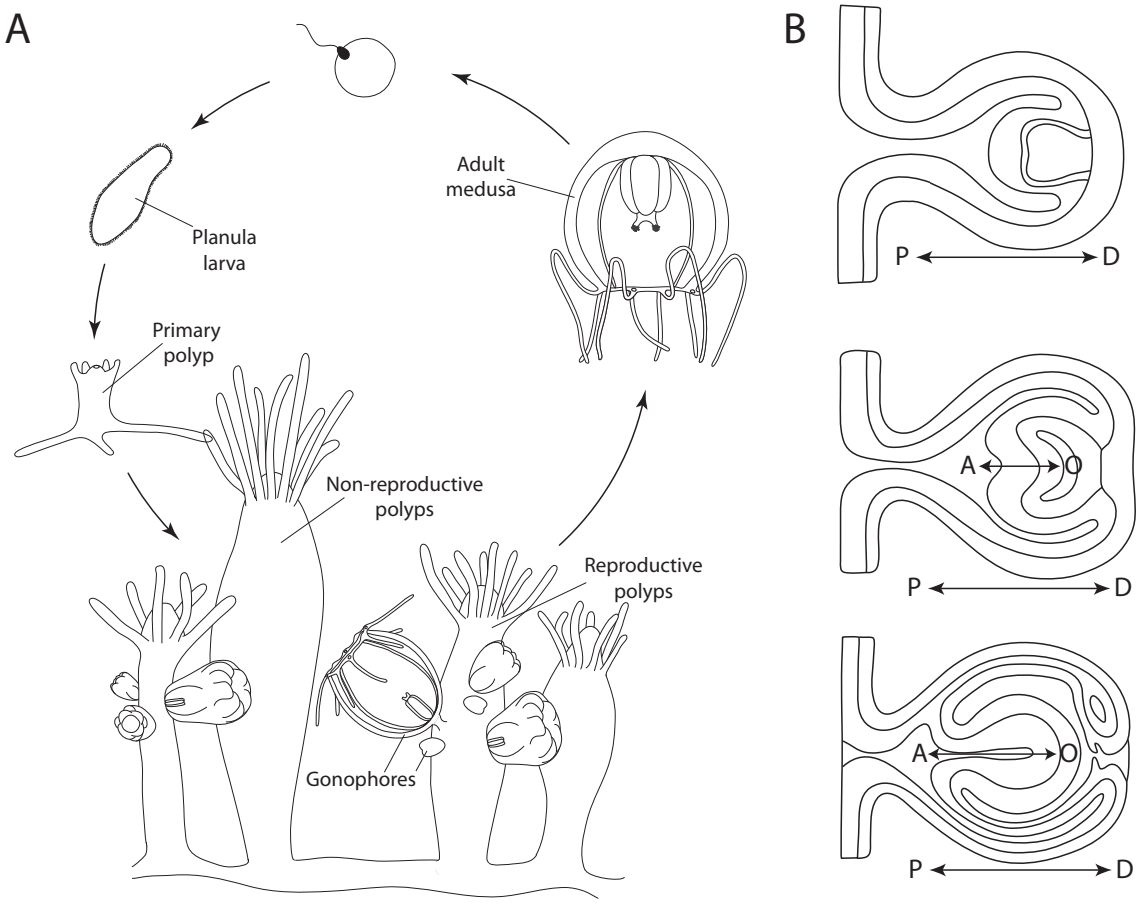


Figure 1. Illustration of *P. carnea*'s life cycle and developmental axes of medusae buds.

A) *Podocoryna carnea* buds medusae along the body column of reproductive polyps. These eventually detach from the colony become sexually mature and spawn in the water column. Planula larvae settle onto a substrate and metamorphoses into a primary polyp. This polyp will asexually produce other polyps to form a colony and eventually but medusae, repeating the cycle. B) Illustration of the two developmental axes of a medusae bud (gonophore) through several developmental stages: the oral-aboral axis of the developing manubrium and the proximal-distal axis of the bell. A = aboral. O = oral. P = proximal. D = distal.

reported expression patterns of some canonical Wnt pathway components in sporosacs [14,15,27] and medusae [8]. In the adult medusae of the leptothecate hydrozoan *Clytia hemisphaerica*, expression of two membrane bound Wnt receptors, *frizzled1* and *frizzled3*, have been characterized. While *frizzled1* exhibits expression specific to the proximal region of the tentacles, called tentacle bulbs, *frizzled3* was reported at the oral end of the fully developed manubrium, and at the distal portion of the bell along the ring canal (part of the digestive tract)[8].

In the aplanulate hydrozoan *Ectopleura larynx*, female and male sporosacs are sexually dimorphic, differing in their level of developmental truncation. Female sporosacs develop tentacle buds whereas males do not and instead have a thickening of the epithelia at the distal tip of the gonophore, called an apical cap. Nawrocki and Cartwright [27] revealed co-expression of the ligand, *Wnt3*, and membrane bound Wnt receptor, *frizzled1*, in both the tentacle buds of the female gonophore and the apical cap of the male gonophore. Furthermore, at the oral end of the spadix (anlage of the manubrium), male gonophores express a putative Wnt antagonist, *sFRP* (secreted frizzled related protein), while female gonophores express *frizzled1* in the corresponding region. In the sporosacs of the hydractiniid hydrozoan *Hydractinia echinata*, which lack all medusae-like structures, *Wnt3* is expressed at the distal tip of the sporosac [14,15], a manner reminiscent of *Wnt3* expression along the oral end of the polyp in this same species. No *frizzled1* expression was detected in the sporosac [14]. The lack of co-expression of a Wnt ligand and *frizzled* receptor in the distal axes of reduced gonophores suggest that down regulation of Wnt pathway elements may be involved in the arrest of the medusae developmental program in those species that lack medusae [27]. Characterization of expression of Wnt pathway genes in

medusae development should provide further insight into the role of the Wnt pathway in hydrozoan life cycle evolution.

Podocoryna carnea, a close relative of *Hydractinia* [28], has a fully developed pelagic medusa as part of its life cycle (Figure 1A). Here we assemble, annotate, and assess differential expression (DE) of the transcriptome of *P. carnea* using RNA-Seq data collected from three different life cycle stages. Our DE and whole mount *in situ* hybridization (ISH) results, in conjunction with previously published studies in other hydrozoans, suggest that changes in the regulation of the canonical Wnt signaling pathway may be responsible for the evolution and development of hydrozoan medusae and their reduced forms.

Materials and Methods

Animal Care

Transplanted colonies of *P. carnea* were kept on microscope slides, placed in slide racks and kept in seawater (REEF CRYSTALS, Aquarium Systems) aquaria at room temperature (~21°C) with a salinity of 29 ppt. Colonies were fed 2-3-day-old nauplii of *Artemia* three times a week.

Tissue Collection and RNA Isolation

Podocoryna carnea colonies were first starved a minimum of three days (up to five days) prior to tissue collection. For non-reproductive and reproductive polyp samples, roughly 100 polyps were individually dissected and immediately flash-frozen for each RNA extraction. Reproductive polyps sampled spanned all developmental stages of gonophore development. Non-reproductive polyps were sampled from colonies that did not show any evidence of medusa buds. To collect medusae tissue, reproductive colonies were kept in a small container overnight and liberated medusae were collected and flash-frozen the following morning. All tissue samples were stored at -80°C until RNA extractions could be performed. Total RNA was isolated at the University of Kansas Medical Genome Sequencing Facility (KUMC-GSF) using the TriReagent isolation protocol (Invitrogen).

Prior to tissue collection for whole mount *in situ* hybridization, colonies were starved for three days. After the third day, colonies were anesthetized with menthol crystals and fixed in 4% paraformaldehyde overnight at 4°C. Fixed tissue was rinsed with and stored in 100% methanol at -20°C. Developing medusae buds (gonophores) were staged according to Frey [29].

Library Construction and Sequencing

Podocoryna carnea RNA libraries were constructed at KUMC-GSF according to the TruSeq RNA Sample Preparation Guide (Illumina) using the TruSeq RNA Sample Preparation Kit (Box A). All libraries were 100bp paired-end with an average insert size of 160bp. Libraries were then barcoded, pooled, and multiplexed on a single lane of either an Illumina HiSeq2500 or HiSeq2000 flowcell at KUMC-GSF. Sample Med3 was multiplexed and sequenced on a lane of an Illumina HiSeq2000 flowcell, while all other samples were multiplexed and sequenced on the same lane of an Illumina HiSeq 2500. The raw reads have been submitted to the NCBI Sequence Read Archive (SRX529566).

Transcriptome Assembly and Annotation

The workflow from sequencing through transcript annotation and differential expression analyses is shown in Figure 2. Prior to assembly, low quality reads were trimmed or altogether removed using Trimmomatic [30] on each library. Reads passing this quality filter were combined into a single dataset for *de novo* transcriptome assembly using Trinity [31]. Post-assembly filters were applied to remove potential contaminants in the transcript pool using NCBI's UniVec [32]. Sequences identified by this process were trimmed or altogether removed.

This assembly was annotated following the methods described previously by Sanders et al. [33]. Gene Ontology (GO) terms were added with Blast2GO [34,35], using the BLASTX algorithm and a significance threshold of $1E^{-03}$ to search against NCBI's non-redundant (NR) protein database. Conserved 'core' eukaryotic genes were recognized with CEGMA v2.4 [36]. HMM (hidden markov model) protein families from the PFAM [37] and TIGR [38] databases were assigned to the amino acid translation of the most likely reading frame (identified using an

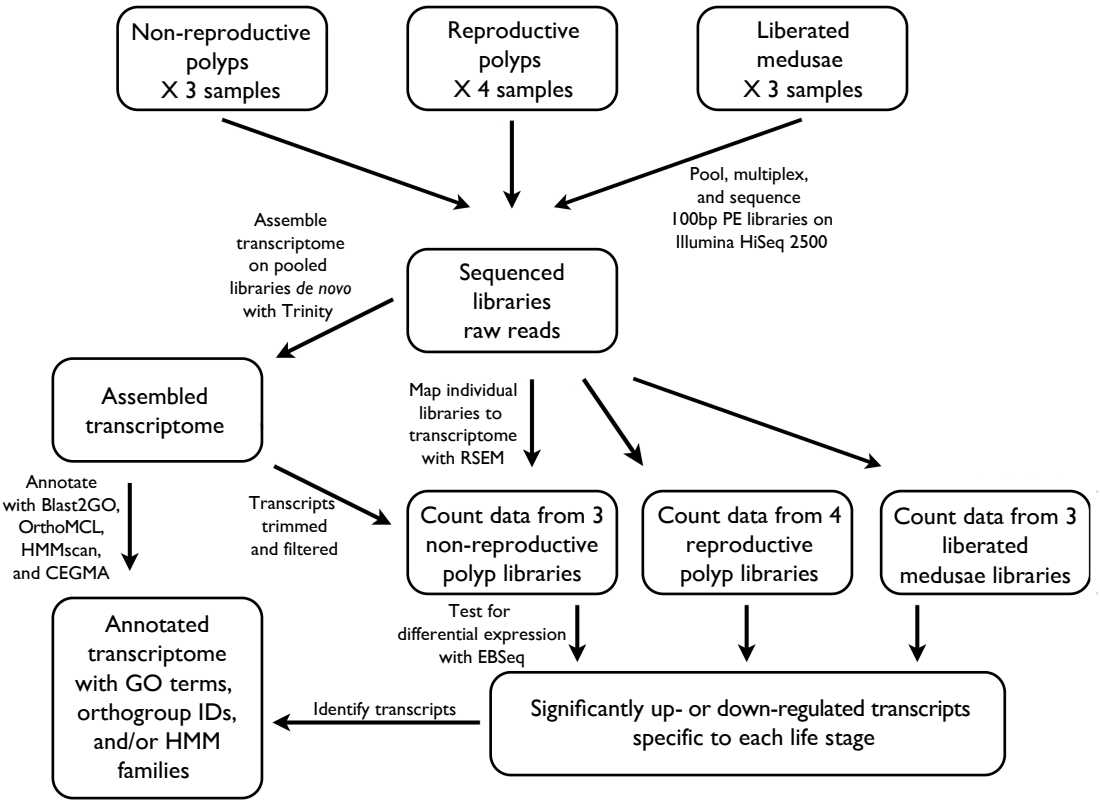


Figure 2. Bioinformatics workflow.

Reads were combined into a single dataset for *de novo* transcriptome assembly using Trinity. Transcripts were annotated with: Gene Ontology (GO) terms added with Blast2GO, conserved eukaryotic genes identified with CEGMA v2.4, HMM (hidden markov model) protein families from the PFAM and TIGR databases using HMMscan, and orthogroup IDs using OrthoMCL. Prior to differential expression, transcripts less than 400bp were excluded and expression for each transcript was estimated using RSEM. Differential expression was then assessed between all three stages simultaneously with EBSeq.

open reading frame prediction tool [39]) of each transcript using HMMscan [40] under default settings. HMMscan annotations were constrained to a significance threshold of 0.01. Amino acid translations of each transcript were also assigned orthogroup ids using OrthoMCL [41].

Differential Expression, KEGG Pathway, and Enrichment Analysis

Prior to differential expression, transcripts less than 400bp were excluded. Isoform- and gene-level expression was calculated using RSEM [42]. The 'isoform' dataset contains the counts for all assembled transcripts while the 'gene' dataset contains a combined count of all potential splice variants of a single 'gene'. Given that there is no genome available for *P. carnea*, we were unable to distinguish isoforms from incompletely assembled transcripts. Thus, to limit redundancy, differential expression was only assessed at gene-level expression. DE was tested simultaneously between all three stages with EBSeq [43]. EBSeq assesses differential expression between multiple conditions (life cycle stages) simultaneously by assigning a gene a posterior probability that it is differentially expressed (PPDE). DE results were constrained to a false discovery rate (FDR = 1 - PPDE) of $1E^{-06}$. Amino acid translations of DE genes specific to each life cycle stage were clustered with CD-HIT [44,45] and subjected to KEGG analysis [46,47]. Enrichment analyses were performed in R using a Fisher's Exact Test, comparing the total number DE genes with functional annotations linking them Wnt signaling to the total number of functionally annotated Wnt signaling genes in the transcriptome.

Probe Synthesis and *in situ* Hybridization

Sequences for *Wnt3, frizzled3,* and *frizzled1* transcripts were identified in the assembly, amplified from cDNA, cloned using the Invitrogen TOPO-TA Cloning Kit, and DIG labeled

riboprobes were synthesized from clones using the Invitrogen T7/T3 Megascript kit. Whole mount *in situ* hybridization (ISH) was adapted from Gajewsky et al. [48]. For sectioning, stained medusa were fixed in 4% PFA overnight at 4˚C and then dehydrated in 100% EtOH. Samples were then embedded in wax and cut into 7µm sections, counter-stained in eosin, and mounted with Canada Balsam.

Molecular Phylogenetic Analyses

Orthology of several DE genes was determined with gene trees constructed for several gene families involved in Wnt signaling. Cnidarian *dickkopf* (*dkk*) sequences were mined from both NCBI's nr protein and Compagen [49] databases and aligned using ClustalΩ [50]. DE expressed Wnts were combined with hydrozoan Wnt sequences recently published by Hensel et al. [51] and aligned with the L-insi alignment algorithm in Mafft [52]. Medusozoan sequences of membrane bound *frizzled* receptors were mined from NCBI's nr protein database and aligned using Mafft with the L-insi alignment algorithm [52]. Using RAxML [53] on the CIPRES portal [54], a ML (maximum likelihood) estimate molecular phylogenies were produced using the rapid bootstrapping (-f a) algorithm with 500 bootstrap replicates under the PROTGAMMA+WAG model.

Results and Discussion

Transcriptome Assembly and Annotation

After sequencing, reads from each library were pooled and subjected to quality filtering/trimming prior to assembly, leaving approximately 180 million paired-end (PE) reads of the original ~270 million total PE reads (Figure 2). From these, Trinity assembled ~196 thousand transcripts (≥ 200bp). NCBI's contamination detection analyses identified numerous foreign contaminant sequences (either parts of or the entire transcript). These sequences were trimmed to eliminate suspect regions, resulting in the removal of 11,593 transcripts (generally as a result of transcript length decreasing to < 200bp). This assembly has been submitted to the NCBI Transcriptome Shotgun Assembly (TSA) database (Accession Number GBEH00000000)[55]. Transcripts were further filtered by sequence length (≥ 400bp) to reduce the number of small, incomplete transcripts in the assembly. While the NCBI trimming step did very little to the assembly, filtering all transcripts less than 400 bases showed a significant improvement in the summary statistics (Table 1).

These trimming and filtering measures not only improved several of the assembly summary statistics, but also increased the proportion of the transcriptome that was annotated. Of the original 196,686 transcripts assembled, approximately 58% (115,055) were annotated in our pipeline. After the final filtering step, nearly 70% (73,894) of the 105,808 transcripts remaining in the final assembly were annotated by at least one of the annotation methods (Figures 2,3; Additional file 1). These annotations contain 7,784 GO terms, 13,351 HMM domains, 10,307 Orthogroups, and 452 'core' eukaryotic genes were annotated (Figure 3, Additional files 1,2).

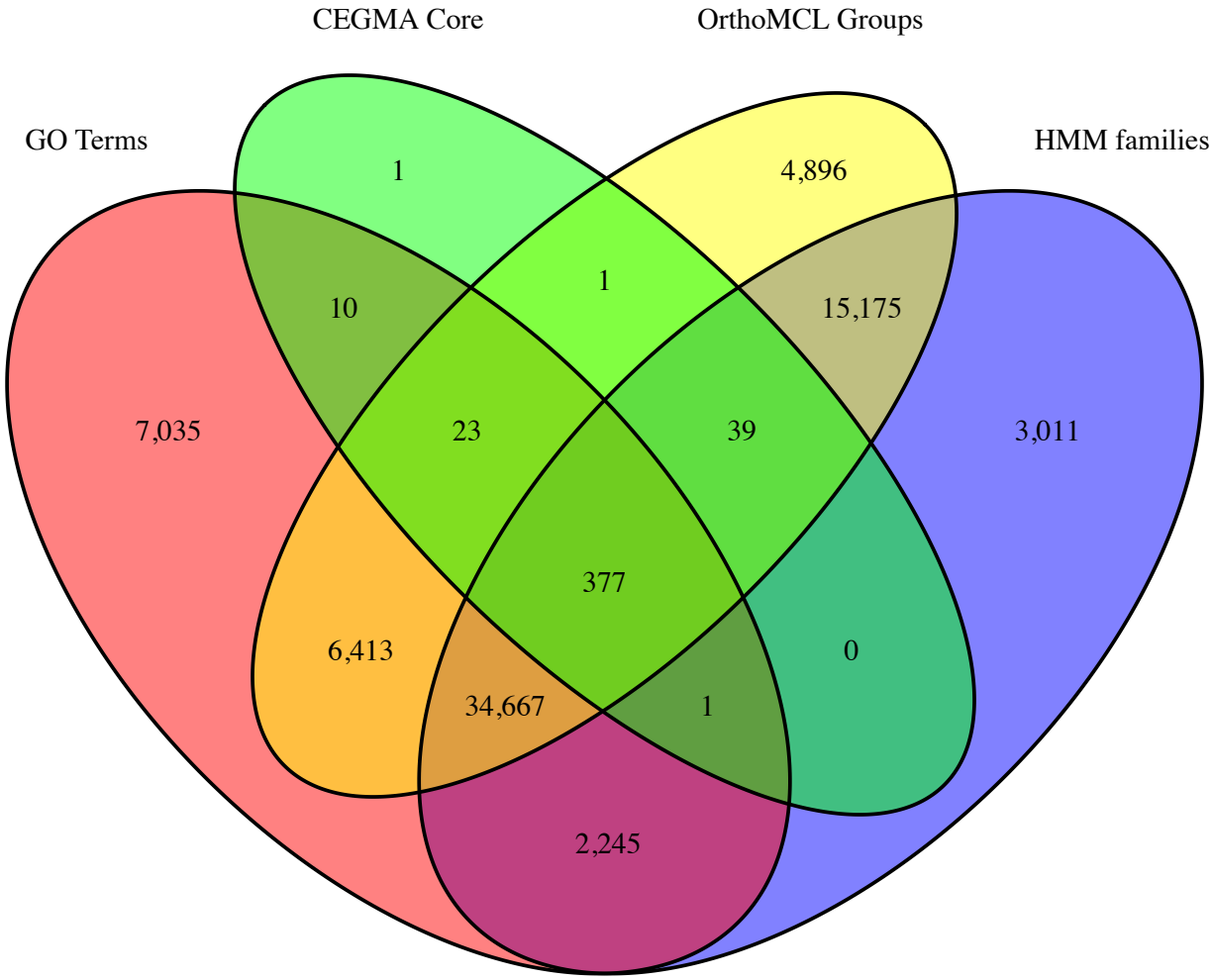


Figure 3. Venn-Diagram of annotated transcripts.

Four-way Venn-Diagram of the number of transcripts annotated in the final 'Filtered' assembly by each of four annotation methods in the bioinformatics workflow.

Table 1. Summary of assembly statistics at various stages of post-assembly filtering.

	All	Trimmed	Filtered
# Transcripts	196,686	185,093	105,808
N25 (bp)	2,459	2,500	2,712
N50 (bp)	1,411	1,444	1,684
N75 (bp)	640	672	984
GC Content	39.53%	39.20%	39.17%

All = Original *de novo* transcriptome assembly. Trimmed = transcriptome assembly after NCBI contamination filtering. Filtered = final transcriptome assembly after filtering by 400bp minimum transcript length.

Components and regulators of the Wnt signaling pathways are represented by 395 of the 50,771 functionally annotated (by Blast2Go) transcripts, with 164 of these specific to canonical Wnt signaling (Additional file 3). While identifying most key components, many other lineage-specific components were not identified by the Blast2Go functional annotations (discussed below).

Differential Expression Analysis

We used the program EBSeq [43] to compute the posterior probability that a given gene is differentially expressed between each of the three sampled tissues. Although medusa buds were not sampled in isolation, up- or down-regulation of a given gene in the reproductive polyp libraries should be specific to the developing medusa buds, as expression specific to the polyp should be equivalent to the expression captured in non-reproductive polyp libraries. Figure 4A summarizes the distance between the expression profiles of each of the ten libraries with a heatmap of the Euclidean distances. EBSeq recovered 38,114 differentially expressed genes with

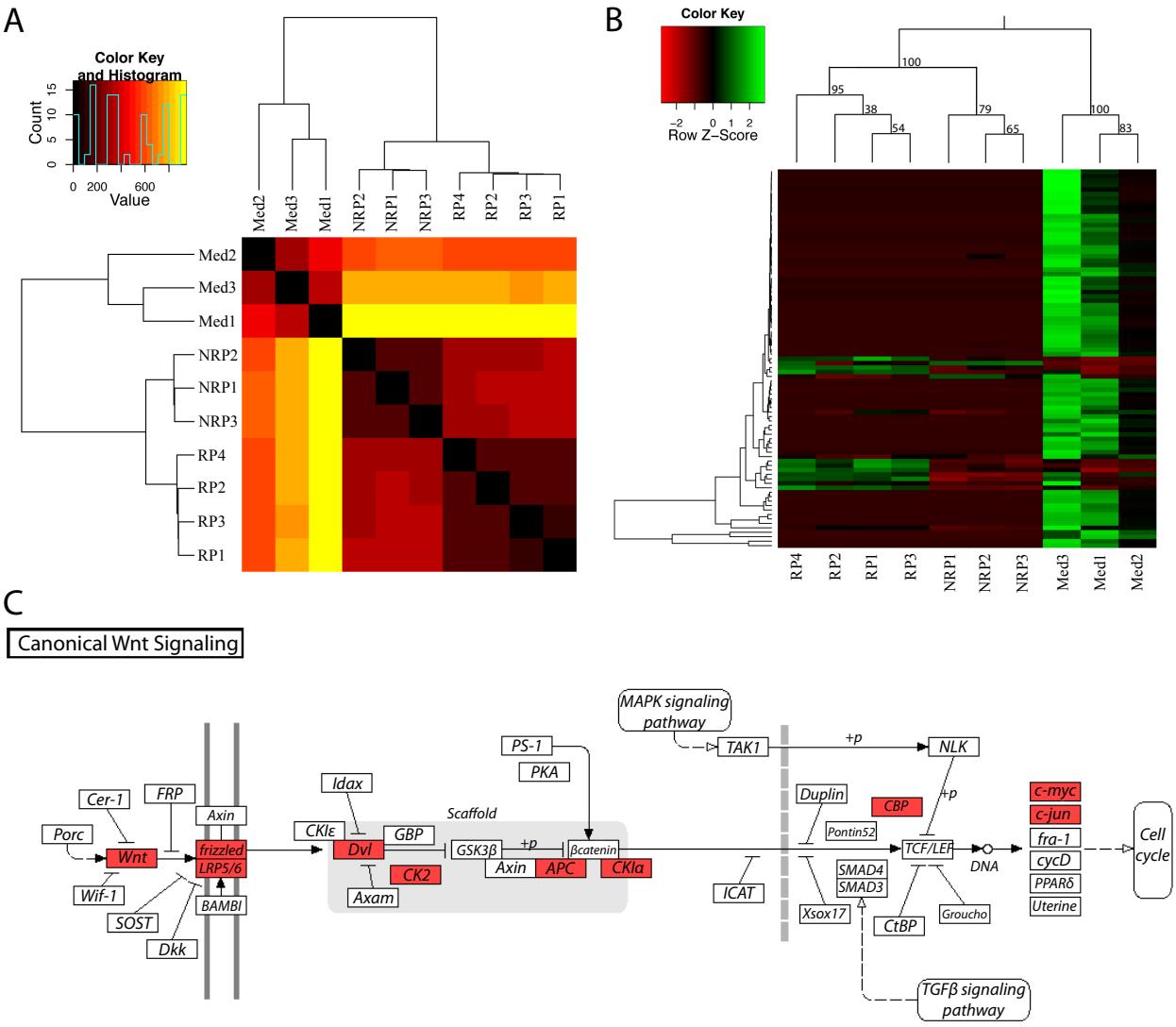


Figure 4. Heatmaps of RNA-Seq libraries and KEGG analysis.

A) Heatmap of the Euclidean distance of each library summed across all genes. B) Z-normalized heatmap of differentially expressed genes with functional annotations connected to canonical Wnt signaling. Dendrogram on the top generated using the *pvclust* package in R with 1000 bootstrap replicates. Numbers at each node in the dendrogram represent percentage of bootstrap support for that node. 'NRP' = non-reproductive polyps. 'RP' = reproductive polyps. 'Med' = adult medusa. C) KEGG pathway map of the canonical Wnt signaling pathway. Red boxes are genes up-regulated in the reproductive polyp.

an FDR ≤ 1e-6 (Table 2; Additional file 4). Table 2 summarizes the number of genes identified as either up or down in a specific life cycle stage. Approximately 68.5% of the DE genes are up-regulated in the medusae, indicating a highly divergent expression profile in this stage (Table 2; Figure 4A; Additional file 4). The high percentage of transcripts up-regulated in the adult medusa is not entirely surprising given the biological differences between this life cycle stage and the other two conditions which both include polyp tissue (Figure 1A). While there could be many factors driving the disparity between the adult medusae and other tissues sampled, gene expression associated with oogenesis is likely the most significant contributing factor. Genes of primary interest to this study are those up-regulated in the reproductive polyp (3,051; Table 2; Additional file 4), as this is the stage where most of the morphogenic patterning of the medusa occurs.

Table 2. Number of DE genes specific to each life cycle stage.

	# of Replicates	Up-Regulated	Down-Regulated
Medusae	3	26,104 (150/75)	409 (2/0)
Reproductive Polyps	4	3,051 (15/7)	1,192 (3/1)
Non-reproductive Polyps	3	5,358 (15/1)	2,000 (6/1)

Total number of genes (all Wnt signaling/canonical Wnt signaling).

Wnt signaling

Given the critical role of the Wnt pathway in developmental patterning of hydrozoans, we wanted to determine if GO terms connected to Wnt signaling were enriched in the pool of genes differentially expressed between life cycle stages. Enrichment analysis of the Blast2Go annotations revealed a significant deviation from the expected number of differentially expressed genes involved in Wnt signaling when compared to the total number of functionally annotated

Wnt genes (p≈0.02; Table 2, 3; Additional file 5). Furthermore, 83 of the 85 DE canonical Wnt genes were specific to developing and adult medusa stages (enrichment p≈0.0041; Table 2, 3; Figure 4B; Additional file 5).

Table 3. Enriched number of DE Wnt signaling genes.

	Wnt	Non-Wnt
DE	191 (83)	19,648 (16,033)
Total	268 (118)	33,679 (33,829)

Contigency table used to test for enrichment DE genes Wnt signaling functional annotations. DE analysis was performed with gene level estimates of expression (as estimated by RSEM). Thus, numbers in the columns represent the total number Wnt signaling (canonical Wnt signaling) genes, and not total transcripts.

Dickkopf family genes, which have been shown to be involved in Wnt signaling in hydrozoans, are also differentially expressed between life cycle stages. Homologs of *dkk1/2/4* are up-regulated in the non-reproductive polyps, while *dkk3* is up-regulated in the adult medusae and reproductive polyps (Additional file 6). Up-regulation of *dkk1/2/4* in the non-reproductive polyps is consistent with findings in *Hydra* [16,24], where *dkk1/2/4* maintains the *Wnt3* expression boundary in the oral end of the polyp. Similarly, specificity of *dkk3* to the medusae and reproductive polyps is consistent with expression in *C. hemisphaerica* where of *dkk3* expression was restricted to differentiating nematoblasts (stem cells that give rise to stinging cells) in the tentacle bulbs in the medusa of *C. hemisphaerica* [56]. Likewise, in *Hydra dkk3* expression is also restricted to nematoblasts, although these are found in the body column of the polyp as *Hydra* lacks a gonophore entirely [57].

Our DE analysis identified 10 *Wnt* genes differentially expressed between *P. carnea* life cycle stages (Additional file 7). Three of these genes are up-regulated in the adult medusa. Interestingly, phylogenetic analysis of these and other hydrozoan *Wnts* reveal that these three medusa-specific *P. carnea Wnts* lack clear orthologous relationships with the 10 well-supported hydrozoan *Wnt* genes (Additional file 7)[51]. These results suggest an expansion in the number of Wnt genes correlated with a fully developed medusa in the *P. carnea* life cycle.

The remaining DE *Wnts* include *Wnt11a*, *Wnt11b, Wnt9/10, Wnt8, Wnt7, Wnt5b,* and *Wnt3*. (Additional file 7). Of these, *Wnt11a, Wnt9/10, Wnt7,* and *Wnt5b* are up-regulated in the non-reproductive polyps, while *Wnt11b* and *Wnt8* are down-regulated in the reproductive or non-reproductive polyps, respectively. Specificity of these *Wnts* with the abovementioned life cycle stages is not inconsistent with previously described expression patterns in *H. echinata* [51] and *Hydra* [19,23]. DE of *Wnt3* is discussed below in detail.

Spatial Expression of Canonical Wnt Components

Of primary interest for this study was Wnt pathway components up-regulated in the reproductive polyp (Table 2; Additional file 5), as this is the stage where the morphogenic patterning of the medusa occurs. KEGG pathway analysis of the DE genes specific to the reproductive polyps mapped to 10 genes in the canonical Wnt signaling pathway, none of which were inhibitors/antagonists (Figure 4C). Specifically, a Wnt ligand and two *frizzled* receptors are up-regulated in the reproductive polyps. Phylogenetic analyses confirm these DE expressed Wnt and *frizzled* genes as orthologs of *Wnt3, frizzled1*, and *frizzled3* (Additional file 7, 8). As these three Wnt signaling genes have been previously implicated in medusae development and evolution [8,14,15,27], we explored their spatial expression in all life cycle stages sampled.

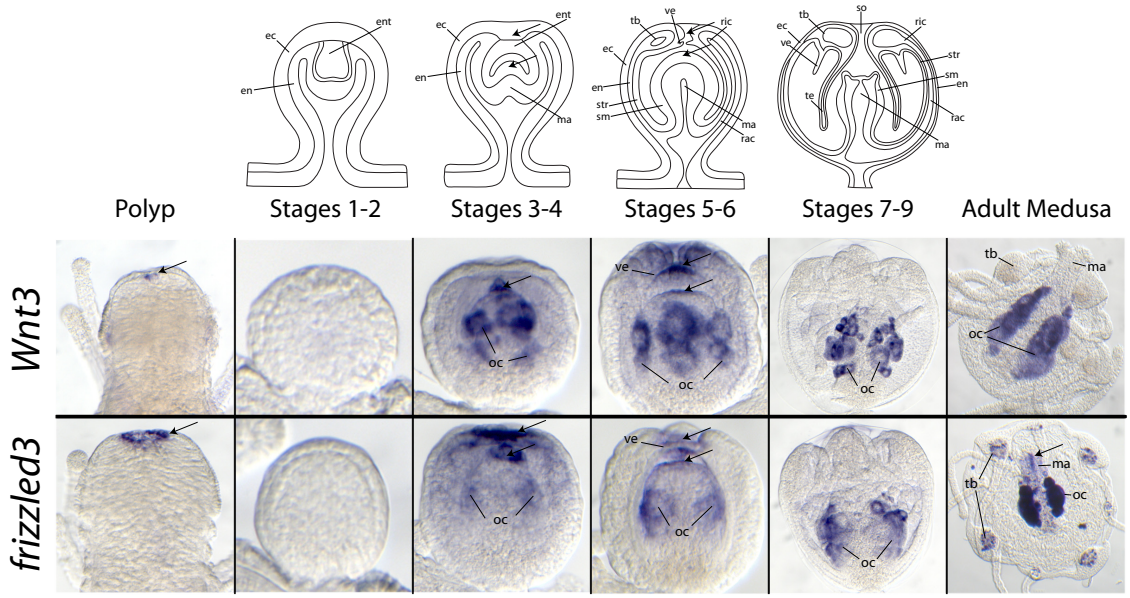


Figure 5. Whole mount *in situ* hybridization of *Wnt3* and *frizzled3*.

Arrows mark expression detected at the distal and oral ends of developmental axes in the polyp, gonophore, and adult medusa. Developing gonophores are staged according to Frey [29] and accompanied by illustrations of cross-sections representing the general development of the medusa bud at that stage. Legend: ec – ectoderm; en – endoderm; ent – entocodon; gt – gametic tissue; ma – manubrium, oc – oocytes; rac – radial canal; ric – ring canal; sm – smooth muscle; spa – spadix; str – striated muscle; tb – tentacle bulb; ve - velum.

As seen in other hydrozoan feeding polyps [11,14-16,19,20], *Wnt3* is expressed in epithelial cells along the oral tip of the hypostome in *P. carnea* (Figure 5). During the earliest stage of gonophore formation, *Wnt3* is absent (Figure 5). By bud stages three and four *Wnt3* expression is observed at the oral end of the developing manubrium as well as in developing oocytes. Expression of *Wnt3* at the oral tip of the developing manubrium continues until stage seven, where expression is from there on limited to developing oocytes (Figure 5). During stages five and six, when the subumbrella is beginning to open at the distal end of the bell margin, *Wnt3* is expressed along the ectodermal tissue surrounding the opening (Figure 5). This is consistent with apoptotic-induced Wnt signaling during head regeneration in *Hydra* [58,59] and normal development in *Hydractinia* [60,61], suggesting further conservation of the role of apoptotic-mediated Wnt signaling in different developmental contexts. By stage seven through medusae liberation, *Wnt3* is not expressed in any tissue except the endodermal tissues surrounding the oocytes of the maturing eggs.

Similar to *Wnt3* expression, ISH detected *frizzled3* expression along the oral tip of the hypostome of the polyp (Figure 5). Throughout medusae bud development, *frizzled3* expression mirrors that of *Wnt3* (Figure 5), except at stages three and four, where *frizzled3* expression precedes *Wnt3* expression at the distal tip of the developing bell. From then on, each gene is co-expressed along the distal and oral regions of each developmental axis, until medusae liberation where there is another noticeable difference in the spatial expression of *frizzled3* and *Wnt3*. While transcripts of both genes are found in the endodermal tissues surrounding maturing oocytes (Figures 5,6A), *frizzled3* is also expressed around differentiating stem cells (presumably nematoblasts) in the tentacle bulb (Figures 5,6B,C) as well as stem cells that accumulate at the distal portion of the mouth after migrating from the aboral end of the manubrium (Figures 5,

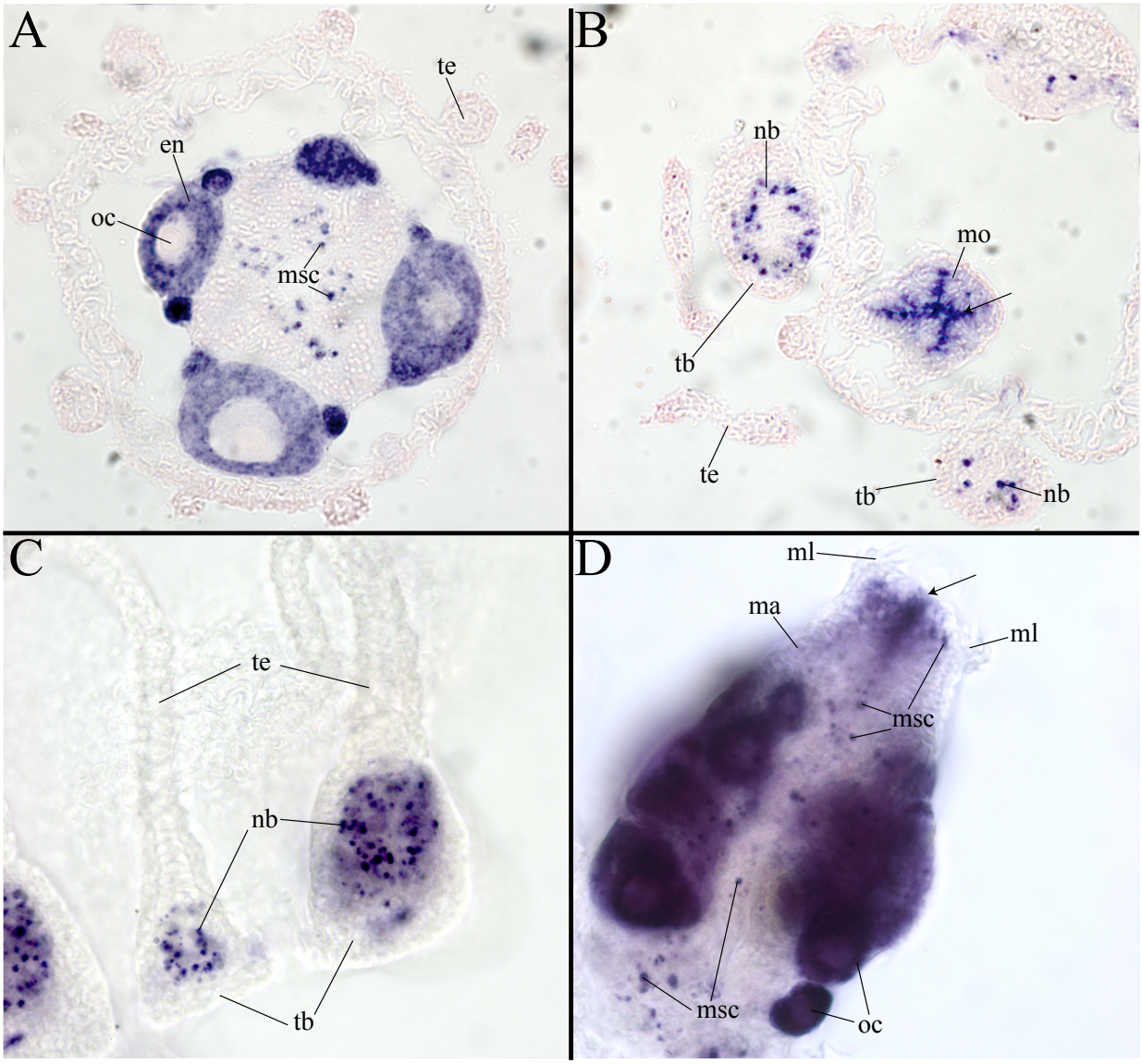

A
te
en
oc
msc
B
nb
mo
tb
te
tb
nb
C
te
nb
tb
D
ml
ma
ml
msc
msc
oc

Figure 6. High-magnification and cross-sections of *frizzled3* ISH in adult medusa.

7μm cross-section of an adult medusae, counter-stained in eosin after *frizzled3* ISH (A,B). Higher magnification images of *frizzled3* ISH of adult medusae (C,D). *Frizzled3* expression observed in endodermal tissue surround mature oocyte (A), in presumptive nematoblasts migrating towards the oral end of the manubrium (A,D) that accumulate at the distal portion of the mouth (B,D), and around differentiating nematoblasts in the tentacle bulb (B,C). Arrows point to oral tip of the manubrium. Legend: en – endoderm; ma – manubrium; ml – manubrium lips; mo – manubrium mouth; msc – migrating stem cells; nb – nematoblasts; oc – oocytes; tb – tentacle bulb; te – tentacles.

6B,D). These presumptive stem cells are most likely nematoblasts moving towards the nematocyst-rich manubrium lips (Figure 6D).

Unlike the other two genes, *frizzled1* expression was not detected in the polyps (not shown). Furthermore, throughout most stages of gonophore development, *frizzled1* expression was not observed. It isn't until stages five and six that ISH revealed expression around cells (presumptive nematoblasts) migrating towards the tentacle bulbs and around developing oocytes (not shown). After medusa liberation, *frizzled1* expression is very similar to that of *frizzled3*, where expression continues around nematoblasts in the tentacle bulbs and developing oocytes (Additional file 9). Expression patterns in early *P. carnea* medusa buds are summarized in Figure 7.

Wnt Expression in Hydrozoan Medusae and Reduced Forms

Previous studies have reported Wnt pathway genes expressed in medusae [8] and reduced forms [8,15,27] (Table 4). *Clytia hemisphaerica*, a leptothecate hydrozoan distantly related to the hydractiniids *Podocoryna* and *Hydractinia* [62], possesses a medusa as part of its life cycle. In the adult medusa, *frizzled1* expression is similar to our finding in *P. carnea*, where it was primarily restricted to the tentacle bulbs (Table 4)[8]. Similarly, adult medusa expression of *frizzled3* in *C. hemisphaerica* is very similar to that reported here in *P. carnea* were it was reported at the oral end of the fully developed manubrium [8]. Where they differ, is at the distal end of the bell margin, as *C. hemisphaerica*'s expression of *frizzled3* was observed in the ring canal [8], and not the tentacle bulbs, while *P. carnea* exhibits the opposite expression. Also, expression of *frizzled3* at the oral ends of the polyp and the lack *frizzled1* expression detected in

the polyp suggest further similarity between *C. hemisphaerica* and *P. carnea* (Table 4)[8]. *Wnt3* expression has not been reported in *C. hemisphaerica.*

Table 4. Summary of reported Wnt signaling expression across hydrozoan life cycle stages.

		Wnt3	***frizzled3***	***frizzled1***
Feeding polyp	*P. carnea*	Oral tip	Oral tip	NONE
	C. hemisphaerica	?	Oral tip	NONE
	E. larynx	Oral tip; tentacle	N/A	Tentacle base
	H. echinata	Oral tip	?	Ubiquitous
	Hydra	Oral tip	N/A	Ubiquitous
Gonophore	*P. carnea*	Distal and oral tips	Distal and oral tips	Tentacle bulbs
	C. hemisphaerica	?	?	?
	E. larynx	Distal tip	N/A	Distal and oral tips
	H. echinata	Distal tip	?	NONE
Adult medusa	*P. carnea*	N/E	Oral tip; tentacle bulbs	Tentacle bulbs
	C. hemisphaerica	?	Oral tip; Ring canal	Tentacle bulbs

? = not studied. NONE = studies and no expression detected in somatic tissues. N/A = gene absent in that taxon.

In the aplanulate hydrozoan *E. larynx*, which is a close relative to *Hydra* [27], co-expression of *Wnt3* and *frizzled1* was observed at the distal end of the gonophore, in both the tentacle buds of the female gonophore and the apical cap of the male gonophore (Figure 7; Table 4)[27]. Furthermore, at the oral end of the reduced manubrium (called a spadix), males expressed a putative Wnt antagonist, *sFRP* (secreted frizzled related protein), while females expressed *frizzled1* in the corresponding region. *Wnt3* was not detected at the oral end of the spadix in either gonophore as it is in the developing manubrium of *P. carnea*, suggesting that down-regulation of *Wnt3* could be involved in the developmental truncation of these gonophores (Figure 7; Table 4). *Ectopleura larynx* appears to lack the receptor *frizzled3* (Additional file 8).

Hydractinia echinata, a close relative of *P. carnea* [28], possesses a sporosac lacking any noticeable medusae characteristics. In *H. echinata*, *Wnt3* expression is restricted to a small patch of cells at the distal tip of the developing sporosac (Figure 7; Table 4)[14,15]. No expression of

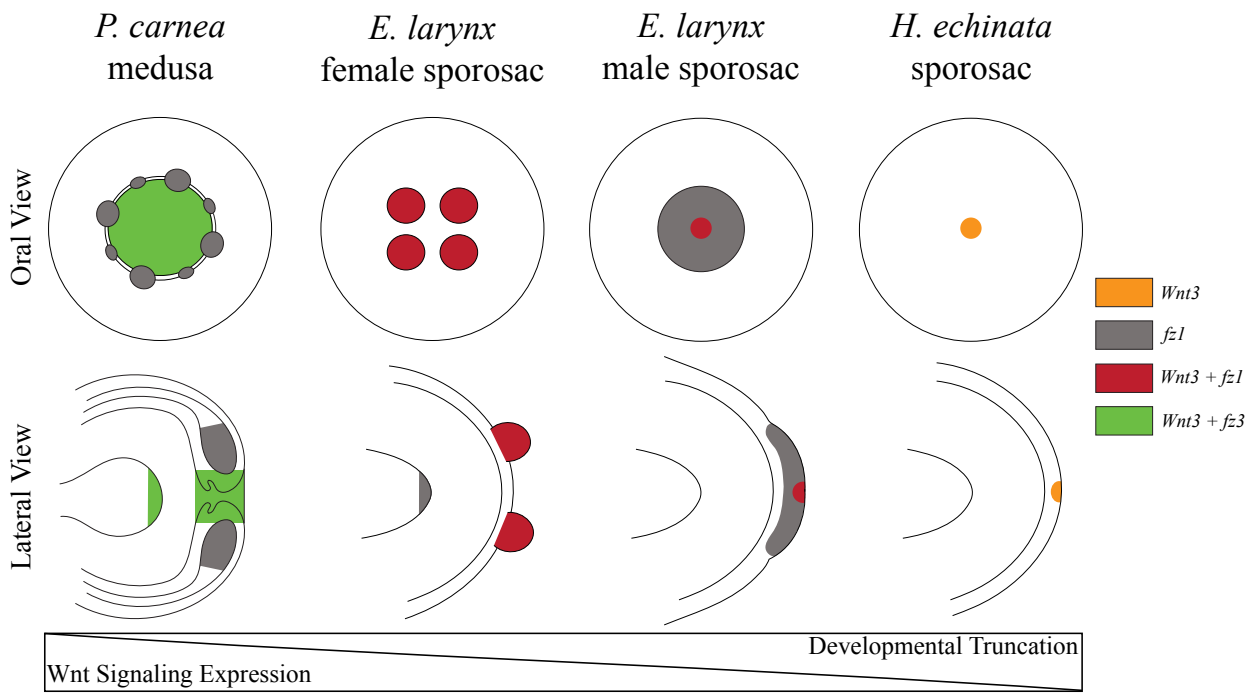


Figure 7. Summary of Wnt signaling expression.

Summary of Wnt signaling expression in three hydrozoan species with different levels of gonophore development. Box on the bottom shows the gradient between Wnt signaling and developmental truncation of the gonophore.

frizzled1 was detected in the sporosac (Figure 7)[14], and *frizzled3* has not been examined in *H. echinata* (Table 4).

DE and Wnt signaling expression patterns presented here shown that co-expression of *Wnt3* ligand and a *frizzled* receptor (*fz1* or *fz3*) are correlated with distal patterning of the medusae bud. The lack of co-expression at the oral end of the spadix of *E. larynx* and *H. echinata*, and at the distal tip of the *H. echinata* sporosac is consistent with developmental truncation of these axes. These expression patterns suggest that down-regulation of Wnt pathway elements could be responsible for the repeated loss of medusae in hydrozoan evolution (Figure 7; Table 4).

Conclusion

While previous studies report evidence that *frizzled1* functions as the canonical Wnt receptor in various developmental contexts [8-11,14,63], here we present repeated co-expression of *Wnt3* and *frizzled3* at the distal/oral ends of each developmental axis of the medusa and the polyp in *P. carnea*. This suggests both lineage and context dependent scenarios where *frizzled1* or *frizzled3* are functioning in canonical Wnt signaling.

Duffy [26] proposed that over evolutionary time, progressive expansion of canonical Wnt signaling resulted in the drastic alterations to the hydrozoan body plan and could have given rise to pelagic (medusa) forms. Our newly generated data in conjunction with previous studies reveal striking similarities in Wnt3 expression between hydrozoan polyps and developing medusae across several hydrozoan taxa (Table 4). In all taxa studied, *Wnt3* is expressed at the oral tip of the polyp. Similarly, in lineages with a gonophore, *Wnt3* is also expressed at the oral and/or distal end of the gonophore axes (Table 4, Figure 7). These patterns suggest co-option of preexisting axial patterning mechanisms in the polyp for the origin of the medusa life cycle stage.

DE analysis between life cycle stages of *P. carnea* revealed an enrichment of differentially expressed canonical and non-canonical Wnt signaling genes, most of which are specific to developing and adult medusae stages. Moreover, ISH reveal the first observed expression pattern *Wnt3* in the developmental axes of medusae. The results presented here provide compelling support the hypothesis of Nawrocki and Cartwright [27] that co-expression of the canonical Wnt ligand, *Wnt3*, and a membrane bound receptor, *frizzled*, at a given developmental axis of the gonophore is necessary for further morphogenic development of the

medusa, and that down-regulation of Wnt pathway elements may be responsible to the loss of the medusa life cycle stage in hydrozoan evolution (Figure 7).

Competing interests

The authors declare they have no competing interests.

Abbreviations

DE: differential expression or differentially expressed; bp: base pair; ISH: whole mount *in situ* hybridization; GO: gene ontology; ORF: open reading frame; HMM: hidden markov model.

Authors' contributions

SMS implemented all bioinformatics analyses, performed ISH, and drafted the manuscript. PC oversaw the design and implementation of the study and helped draft the manuscript. Both authors have read and approve of the final manuscript.

Acknowledgments

We would like to thank: Neil Blackstone and Leo Buss for lab cultures of *P. carnea*, Susan Boyles for illustrations in Figures 1B and 5, Uri Frank for discussion and help with the ISH protocol, and Kirsten Jensen for illustrations in Figure 1A and help with cross-sectioning and counter-staining. This work was supported by NSF grant DEB-095357 to PC.

References

1. Gibbons MJ, Janson LA, Ismail A, Samaai T: **Life cycle strategy, species richness and distribution in marine Hydrozoa (Cnidaria: Medusozoa).** *J. Biogeogr.* 2010, **37**:441-448.
2. Schuchert P: **The European athecate hydroids and their medusae (Hydrozoa, Cnidaria): Filifera part 3.** *Rev. Suisse de Zool.* 2008, **115**:677-757.
3. Cunningham CW, Buss LW: **Molecular evidence for multiple episodes of paedomorphosis in the family Hydractiniidae.** *Biochem. Sys. Ecol.* 1993, **21**:57-69.
4. Leclère L, Schuchert P, Manuel M: **Phylogeny of the Plumularioidea (Hydrozoa, Leptothecata): evolution of colonial organisation and life cycle.** *Zool. Scripta* 2007, **36**:371-394.
5. Leclère L, Schuchert P, Cruaud C, Couloux A, Manuel M: **Molecular phylogenetics of Thecata (Hydrozoa, Cnidaria) reveals long-term maintenance of life history traits despite high frequency of recent character changes.** *Sys. Biol.* 2009, **58**:509-526.
6. Cartwright P, Nawrocki AM: **Character evolution in Hydrozoa (phylum Cnidaria).** *J. Integ. Comp. Biol.* 2010, **50**:456-472.
7. Miglietta MP, Cunningham CW: **Evolution of life cycle, colony morphology, and host-specificity in the Family Hydractiniidae (Hydrozoa, Cnidaria).** *Evolution* 2012, **doi**:10.1111/j.1558-5646.2012.01717.x.
8. Momose T, Houliston E: **Two oppositely localized frizzled RNAs as axis determinants in cnidarian embryo.** *PLoS Biol.* 2007, **5**:e70.
9. Momose T, Derelle R, Houliston E: **A maternally localized Wnt ligand required for axial patterning in the cnidarian *Clytia hemisphaerica*.** *Dev.* 2008, **135**:2105-2113.
10. Amiel A, Houliston E: **Three distinct RNA localization mechanisms contribute to oocyte**

polarity establishment in the cnidarian *Clytia hemisphaerica*. *Dev. Biol.* 2009, **327**:191-203.

11. Plickert G, Jacoby V, Frank U, Müller WA, Mokady O: **Wnt signaling in hydroid development: Formation of the primary body axis in embryogenesis and its subsequent patterning.** *Dev. Biol.* 2006, **298**:368-378.
12. Lapébie P, Ruggiero A, Barreau C, Chevalier S, Chang P, Dru P, Houliston E, Momose T: **Differential responses to Wnt and PCP disruption predict expression and developmental function of conserved and novel genes in a cnidarian.** *PLoS Gen.* 2014, **10**:e1004590.
13. Müller WA, Teo R, Möhrlen F: **Patterning a multi-headed mutant in *Hydractinia*: enhancement of head formation and its phenotypic normalization.** *Int. J. Dev. Biol.* 2004, **48**:9-15.
14. Müller W, Frank U, Teo R, Mokady O, Guette C, Plickert G: **Wnt signaling in hydroid development: ectopic heads and giant buds induced by GSK-3*ß* inhibitors.** *Int. J. Dev. Biol.* 2007, **51**:211-220.
15. Duffy DJ, Plickert G, Kuenzel T, Tilmann W, Frank U: **Wnt signaling promotes oral but suppresses aboral structures in *Hydractinia* metamorphosis and regeneration.** *Dev*. 2010, **137**:3057-3066.
16. Guder, C, Pinho S, Nacak TG, Schmidt HA, Hobmayer B, Niehrs C, Holstein TW: **An ancient Wnt-Dickkopf anatagonism in *Hydra*.** *Dev*. 2005, **133**:901-911.
17. Guder C, Philipp L, Lengfeld T, Watanabe H, Hobmayer B, Holstein TW: **The Wnt code: cnidarians signal the way.** *Oncogene* 2006, **25**:7450–7460.
18. Hobmayer B, Rentzsch F, Kuhn K, Happel CM, von Laue CC, Snyde P, Rothbacher U,

Holstein TW: **WNT signalling molecules act in axis formation in the diploblastic metazoan Hydra.** *Nature* 2000, 407:186–189.

19. Lengfeld T, Watanabe H, Simakov O, Lindgens D, Gee L, Law L, Schmidt HA, Özbek S, Bode H, Holstein TW: **Multiple Wnts are involved in *Hydra* organizer formation and regeneration.** *Dev. Biol.* 2009, **330**:186-199.
20. Gee L, Hartig J, Law L, Wittlieb J, Khalturin K, Bosch TCG, Bode HR: ***ß*-catenin plays a central role in setting up the head organizer in hydra.** *Dev. Biol.* 2010, **340**:116-124.
21. Broun M, Gee L, Reinhardt B, Bode HR: **Formation of the head organizer in hydra involves the canonical Wnt pathway.** *Dev*. 2005, **132**:2907-2916.
22. Nakamura Y, Tsiairis CD, Özbek S, Holstein TW: **Autoregulatory and repressive inputs localize *Hydra Wnt3* to the head organizer.** *PNAS* 2011, **108**:9137-9142.
23. Philipp I, Aufschnaiter R, Özbek S, Pontasch S, Jenewein M, Watanabe H, Rentzsch F, Holstein TW, Hobmayer B: **Wnt/*ß*-catenin and noncanonical Wnt signaling interact in tissue evagination in the simple eumetazoan *Hydra*.** *PNAS* 2009, **106**:4290-4295.
24. Augustin R, Franke A, Khalturin K, Kiko R, Siebert S, Hemmrich G, Bosch TCG: **Dickkopf related genes are compnents of the positional value gradient in Hydra.** *Dev. Biol.* 2006, **296**:62-70.
25. Khalturin K, Anton-Erxleben F, Milde S, Plötx C, Wittlieb J, Hemmrich G, Bosch TCG: **Transgenic stem cells in *Hydra* reveal an early evolutionary origin for key elements controlling self-renewal and differentiation.** *Dev. Biol.* 2007, **309**:32-44.
26. Duffy DJ: **Modulation of Wnt signaling – A route to speciation?** *Comm. Integ. Biol.* 2011, **4**:59-61.

27. Nawrocki AM, Cartwright P: **Expression of Wnt pathway genes in polyps and medusa-like structures of *Ectopleura larynx* (Cnidaria: Hydrozoa).** *Evol. & Dev.* 2013, **15**:373-384.

28. Miglietta MP, Schuchert P, Cunningham CW: **Reconciling genealogical and morphological species in a worldwide study of the Family Hydractiniidae (Cnidaria, Hydrozoa).** *Zool. Scripta* 2009, **38**:403-430.

29. Frey J: **Die Entwicklungsleistungen der Medusenknospen und Medusen von *Podocoryne carnea* M. Sars nach Isolation und Dissoziation.** *Wilhelm Roux' Arch Entwicklungsmech. Org.* 1968, **168**:428-464.

30. Bolger AM, Lohse M, Usadel B: **Trimmomatic: A flexible trimmer for Illumina sequence data.** *Bioinformatics* 2014, **30**:2114-2120.

31. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A: **Full-length transcriptome assembly from RNA-seq data without a reference genome.** *Nat. Biotechnol.* 2011, **29**:644-52.

32. **NCBI UniVec [**http://www.ncbi.nlm.nih.gov/tools/vecscreen/univec]

33. Sanders SM, Shcheglovitova M, Cartwright P: **Differential gene expression between functionally specialized polyps of the colonial hydrozoan *Hydractinia symbiolongicarpus* (Phylum Cnidaria).** *BMC Genomics* 2014, **15**:406.

34. Conesa A, Götz S, Garcia-Gomez JM, Terol J, Talon M, Robles M: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.** *Bioinformatics* 2005, **21**:3674-3676.

35. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A: **High-throughput functional annotation and data mining with the Blast2GO suite.** *Nucl. Acids Res.* 2008, **36**:3420-3435.

36. Parra G, Bradnam K, Korf I: **CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes.** *Bioinformatics* 2007, **23**:1061-1067.

37. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD: **The Pfam protein families database.** *Nucl. Acid Res.* 2012, **Database Issue**:doi:10.1093/nar/gkr1065.

38. **TIGR Database** [http://blast.jcvi.org/web-hmm/]

39. Min XJ, Butler G, Storms R, Tsang A: **OrfPredictor: predicting protein-coding regions in EST-derived sequences.** *Nucleic Acids Res.* 2005, **Web Server Issue**:W677-W680. [http://proteomics.ysu.edu/tools/OrfPredictor.html]

40. **HMMER** [http://hmmer.org/]

41. Li L, Stoeckert CJ, Roos DS: **OrthoMCL: Identification of orthologous groups for Eukaryotic genomes.** *Genome Res.* 2003, **13**:2178-2189.

42. Li B, Dewey CN: **RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome.** *BMC Bioinform.* 2011, **12**:323.

43. Leng N, Dawson JA, Thomson HA, Ruotti V, Rissman AI, Smits BMG, Haag JD, Gould MN, Stewart RM, Kendziorski C: **EBSeq: an emipirical Bayes hierarchical model for inference in RNA-seq experiments.** *Bioinform.* 2013, **29**:1035-1043.

44. Li W, Godzik A: **Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences.** *Bioinformatics* 2006, **22**:1658-1659.

45. Fu L, Niu B, Zhu Z, Wu S, Li W: **CH-HIT: accelerated for clustering the next generation sequencing data.** *Bioinformatics* 2012, **28**:3150-3152.

46. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M: **Data, information, knowledge and principle: back to metabolism in KEGG.** *Nucl. Acids Res.* 2014, **42**:D199-D205.

47. Kanehisa M, Goto S: **KEGG: Kyoto Encyclopedia of Genes and Genomes.** *Nucl. Acids Res.* 2000, **28**:27-30.

48. Gajewsky M, Leitz T, Schlossherr J, Plickert G: **LWamides from Cnidaria constitute a novel family of neuropeptide with morphogenetic activity.** *Roux's Arch. Dev. Biol.* 1996, **205**:232– 242.

49. Hemmrich G, Bosch TCG: **Compagen, a comparative genomics platform for early branching metazoan animals, reveals early origins of genes regulating stem-cell differentiation.** *BioEssays* 2008, **30**:1010-1018.

50. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG: **Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.** *Mol. Sys. Biol.* 2011, **7**:539.

51. Hensel K, Lotan T, Sanders SM, Cartwright P, Frank U: **Lineage-specific evolution of cnidarian Wnt ligands.** *Evol. & Dev.* 2014, **16**:259-269.

52. Katoh K, Kuma K, Toh H, Miyata T: **MAFFT version 5: improvement in the accuracy of multiple sequence alignment**. *Nucl. Acids Res.* 2005, **33**:511-518.

53. Stamatakis A, Hoover P, Rougemont J: **A fast bootstrapping alogorithm for the RAxML web-servers.** *Sys. Biol.* 2008, **57**:758-771.

54. Miller MA, Pfeiffer W, Schwartz T: **Creating the CIPRES Science Gateway for inference of large phylogenetic trees.** *GCE* 2010, 1-8.

55. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GBEH00000000.

56. Denker E, Manuel M, Leclère L, Guyader H, Rabet N: **Ordered progression of nematogenesis from stem cells through differentiation stages in the tentacle bulb of *Clytia hemisphaerica* (Hydrozoa, Cnidaria).** *Dev. Biol.* 2008, **315**:99-113.

57. Fedders H, Augustin R, Bosch TCG: **A *Dickkopf-3*-related gene is expressed in differentiating nematocytes in the basal metazoan *Hydra*.** *Dev. Genes Evol.* 2004, **214**:72-80.

58. Chera S, Ghila L, Dobretz K, Wenger Y, Bauer C, Buzgariu W, Martinou JC, Galliot B: **Apoptotic cells provide an unexpected source of Wnt3 signaling to drive *Hydra* head regeneration.** *Dev. Cell* 2009, **17**:279-289.

59. Galliot B, Chera S: **The *Hydra* model: disclosing an apoptosis-driven generator of Wnt-based regeneration.** *Trends in Cell Bio.* 2010, **20**:514-523.

60. Duffy DJ, Millane RC, Frank U: **A heat shock protein and Wnt signaling crosstalk during axial patterning and stem cell proliferation.** *Dev. Biol.* 2012, **362**:271-281.

61. Duffy DJ: **Instructive reconstruction: A new role for apoptosis in pattern formation.** *Bioessays* , 2012 **34**:561-564.

62. Cartwright P, Evans NM, Dunn CW, Marques AC, Miglietta MP, Schuchert P, Collins AG: **Phylogenetics of Hydroidolina (Hydrozoa: Cnidaria).** *J. Mar. Biol. Assoc.* 2008, **88**:1163-1172.

63. Teo R, Möhrlen F, Plickert G, Müller WA, Frank U: **An evolutionary conserved role of Wnt-signaling in stem cell fate decision.** *Dev. Biol.* 2006, **289**:91-99.

Additional files

Additional file 1. Transcriptome Annotations.

Tab delimited text file containing Blast2Go, HMM, and orthoMCL annotations for the filtered transcriptome.

Additional file 2. CEGMA output.

Compressed (.tgz) file containing output from CEGMA analysis of *P. carnea*'s transcriptome assembly.

Additional file 3. All functionally annotated Wnt signaling transcripts.

Tab delimited text file containing the Blast2Go, HMM, and orthoMCL annotations of all functionally annotated Wnt signaling transcripts.

Additional file 4. All DE genes with annotations.

Tab delimited text file containing the mean expression for each life cycle stage sampled, the inferred posterior probability of differential expression (PPDE), life cycle specificity, and Blast2Go and HMM annotations for all DE genes. The 'Pattern' column indicates which life cycle stage each gene is specific to and whether UP or DOWN regulated in that stage. As DE was performed at the gene-level expression (as estimated by RSEM), annotations cover all isoforms of each gene.

Additional file 5. All DE Wnt signaling genes with annotations.

Tab delimited text file containing the mean expression for each life cycle stage sampled, the inferred posterior probability of differential expression (PPDE), life cycle specificity, and Blast2Go and HMM annotations for all DE genes functionally annotated with Wnt signaling. The 'Pattern' column indicates which life cycle stage each gene is specific to and whether UP or DOWN regulated in that stage. As DE was performed at the gene-level expression (as estimated by RSEM), annotations cover all isoforms of each gene.

Additional file 6. *Dickkopf* (*dkk*) gene tree.

Phylogeny of cnidarian *dickkopf* genes. Highlighted in red are sequences from *P. carnea*'s assembly. Maximum likelihood bootstrap support values given at the base of each node. Values not shown for nodes with less than 70% support.

Additional file 7. Wnt gene tree.

Phylogeny of hydrozoan Wnt genes from Hensel et al. [51] and Wnt genes DE between *P. carnea* life-cycle stages. Highlighted in red are DE genes from *P. carnea*'s assembly. Boxed off sequences indicate those analyzed with ISH. Arrows mark *Wnt* genes that are not orthologous to the 10 well-supported hydrozoan *Wnts*, while '*' marks sequences orthologous to those reported by Hensel et al. [51]. To be consistent with Hensel et al. [51], the tree is rooted on the branch leading to *Wnt9/10*. Maximum likelihood bootstrap support values given at the base of each node. Values not shown for nodes with less than 70% support.

Additional file 8. *Frizzled* gene tree.

Phylogeny of medusozoan *frizzled* genes. Highlighted in red are sequences from *P. carnea*'s assembly. Boxed off sequences indicate those analyzed with ISH. Maximum likelihood bootstrap support values given at the base of each node. Values not shown for nodes with less than 70% support.

Additional file 9. ISH of *frizzled1* in the adult medusa.

Legend: ma – manubrium; oc – oocytes; tb – tentacle bulb.

CHAPTER 3:

Interspecific differential expression analysis of RNA-Seq data yields insight into medusae evolution in hydrozoans (Phylum Cnidaria)

Abstract

The evolution of the hydrozoan medusa (jellyfish), and its reduced forms, has been a widely examined and controversial topic in invertebrate zoology. Although many hydrozoans alternate between asexually reproducing polyps and sexually reproducing medusae, most species display developmentally truncated forms of the medusa stage. While evolutionary transitions in medusa truncation and loss have been investigated phylogenetically, little is known about the genetic pathways involved. Here we present a new workflow for evaluating differential expression (DE) between species using short read Illumina RNA-seq data. Through interspecific DE analyses between the closely related hydrozoans, *Hydractinia symbiolongicarpus* and *Podocoryna carnea*, we identified genes potentially involved in the evolutionary transition between fully developed medusa and truncated sporosacs. A dataset of 10,942 orthologous *H. symbiolongicarpus* and *P. carnea* genes was constructed from *de novo* assemblies of short read Illumina data. Differential expression analysis identified 941 orthologs that were differentially expressed in *P. carnea* developing and adult medusa when compared to *H. symbiolongicarpus* sporosacs. In addition, two genes with no corresponding ortholog in *H. symbiolongicarpus* were found to be expressed in developing medusa of *P. carnea*. Results presented here show interspecific differential expression analyses of RNA-seq data to be a sensitive and reliable method for identifying genes and gene pathways potentially involved in morphological and life cycle differences between species. Our study finds that while differential regulation of orthologous genes can help explain differences in homologous structures, evolutionary shifts in Hydrozoa life cycles are also likely accompanied by gene gain or loss.

Introduction

In Hydrozoa (Phylum Cnidaria), many species exhibit an alternation of generations, where asexually reproducing polyps give rise to sexually reproducing jellyfish (medusae). However, across hydrozoans there is much variation in this sexually reproducing life cycle stage. In most hydrozoan species (~70%), development of the medusa bud (gonophore) is truncated to some degree or entirely absent (Leclère et al. 2009; Cartwright and Nawrocki 2010; Gibson et al. 2010). In these taxa, sexual maturity is reached in a gonophore that resembles an early ontogenetic stage of medusae development. The degree of gonophore development ranges from completely reduced structures called sporosacs that lack any resemblance of the medusa (Figure 1A), to more developed forms called medusoids, that may or may not detach and swim, but lack the ability to feed (not shown), to the fully developed medusa stage that detaches from the hydroid polyp and can feed, swim, and sexually reproduce in the water column (Figure 1B).

The evolution of this structure and its reduced forms has been a topic of investigation for the last 150 years (Allman 1864; Cornelius 1992; Cunningham and Buss 1993; Marques and Migotto 2001; Leclère et al. 2007, 2009; Miglietta et al. 2009, 2010; Cartwright and Nawrocki 2010; Miglietta and Cunningham 2012). Phylogenetic studies have revealed multiple independent losses of medusae (Cunningham and Buss 1993; Leclère et al. 2007, 2009; Cartwright and Nawrocki 2010; Miglietta and Cunningham 2012), and possibly even re-gain Leclère et al. 2009; Cartwright and Nawrocki 2010; Miglietta and Cunningham 2012). While phylogenies are important for recognizing evolutionary patterns of character transitions, understanding complex patterns of character loss and possible re-gain will come from insight about their development. Specifically, maintenance of developmental regulatory pathways underlying medusae ontogeny in reduced forms, could add support to arguments for medusae re-

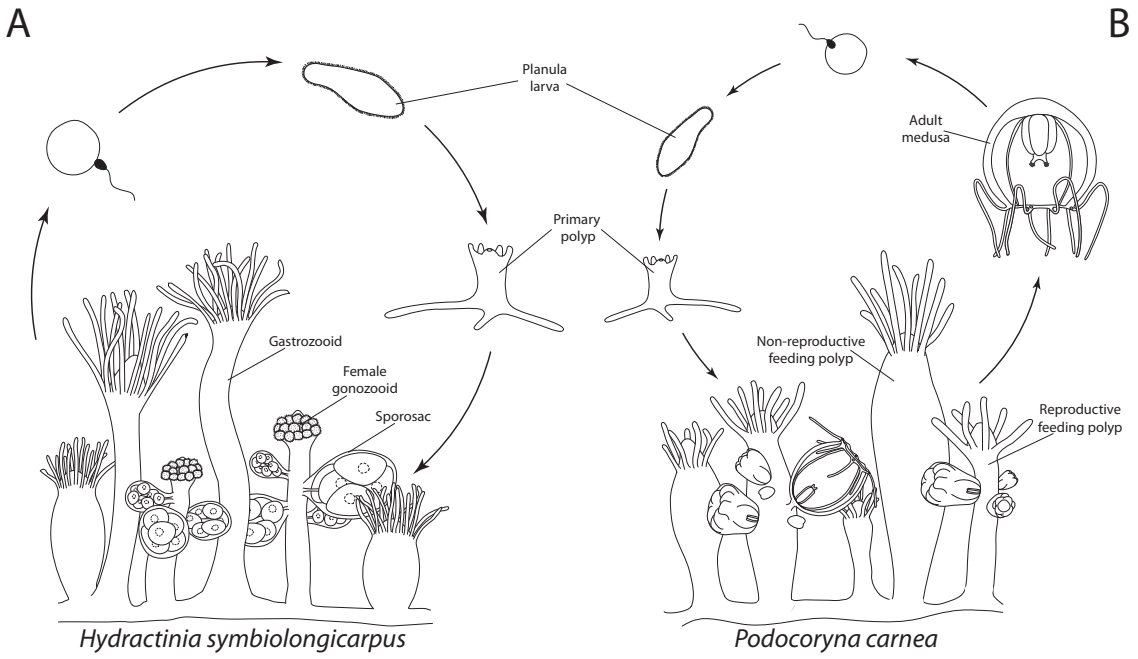


Figure 1. Illustration of hydrozoan life cycles.

1A) In the life cycle of *H. symbiolongicarpus* gonophores develop into sporosacs that lack all medusa features and remain attached to the colony on specialized reproductive polyps called gonozooids. Sexual reproduction occurs in the water column after the sporosacs release their gametes. Sexual reproduction results in a planula larva that eventually settles onto a suitable substrate and metamorphoses into a primary polyp. This polyp will asexually produce other polyps to form a colony and the cycle repeats. 1B) *Podocoryna carnea'*s life cycle is similar to that of *H. symbiolongicarpus* except that medusae asexually bud from reproductive polyps and detach from the colony to sexually reproduce in the water column.

evolving in the Hydrozoa. The hydrozoan family Hydractiniidae provides an excellent system for identifying key components in the development and evolution of both medusa truncation and possible re-evolution, as the entire spectrum of gonophore development is exhibited within this group. The hydractiniid species *H. symbiolongicarpus* and *P. carnea* exhibit either ends of this developmental spectrum, possessing a sporosac (Figure 1A) and medusa (Figure 1B), respectively.

Now that transcriptomes of non-traditional model systems can be readily obtained and characterized in different stages or parts of an organism (Hao et al. 2011; Siebert et al. 2011; Helm et al. 2013; Sanders et al. 2014; Schunter et al. 2014), comparing transcriptomes between species is the obvious next step. Dunn et al. (2013b) extensively reviewed the utility of comparative expression across multiple species, as well as its challenges. While not as abundant as intraspecific transcriptomic studies, interspecific analyses have proved illuminating on a diversity of topics (Yang and Wang 2013; Boyle et al. 2014; Pankey et al. 2014). These studies took a general approach to comparing whole transcriptomes, but did not apply interspecific differential expression (DE) in an unbiased approach to identify genes potentially involved in differences between species.

Here we present a workflow for performing differential expression analyses between species from short read Illumina RNA-Seq data. Specifically, we use newly generated RNA-Seq data from *P. carnea* and previously published data from *H. symbiolongicarpus* (Sanders et al. 2014), to identify genes and gene pathways that are potentially involved in the evolutionary transition between truncated and fully developed medusae in the Hydractiniidae.

Results and Discussion

Library Summary

Colonies of *H. symbiolongicarpus* possess morphologically and functionally distinct polyp types. Raw Illumina RNA-Seq data for *H. symbiolongicarpus* (SRA archive SRX474462) were generated from four different polyp types, representing four different conditions: gastrozooids (feeding polyps), dactylozooids (defensive polyps); gonozooids (reproductive polyps) bearing male gonophores; and gonozooids bearing female gonophores (Table 1). For *P. carnea*, three tissue types were sampled: non-reproductive feeding polyps (gastrozooids), medusa-budding feeding polyps, and free-living medusae. At the time of this study we did not have access to male colonies, so only females were sampled for *P. carnea*.

Table 1. RNA-Seq Illumina libraries.

Condition	# replicates
H. symbiolongicarpus[1]	
gastrozooid	4
*female gonozooid	2
*male gonozooid	2
P. carnea	
non-reproductive polyp	3
*medusae-budding polyp	4
*adult medusae	3

[1] The other condition, dactylozooid was included in the assembly but not in the DE analysis.*Conditions of interest.

Transcriptome Assembly and Annotation, Enrichment Analyses, and Orthology Prediction

All of the libraries (described above; Table 2) were used to maximize transcript discovery in the assembly. Final assemblies consisted of 127,716 and 178,396 transcripts for *H. symbiolongicarpus* and *P. carnea*, respectively (Table 2, Figure 2).

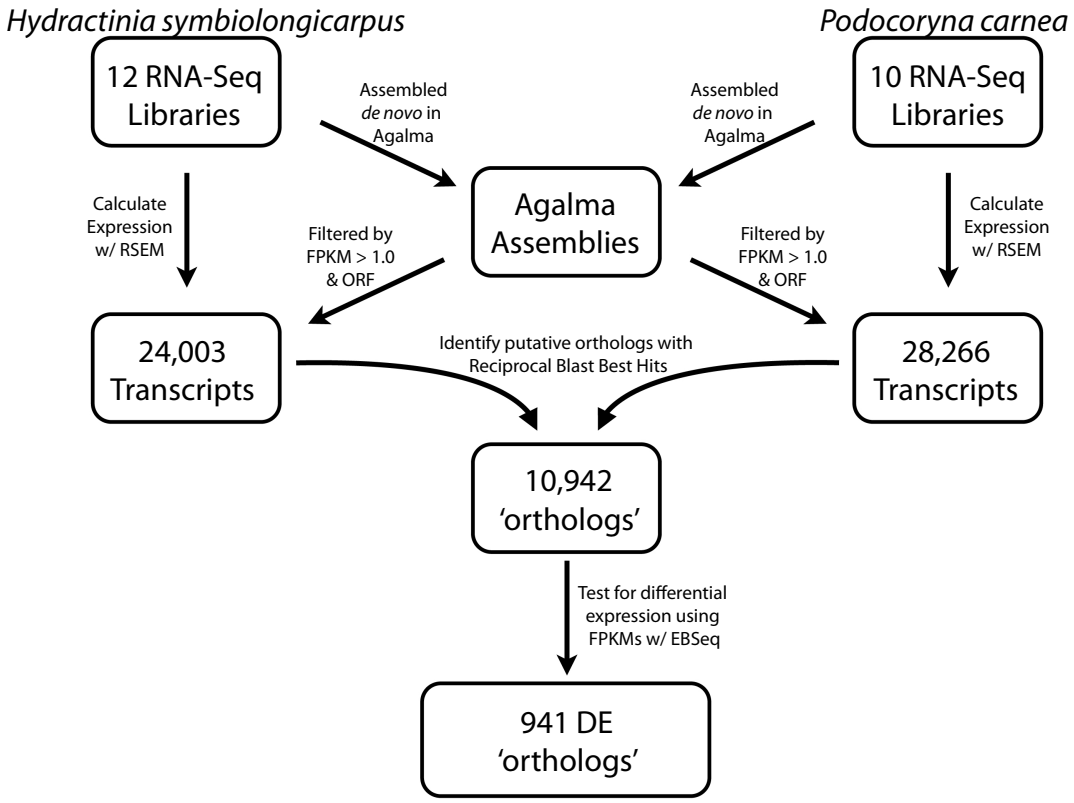


Figure 2. Schematic of bioinformatics workflow.

Initial transcriptomes for *H. symbiolongicarpus* and *P. carnea* are assembled from 12 and 10 100bp paired-end Illumina libraries respectively, using the program Agalma. Each transcriptome is filtered for transcripts that meet the Transdecoder reading frame criteria (as implemented in Agalma) and have an FPKM $\geq$ 1.0. Expression values are estimated for these remaining transcripts for each library independently using RSEM. Orthologs are identified using one-to-one reciprocal blast best hits between the Transdecoder protein translations of the subsetted transcriptome using Blastp under default setting. Differential expression analyses are performed with EBSeq using FPKM and TPM expression normalizations.

Table 2. Assembly statistics summary.

	H. symbiolongicarpus			*P. carnea*		
	Initial	**Filtered**	**RBBH**	**Initial**	**Filtered**	**RBBH**
# Transcripts	**127,716**	**24,003**	**10,942**	**178,396**	**28,266**	**10,942**
N25 (bp)	**3,960**	**4,389**	**4,536**	**3,342**	**3,890**	**4,341**
N50 (bp)	**2,459**	**2,894**	**2,990**	**1,977**	**2,610**	**2,881**
N75 (bp)	**1,290**	**1,928**	**2,015**	**945**	**1,784**	**1,957**
GC Content	**35.40%**	**36.50%**	**36.72%**	**38.60%**	**38.32%**	**36.58%**

Initial = transcriptomes assembled with Agalma. Filtered = transcripts remaining after transcriptomes were filtered by FPKM and Transdecoder reading frame criteria. RBBH = transcripts with a one-to-one reciprocal best blast hit match between the two filtered transcriptomes.

Gene ontology (GO) analyses identified a total of 11,196 and 16,386 unique GO terms and Hidden Markov Model (HMM) domains in at least one transcriptome, respectively. As a means of identifying candidate 'medusa' genes, enrichment analyses (Fisher's exact test) were performed on the abundance of each GO term and HMM domain in either assembly. Gene ontology enrichment analyses did not identify any over abundant GO terms in *P. carnea*'s transcriptome when compared to the total number of GO terms for each species combined. Similarly, enrichment analyses of HMM domains did not identify any over-represented domains in the *P. carnea*'s transcriptome. By contrast, 110 GO terms and 27 HMM domains were over-represented in *H. symbiolongicarpus*' transcriptome. This is most likely due to the inclusion of dactylozooids and both male and female gametic tissues in the assembly for *H. symbiolongicarpus,* whereas *P. carnea* dactylozooids and male gametic tissue were not sampled. Since neither sets of enrichment analyses yielded insight into gene and/or signaling pathways involved in medusa development and evolution, we performed interspecific differential expression analyses to detect quantitative differences in gene expression levels associated with the phenotypic differences between these species' gonophores.

When comparing gene expression between species, the first critical step is to establish robust orthology assignments. In order to avoid artifactual differences due to different assembly methods, each transcriptome was assembled *de novo* with the pipeline Agalma (Dunn et al. 2013a) under identical settings, as opposed to using a previously published genome-guided transcriptome for *H. symbiolongicarpus* (Sanders et al. 2014). Of further concern, is the effect of polymorphisms on transcript/gene redundancy in the assembly. Polymorphisms (common in data collected from non-inbred lines) can lead to an increase in the number of paths to reconcile during the assembly process, thus increasing the number of fragmented and rare variants of a transcript/gene. In order to minimize the number of fragmented and redundant transcripts, each assembly was filtered for transcripts with a minimum relative expression level (FPKM $\geq$ 1.0) and reading frame criteria prior to orthology prediction.

After initial filtering, approximately 24K and 28K transcripts remained (referred to from here on as the filtered transcriptomes) in the *H. symbiolongicarpus* [cite assembly accession] and *P. carnea* [cite assembly accession] assemblies, respectively (Table 2). This reduction in transcript number greatly reduced the differences between each transcriptome assemblage characteristics, including the distribution of transcript size, N50, and GC content (Table 2, Figure 2-3). Most importantly, removing incomplete transcripts and under-represented variants increases our confidence in the transcripts remaining for orthology prediction and subsequently, the reliability of the inferred relative expression of each predicted ortholog. A total of 10,942 putative orthologs were identified (Tables 2,S1,S2) between our filtered assemblies. Not surprisingly, the resulting orthologous gene data set further decreased the disparity between the summary statistics for each species (Table 2, Figure 3).

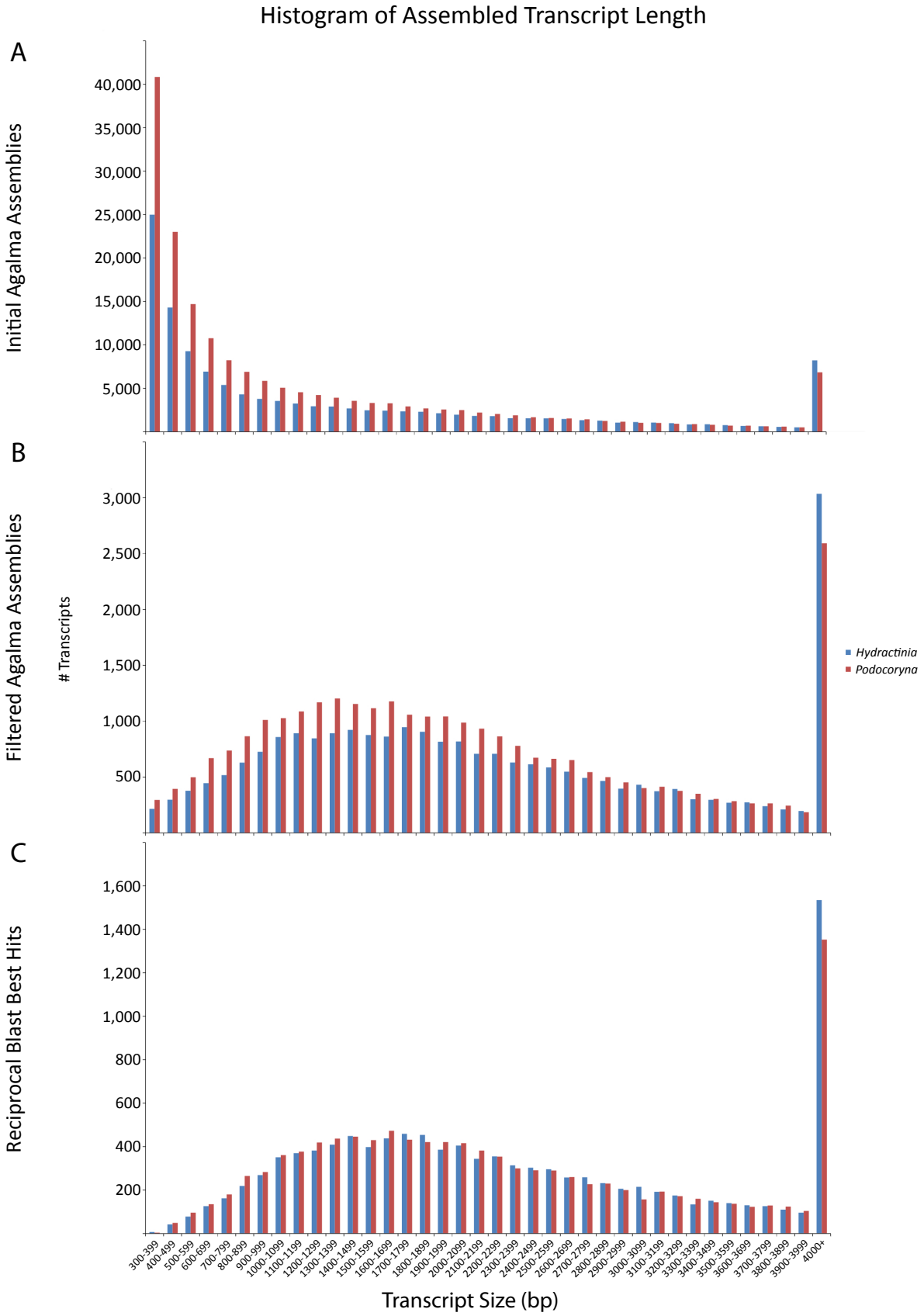

Histogram of Assembled Transcript Length
A
Initial Agalma Assemblies
40,000
35,000
30,000
25,000
20,000
15,000
10,000
5,000
B
Filtered Agalma Assemblies
Transcripts
3,000
2,500
2,000
1,500
1,000
500
Hydractinia
Podocoryna
C
Reciprocal Blast Best Hits
1,600
1,400
1,200
1,000
800
600
400
200
300-399
400-499
500-599
600-699
700-799
800-899
900-999
1000-1099
1100-1199
1200-1299
1300-1399
1400-1499
1500-1599
1600-1699
1700-1799
1800-1899
1900-1999
2000-2099
2100-2199
2200-2299
2300-2399
2400-2499
2500-2599
2600-2699
2700-2799
2800-2899
2900-2999
3000-3099
3100-3199
3200-3299
3300-3399
3400-3499
3500-3599
3600-3699
3700-3799
3800-3899
3900-3999
4000+
Transcript Size (bp)

Figure 3. Distributions of transcript size.

Histograms of the: A) Initial Agalma assembled transcriptomes; B) Assemblies filtered by FPKMs and reading frame criteria; C) Transcripts with a one-to-one reciprocal blast best hit (orthologs). Red = *P. carnea*; Blue = *H. symbiolongicarpus*. X-axis is constant. Y-axis changes with each assembly.

Differential Expression Results

Two separate expression matrices, FPKM (fragments per kilobase of transcript per million mapped reads) and TPM (transcripts per million), were generated for the 10,942 orthologs by RSEM (Li and Dewey 2011) and analyzed with EBSeq (Leng et al. 2013). EBSeq is an ideal software for assessing DE between species as EBSeq's false discovery rate (FDR) and statistical power have been shown to be less sensitive to overdispersal of expression values between conditions when compared to other DE software (Leng et al. 2013). Furthermore, EBSeq tests for differential expression between multiple conditions (i.e. tissue types) simultaneously, by assigning a posterior probability to each possible expression pattern in an enumerated list of all possible expression patterns, given a set of conditions. These patterns are defined as the unique combination of significant differences in expression values between a given number of conditions. As more conditions are present, the number of possible patterns increases.

To take advantage of this feature, we included four conditions that have a gonophore stage present (male and female gonozooids in *H. symbiolongicarpus*, medusa-budding polyps and free-living medusa in *P. carnea*), as well as non-reproductive gastrozooids from each species; see description above). Only the dactylozooid libraries were excluded as no such homologous tissue was sampled in *P. carnea* (Table 1). Inclusion of the non-reproductive tissue types increases the complexity of the expression landscape within and between each species (i.e. more patterns), effectively increasing the power of the DE analysis. With six conditions in the analyses (Table 1), EBSeq identified a total of 203 possible expression patterns, although not all of them are informative to our question (Figure 4A; Table S3). This is another advantage of EBSeq, as the researcher can insert biologically relevant constraints on expression patterns *a*

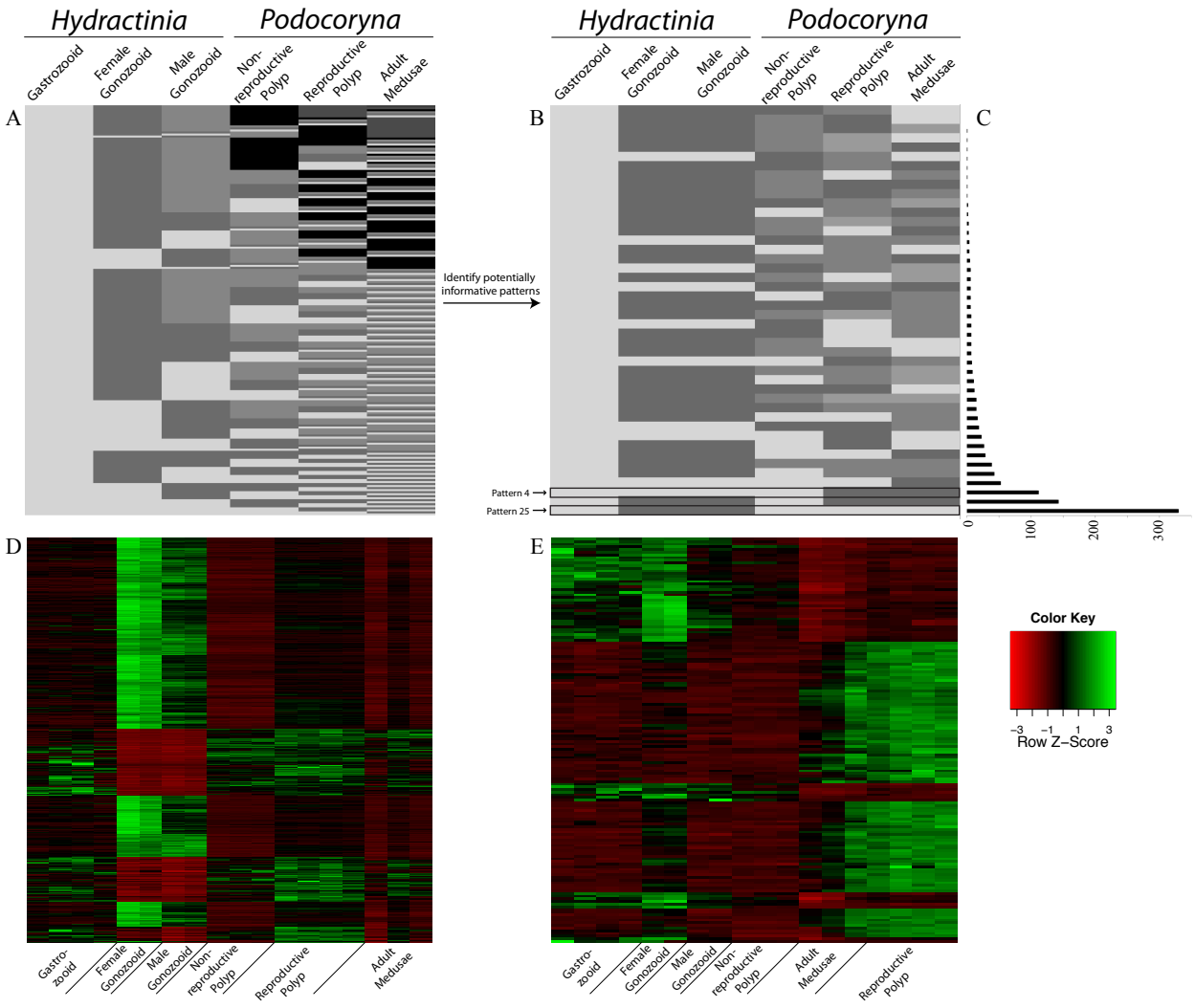

Hydractinia
Podocoryna
Gastrozooid
Female Gonozooid
Male Gonozooid
Non-reproductive Polyp
Reproductive Polyp
Adult Medusae
A
B
C
Identify potentially informative patterns
Pattern 4
Pattern 25
0
100
200
300
D
E
Color Key
−3
−1
1
3
Row Z−Score

Figure 4. Pattern reduction to informative patterns.

A) With six conditions (columns) present in the DE analysis, EBSeq identifies 203 possible expression patterns (rows). B) Using biologically relevant constraints on expression in an attempt to reduce the noise in the DE signal, the number of patterns is reduced to 44 potentially informative patterns. Colors in this schematic do not indicate magnitude of expression, just non-directional levels of expression to show statistically equivalent and non-equivalent levels of expression between conditions in the analysis. C) Bar graph of the number of DE genes (FDR < 0.05) specific to each 44 of the potentially informative patterns. D) Z-normalized heatmap of all orthologs whose expression is consistent with 'Pattern 25' in the FPKM dataset, an expression that should contain sporosac-specific orthologs. E) Z-normalized heatmap of all orthologs whose expression is consistent with 'Pattern 4' in the FPKM dataset, an expression that should contain genes specific to *P. carnea* reproductive polyps and adult medusae.

priori, retaining only those patterns that are specific to the conditions of interest for that particular study (Figure 4B).

Capitalizing on this aspect, we identified 44 expression patterns that were potentially informative, greatly reducing the number of potential results (Figure 4B; Table S4). These patterns were selected with an initial constraint that *Hydractinia* male and female gonozooid expressions are statistically equivalent. Following assumptions made by Sanders et al. (2014), transcripts differentially expressed between male and female gonozooids can be attributed to differences in gametogenesis (either spermatogenesis or oogenesis). Since only female gametic tissues were sampled in *P. carnea*, gene expression driven by maternal transcript generation during oogenesis will be highly expressed in the budding and adult medusae, potentially skewing the DE results. Assuming maternally loaded genes are conserved between closely related species, patterns where expression of male and female sporosacs are not statistically equivalent are removed, thus reducing the number of patterns to 52. Further pattern restrictions were added to increase the chance of finding developing gonophore and/or medusa-specific genes (i.e. patterns where expression is statistically equivalent between non-homologous interspecific tissue samples but are differentially expressed between non-homologous intraspecific samples were removed), resulting in a total of 44 potentially informative patterns (Figure 4B; Table S4).

Approximately 75% (8,237) of the orthologs are recovered as significantly differentially expressed in at least one of the datasets along one of the 203 expression patterns at a posterior probability of differential expression (PPDE) > 0.95 (Table S2). This high proportion of DE genes is not entirely surprising given the shear complexity of the dataset being analyzed (Figure 4), since only one condition needs to significantly vary from any of the others to be recovered as such. In either case, both datasets perform similarly, with only 694 and 862 of those DE

transcripts unique to the FPKM and TPM datasets at this significance threshold, respectively. The percentage of transcripts identified as differentially expressed in both datasets remains roughly constant (between 79-81%), until the FDR (false discovery rate) decreases to 0.00 (PPDE = 1.00). At this threshold, FPKM performs more conservatively than TPM, with only 5.3% (273) of the total DE orthologs (5,138) specific to the FPKM dataset, while 23.1% (1,190) are unique to the TPM dataset (Figure S1). This increased conservation can be attributed to the added scaling by transcript length for FPKM expression values.

Table 3. Number of differentially expressed transcripts.

	FDR < 0.05				FDR = 0.00			
	FPKM	**TPM**	**Shared**	**Total**	**FPKM**	**TPM**	**Shared**	**Total**
Tot. DE	694	862	6,681	8,237	273	1,190	3,675	5,138
Tot. *PIT*	406	349	1,690	2,445				
Tot. DE *PIT*	362	368	1,613	2,343	176	403	848	1,427
Up-reg. DE *PIT*	184	152	605	941	83	168	244	495

FPKM and TPM columns correspond to the number of transcripts unique to that dataset. *PIT* = potentially informative transcripts. Up-reg. DE *PIT* = counts for transcripts specific to one of the conditions of interest shown with an * in Table 1.

A total of 2,445 potentially informative transcripts (*PIT*; transcripts whose expression is consistent with one of the 44 potentially informative patterns) were identified in at least one of the two datasets. Of those 1,613 were found significantly differentially expressed in both datasets, while an additional 362 and 368 were specific to the FPKM and TPM sets, respectively (FDR < 0.05; Table 3). While 2,343 *PIT* are differentially expressed between the test conditions in at least one of the datasets, they are not necessarily expressed in a gonophore/medusa specific manner since the predefined expression patterns don't contain information about the magnitude of expression for each condition (Figure 4D, E). Further examination of these DE *PIT* revealed

only 941 are up-regulated in one of the gonophore/medusa containing conditions, of which 605 were significant in both (FDR < 0.05; Table 3; Figure 4C; Tables S2,S5).

Discrepancies between the different normalization methods on which orthologs are differentially expressed can be seen across all levels of the DE analysis and are explained by differences in the expression pattern with the highest posterior probability. This is largely due to disagreement on constraints imposed to identify potentially informative patterns. Disagreement along these constraints is most likely due to the result of scaling by transcript length in FPKM. Although Li and Dewey (2011) suggest that TPM is better expression measure for comparisons between species, there has been no comprehensive evaluation of normalization methods for RNA-Seq in an interspecific DE framework. So to minimize the effect that either normalization method has on the DE results, any ortholog identified in either dataset are considered candidates for future study, while ones shown significant in both datasets should be considered most reliable.

Table 4. Differentially expressed orthologs consistent with previously published studies in *P. carnea*.

RBBH ID	Name	Source	FPKM	TPM	Specificity
RBBH_4080	*Myo1*	Yanze et al. 1999	**	**	All medusa stages
RBBH_5585	*Otx*	Yanze et al. 1999	**NS**	*	Adult medusa
RBBH_8705	*Twist*-like	Spring et al. 2000	**	**	All medusa stages
RBBH_6387	*Snail*	Spring et al. 2002	*	**	All medusa stages
RBBH_9832	*MafL*	Seipel et al. 2004	**	**	All medusa stages
RBBH_540	*Msx*	Galle et al. 2005	*	*	Adult medusa

** in FPKM and TPM columns mark orthologs significant in either dataset at an PPDE = 1.00. * - PPDE > 0.99. NS = not significant.

Previously Published Medusa-specific Candidate Genes

As an initial validation of our DE results, we surveyed previous studies using candidate gene approaches to identify genes specific to medusae development in *P. carnea* (Aerne et al

1995; Spring et al. 2000; Yanze et al. 2001; Galliot and Schmid 2002; Spring et al. 2002; Seipel et al. 2004; Galle et al. 2005). These studies primarily found genes specific to striated muscle in developing and adult medusae. In our DE analyses, orthologs of striated muscle specific homeobox genes *msx* (Galle et al. 2005) and *orthodenticle* (*otx*) (Yanze et al. 2001), a myosin heavy chain, *myo1* (Yanze et al. 2001), a basic leucine zipper transcription factor, *MafL* (Seipel et al. 2004), and zinc finger transcription factors, *twist* and *snail* (Spring et al. 2000, 2002) are differentially up-regulated in developing and/or adult medusa stages of *P. carnea* relative to *H. symbiolongicarpus* sporosacs, which lack all medusae like features, including striated muscle (Table 4) (Boero and Sará 1987; Bouillon et al. 1997; Miglietta and Cunningham 2012).

Spatial Expression of Differentially Expressed Orthologs

To further validate our unbiased interspecific DE analyses, several of the 495 transcripts identified as significantly up-regulated (FDR = 0.00) in at least one dataset (Table 3) were selected for spatial expression analysis via whole mount *in situ* hybridization (ISH) (Table 5; Figure 5; Table S6). None of the candidates discussed below have been previously characterized in cnidarians. Spatial expression of each gene was examined with ISH in each of the tissues sampled for DE analyses. Medusa buds were examined across all ten stages of medusa development as defined by Frey (1968) (Figure 5). Expression of the candidates discussed below was primarily restricted to tissues in the gonophores and adult medusae (Figure 5), except for *APLP,* which also exhibited expression in polyp tissues (not shown).

Two of the candidates surveyed with ISH, *IF2B2* and *TCF3C6*, exhibited similar endodermal expression patterns in *P. carnea* gonophores. *IF2B2* was recovered as significantly up-regulated in all gonophore stages (including *H. symbiolongicarpus* sporosacs and *P. carnea*

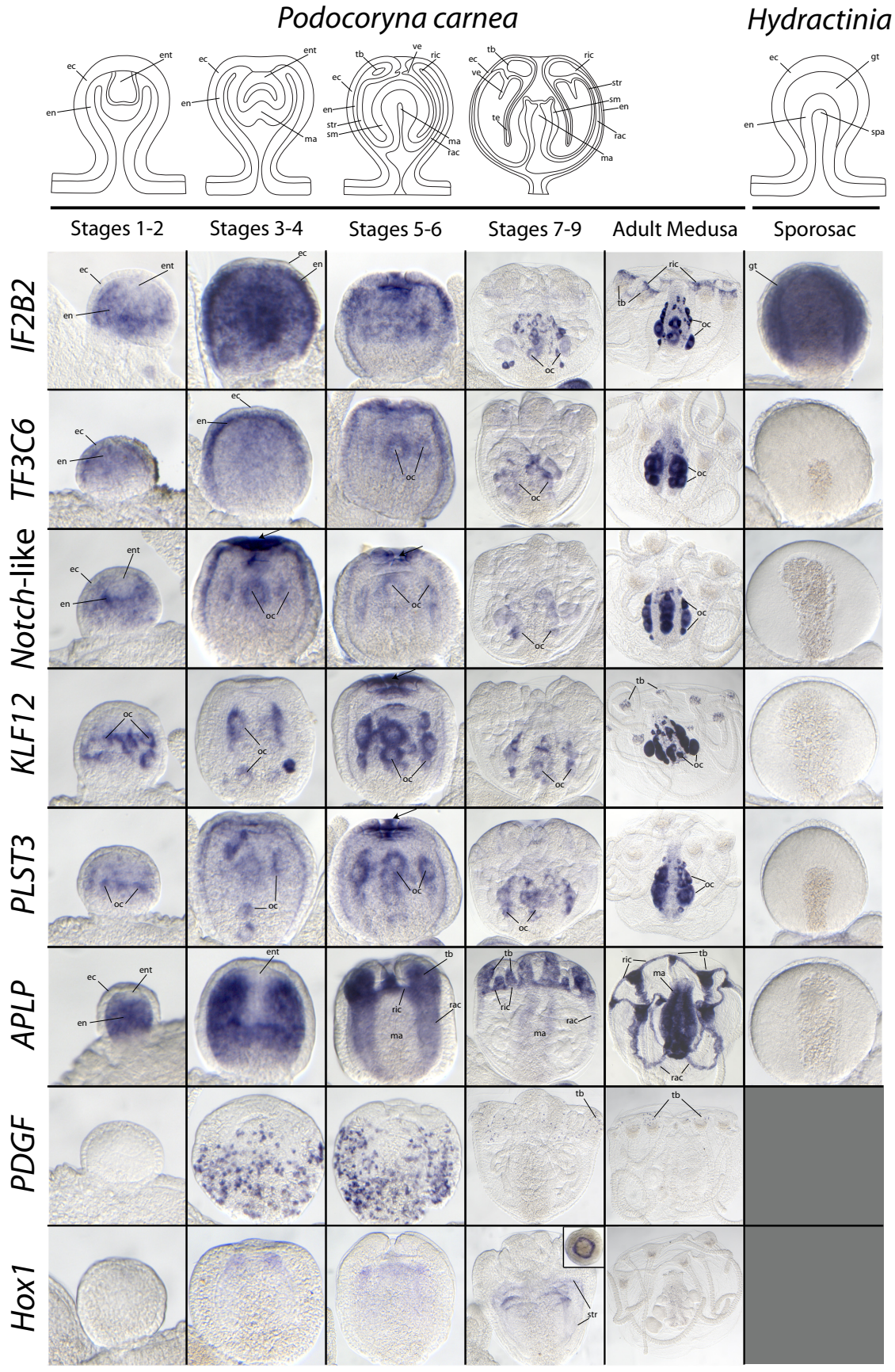

Podocoryna carnea
Hydractinia
Stages 1-2
Stages 3-4
Stages 5-6
Stages 7-9
Adult Medusa
Sporosac
IF2B2
TF3C6
Notch-like
KLF12
PLST3
APLP
PDGF
Hox1

Figure 5. Whole mount *in situ* hybridization results.

Images position the oral end of the gonophores/medusa towards the top. Arrows mark regions of concentrated expression at the distal end of the bell axis or the oral end of the developing manubrium. Only male *H. symbiolongicarpus* sporosacs are shown as eggs in females block the view of the spadix (manubrium anlage). Inset in *Hox1,* stages 7-9 pane is a view from the oral end of the gonophore looking down. Legend: ec – ectoderm; en – endoderm; ent – entocodon; gt – gametic tissue; ma – manubrium, oc – oocytes; rac – radial canal; ric – ring canal; sm – smooth muscle; spa – spadix; str – striated muscle; tb – tentacle bulb; ve - velum.

medusa buds and adult medusa) while *TCF3C6* was recovered as specific to *P. carnea* medusa tissues. ISH of both genes shows strong endodermal expression from stages one through six of medusa development in *P. carnea*, which then ceases by stage seven. After liberation, *IF2B2* is also expressed in part of the gastrovascular system at the medusa bell margin, called the ring canal, primarily in regions where the ring canal meets the proximal portions of the tentacles, called tentacle bulbs. These expression patterns suggest that these genes might have a role inducing cell proliferation as they are expressed in highly proliferative regions of the developing medusa. These spatially restricted expression patterns were not observed in *H. symbiolongicarpus* sporosacs (Fig. 5).

Furthermore, ISH of *IF2B2* revealed expression consistent with a general role in gametogenesis in both *H. symbiolongicarpus* and *P. carnea* (Figure 5). *IF2B2* is expressed around early and late stage oocytes in *P. carnea*. In *H. symbiolongicarpus*, *IF2B2* is expressed around oocytes in the germinal zone (body column) of female gonozooids (not shown), where oogenesis begins (Bunting 1984; Berrill 1953; Müller 1964), and expression continues in the endoderm surrounding oocytes after it moves into the gonophores, where oocyte differentiation continues (Bunting 1984; Berrill 1953; Müller 1964). In males, expression is specific to the endoderm of mature sporosacs (Figure 5), where spermatogenesis occurs (Bunting 1984; Berrill 1953). This expression pattern indicates that *IF2B2* is only operating in gametogenesis and plays no role in patterning the sporosac. This is different from ISH results of *TCF3C6* where no expression of *TCF3C6* was seen either male or female sporosacs, while expression was observed around early and late stage oocytes in *P. carnea*.

ISH of three genes recovered as up-regulated in *P. carnea* medusa tissues, *Notch*-like, *KLF12*, and *PLST3*, revealed similar spatially-restricted expression patterns at the distal tip of

the developing axes of medusa buds of *P. carnea*. For each of these genes, ISH shows minor endodermal expression at various stages of medusa development, but in each case, the prominent expression is seen at the distal end of the developing bell axis by stages five and six. *Notch*-like expression preceeds the expression of both *KLF12* and *PLST3* and is strongly expressed at the distal end of the gonophore prior to opening of the bell margin in stage four, potentially the source of the signal to begin apoptotic patterning of this opening. Past stage six, expression for all genes is specific to maturing oocytes, although after liberation, *KLF12* expression is also expressed in the tentacle bulbs and along the structure called the manubrium (the structure bearing the gonads and mouth at its distal end), in a pattern that appears to be restricted to stem cells. Similar to *TCF3C6* expression, ISH did not detect expression of these genes in the sporosacs of *H. symbiolongicarpus* (Figure 5), except for minor expression around early stage oocytes (not shown).

Although showing high similarity and a top blast hit to other 'neurogenic notch' genes, the ortholog, *Notch*-like, examined in this study is not orthologous to those notch genes examined in *Hydra* and *Nematostella* (Käsbauer et al. 2007; Marlow et al. 2012). Structural analysis of *Notch*-like reveals that it lacks the essential domains that characterizes these genes (LNR and NOD), and only share the common EGF domains present in notch genes (which are not specific to just notch genes) (Käsbauer et al. 2007; Marlow et al. 2012).

APLP is the only candidate selected for ISH that does not exhibit expression consistent with any role in oogenesis in *P. carnea*. All throughout gonophore development, ISH reveals *APLP* expression to be specific to the endodermal tissues that give rise to gastric structures in the adult medusa. Starting at stage one, strong *APLP* expression is detected in the endoderm of the newly formed gonophore. As development proceeds, *APLP* expression remains specific to the

endodermal tissue beginning to form the radial canals in stages three through six. *APLP* expression is excluded from and clearly outlines the entocodon, which is medusa-specific tissue layer formed through evagination of the distal ectoderm of the gonophore that gives rise to striated muscle (Avset 1961). By stages five and six, strong expression is noted in the newly formed ring canal and tentacle bulbs, but is excluded from the developing manubrium. This pattern continues through the later stage buds, but seems to decrease in the strength of expression (especially in the radial canals) until the medusa is liberated from the colony, where expression strongly reappears in all digestive tissues; including the fully developed manubrium, radial and ring canals, and tentacle bulbs (Figure 5). *APLP* expression was not observed in the sporosacs of *H. symbiolongicarpus*.

Throughout medusa development, *APLP* appears to be expressed in a manner consistent with the patterning and development of the digestive tract of the medusa, while after liberation it would appear to play some role in digestion and nutrient movement. This is consistent with *APLP*'s role in other animals, where it functions not only as a lipid trafficking molecule but also plays a critical role in regulating *hedgehog* and Wnt signaling during wing development in *Drosophila* (Panáková et al 2005).

Previous studies have implicated canonical Wnt signaling in medusae evolution (Duffy et al. 2010; Duffy 2011; Nawrocki and Cartwright 2013), yet, given the dual role of the Wnt pathway as both a maternal effect for larval development and in adult patterning (Plicket et al. 2006; Teo et al. 2006; Momose and Houliston 2007; Müller 2007; Momose et al. 2008; Amiel and Houliston 2009; Duffy et al. 2010), DE expression of Wnt signaling genes in medusa would likely be obscured by high expression in female gametic tissue due to maternal loading, violating the initial constraint of identifying potentially informative patterns (equivalent expression of

male and female gonozooids). Yet even without many of the key wnt signaling components recovered by our DE analyses, recovery of *APLP* further implicates the role of Wnt signaling in medusa development and evolution (Nawrocki and Cartwright 2013). Further ties between *APLP* and Wnt signaling come from the expression patterns of *APLP* in both *P. carnea* and *H. symbiolongicarpus* polyps (not shown). In the feeding polyps of both species, *APLP* is expressed in a ring around the distal tip of the hypostome and in the endoderm at the tip of the tentacles (Figure S2), consistent with observed *Wnt3* expression in *Hydra* (Guder et al. 2005; Lengfeld et al. 2009; Gee et al. 2010), *Hydractinia* (Plickert et al. 2006; Müller et al. 2007; Duffy et al. 2010), *Podocoryna* (Sanders and Cartwright, in review), and *Ectopleura* (Nawrocki and Cartwright 2013). Similarly, ISH revealed *APLP* expression at the distal tip of *H. symbiolongicarpus* gonozooids (Figure S2), consistent with *Wnt3* expression in *H. echinata* (Müller et al. 2007; Duffy et al. 2010) and *H. symbiolongicarpus* (name Steven M Sanders, unpublished data).

Table 5. Differentially expressed orthologs and *P. carnea* specific genes validated with *in situ* hybridization.

RBBH ID	Name	Blast hit	FPKM	TPM
RBBH_6358	*IF2B2*	Insulin-like growth factor 2 mRNA-binding protein 2	**	**NS**
RBBH_7273	*TCF3C6*	General transcription factor 3C polypeptide 6	**	**NS**
RBBH_608	*Notch*-like	Neurogenic locus notch homolog protein 1	**	**
RBBH_8405	*KLF12*	Krueppel-like factor 12	**	**
RBBH_2122	*PLST*	Plastin-3	**	**
RBBH_3474	*APLP*	Apolipophorins	**	**
NONE	*PDGF*	Platelet-derived growth factor subunit A	**NA**	**NA**
NONE	*Hox1*	Homeobox protein Hox-B1	**NA**	**NA**

** in FPKM and TPM columns mark orthologs significant in either dataset at an PPDE = 1.00. NS = not significant. NA = not subject to DE analyses. NONE = ortholog not present in *H. symbiolongicarpus*.

Spatial Expression of *Podocoryna carnea* Specific Genes

While differential regulation of orthologous genes does explain differences in homologous structures between species, evolutionary shifts between phenotypes are likely accompanied by gene gain or loss as well. To explore this, a growth factor, *PDGF*, and a homeobox gene, *Hox1,* were selected from those transcripts in *P. carnea*'s filtered transcriptome that did not have a predicted *H. symbiolongicarpus* ortholog (Table 5). In each case, these transcripts could represent genes that were either independently gained in *P. carnea* or lost in *H. symbiolongicarpus*. When comparing *de novo* transcriptomes, potential gene duplications and deletions must be confirmed with phylogenetic and genomic analyses, as the lack of a given gene could be the result of the lack of its expression in the tissues or developmental stages sampled and/or extreme sequence divergence of a given ortholog. Phylogenetic analyses of the cnidarian homeobox and vascular-endothelial growth factor gene families (Figures S3,S4), as well as searching unpublished genomic scaffolds of *H. symbiolongicarpus*, confirm the lack of *H. symbiolongicaparpus* orthologs to *Hox1* and *PDGF*.

ISH of *PDGF* confirmed the specificity of this gene to developing and adult medusa stages with no expression detected in the polyp stage of *P. carnea*. *PDGF* expression begins at bud stages three and four of medusa development, revealing a speckled expression pattern. This expression pattern is consistent with being specific to a type stem cells restricted to the developing and adult medusa. Since this gene is not expressed in the known stem cell populations of hydrozoan polyps (Teo et al. 2006; Müller et al. 2007; Millane et al. 2011; Duffy et al. 2012; Hemmrich et al. 2012), it suggests a potential medusa-specific stem cell lineage. At stages three and four, these *PDGF* positive cells are concentrated at the proximal end of the medusa bud. This pattern lessens as medusa development continues, and *PDGF* positive cells

appear to be evenly distributed except at the most distal tip of the gonophore by stages five and six. By stages seven and eight of medusa development, expression is limited to just a few stem cells in the tentacle bulbs and this continues through the adult medusa.

ISH showed relatively little expression of *Hox1* through medusa development with no expression detected in the polyp. Noticeable (although very faint) expression begins by stage three and continues through the later stages of gonophore development, until the medusa is fully developed, where expression ceases. At the earlier stages, *Hox1* expression is seen as a ring like pattern around the distal region of the differentiating entocodon (Figure 5). This pattern is maintained as gonophore development progresses, broadening the expression ring as the gonophore grows. By stages seven and eight the strongest expression seen at the distal end of the expression domain, yet with more minor expression dispersed more proximally along the striated muscle tissue of the developing medusa.

ISH of *Hox1* is consistent with expression reported by Aerne et al. (1995), where expression was detected in bud stages with developing striated muscle. Later, Yanze et al. (2001) showed *Hox1* expression throughout embryonic development and in the aboral end of the planula, consistent with *Hox1* orthologs (74% BS support; Figure S3) reported in *Hydra* (Schummer et al. 1992; Shenk et al. 1993a,b; Gauchat et al. 2000; Smith et al. 2000; Bode et al. 2001), *Clytia* (Chiori et al. 2009), and *Eleutheria* (Jakob and Schierwater 2007), where it appears to play a role in the oral-aboral patterning. Two of three taxa (*Clytia* and *Eleutheria*) with an ortholog of *Hox1* possess fully developed medusa and in each case *Hox1* expression is not observed in the striated muscle. In *C. hemisphaerica* medusa, *Hox1* expression is specific to the balancing organ (statocyst; not present in *P. carnea* medusa) (Chiori et al. 2009), while *E. dichotoma* benthic medusae exhibit no *Hox1* expression (Jakob and Schierwater 2007).

While results from the phylogenetic analyses suggest that *Hox1* evolved earlier in Hydrozoa and was subsequently lost in *H. symbiolongicarpus*, *PDGF* appears to have be the result of a duplication event in *P. carnea*, although this could be an artifact of limited sampling of this gene family (Figures S3,S4). The lack of an orthologous *Hox1* gene in *H. symbiolongicarpus* is consistent with the loss of striated muscle during truncation of the medusa following a *Hox1* deletion, this however does not explain the observed variability of *Hox1* expression in the medusa of more distantly related hydrozoans. This variable expression of *Hox1* across distantly related taxa suggests a potential evolutionary scenario where *Hox1* was co-opted to be involved in striated muscle development during a transition towards fully developed medusa in *P. carnea* as no other hydrozoan medusa exhibits *Hox1* expression in their striated muscle tissues (Jakob and Schierwater 2007; Chiori et al. 2009). Future areas of research should focus on sampling more intermediate levels of medusa truncation in order to determine if changes in expression correlates with the development of striated muscle tissue.

These results show interspecific differential expression analyses to be a more sensitive method for identifying candidate genes and/or gene networks involved in the evolutionary transitions between different life history forms than more common comparative methods such as enrichment analyses. Albeit more powerful, our method is reliant on identifiable orthologs. Further analyses of the genes *PDGF* and *Hox1*, which were absent in *H. symbiolongicarpus*, revealed expression consistent with an important role in medusa development in *P. carnea*. Thus, both up- and down-regulation of orthologous genes and novel gene gain and loss appear important for the reduction and possible re-evolution of the medusa life cycle stage in the Hydractiniidae. With nearly 100 million years of divergence between these two species (Miglietta and Cunningham 2012), which exhibit the ‘book-end’ phenotypes, the differential

regulation of orthologs and gene duplication and loss, most likely accompanied the transitions between the fully developed and fully truncated medusa.

Addressing questions of parallel incidences of medusa loss, and even re-gain, requires increased taxonomic sampling. Increasing the number of taxa sampled adds a new layer of complexity to ensure the validity of the DE analysis. Dunn et al. (2013b) extensively reviewed not only the utility of comparative expression across multiple species but also the numerous challenges it presents. As with any phylogenetic statistical analysis, the non-independent nature of the data can have large effects on the results (Felsenstein 1985). Future studies sampling more than two species will need to expand current DE software to utilize independent contrasts.

Here we have provided a new workflow with which one can effectively quantify cross-species differences in expression using short read Illumina data. DE results between these two hydractiniid species reveal 941 candidate orthologs potentially involved in the evolution of the medusa life cycle stage. Moreover, orthology assignments and phylogenetic analyses suggest multiple instances of novel gene loss and gain correlated with phenotypic differences of the gonophore in *P. carnea* and *H. symbiolongicarpus*. Expanding this method to include more taxa and utilizing independent contrasts should provide significant insight into the role of these genes in medusa evolution in Hydrozoa.

Materials and Methods

Animal Care

Transplanted colonies of *P. carnea* and *H. symbiolongicarpus* were kept on microscope slides, placed in slide racks and kept in seawater (REEF CRYSTALS, Aquarium Systems) aquaria at room temperature (~21°C) with a salinity of 29 and 32 ppt, respectively. Colonies were fed 2-3-day-old nauplii of *Artemia* three times a week.

Tissue Collection and RNA Isolation

Podocoryna carnea colonies were first starved a minimum of three days (up to five days) prior to tissue collection. For non-reproductive and reproductive polyp samples, roughly 100 polyps were individually dissected and immediately flash-frozen for each RNA extraction. Reproductive polyps sampled spanned all developmental stages of gonophore development, as staged according to Frey (1968). To collect medusae tissue, highly reproductive colonies were kept in a small container overnight, and liberated medusae were collected and flash-frozen the following morning. All tissue samples were stored at -80°C until RNA extractions could be performed. Total RNA was isolated at the University of Kansas Medical Genome Sequencing Facility (KUMC-GSF) using the TriReagent isolation protocol (Invitrogen).

Prior to tissue collection for whole mount *in situ* hybridization, colonies were starved for three days. After the third day, colonies were anesthetized with menthol crystals and fixed in 4% paraformaldehyde overnight at 4°C. Fixed tissue was rinsed with and stored in 100% methanol at -20°C.

Library Construction and Sequencing

Podocoryna carnea RNA libraries were constructed at KUMC-GSF according to the TruSeq RNA Sample Preparation Guide (Illumina) using the TruSeq RNA Sample Preparation Kit (Box A). All libraries were 100bp paired-end with an average insert size of 160bp. Libraries were then barcoded, pooled, and multiplexed on a single lane of either an Illumina HiSeq2500 or HiSeq2000 flowcell at KUMC-GSF. Sample Med3 was multiplexed and sequenced on a lane of an Illumina HiSeq2000 flowcell, while all other samples were multiplexed and sequenced on the same lane of an Illumina HiSeq 2500. The raw reads have been submitted to the NCBI Sequence Read Archive (SRX529566).

Transcriptome Assembly and Annotation

Figure 2 is a schematic of our bionformatic pipeline for identifying differentially expressed orthologs. Raw Illumina RNAseq data for *H. symbiolongicarpus* was downloaded from the SRA archive (SRX474462). *Hydractinia symbiolongicarpus* libraries included: four replicated libraries of feeding (non-reproductive) polyps (gastrozooids), two replicated libraries of female reproductive polyps (gonozooids), two replicated libraries of male reproductive polyps (gonozooids), and four replicated libraries of defensive (non-reproductive) polyps (dactylozooids). *Podocoryna carnea* libraries included: three replicated libraries of non-reproductive gastrozooids, four replicated libraries of female reproductive gastrozooids, and three replicated libraries of female adult medusa. These reads were pooled by species and then assembled with Trinity (Grabherr et al. 2011) through the automated bioinformatics pipeline, Agalma (Dunn et al. 2013a), under default settings. When possible, transcripts were named according to the transcript annotations provided by Agalma.

In order to perform gene ontology analyses, transcripts were blasted against the nr protein database using blastx with the '–outfmt 5' flag for xml formatted output. Blast output was imported into Blast2Go (Conesa et al. 2005; Götz et al. 2008) where GO mapping and annotations were performed. Conserved protein domains were also identified using with the PFAM (Punta et al. 2012) and TIGR (http://blast.jcvi.org/web-hmm/) databases using HMMER (http://hmmer.org/). The enriched GO terms and protein domains was assessed with the Fisher's Exact Test and corrected for a false discovery rate of 0.05 in R.

Ortholog Identification and Differential Expression Analysis

To identify putative orthologs, assemblies were filter based on a minimum relative expression (FPKM) value and default Transdecoder (http://transdecoder.sourceforge.net/) reading frame criteria (Figure 2). An FPKM value was calculated for each transcript across all libraries used in the assemblies and were used as a means of assessing the relative coverage of each transcript. Transcripts that met both filtering criteria (filtered transcriptome) were then translated by their longest reading frame, and blasted against the other filtered transcriptome using the Blastp algorithm. One-to-one reciprocal best blast hits (RBBHs) with both e-values > 1e-03 were treated as orthologous genes. Since dealing with only two taxa, reciprocal blast best hits is an adequate means of establishing orthology and is commonly used (Yang and Wang 2013; Pankey et al. 2014). As the number of taxa considered increases, tree based methods would become necessary to identify true orthologs.

Expression of orthologs was calculated with RSEM (Li and Dewey 2011) by remapping the raw reads from the individual libraries to the filtered transcriptome of the corresponding species, excluding only libraries specific to the *H. symbiolongicarpus* dactylozooids. RSEM

calculates expression levels and computes three different expression values: expected counts, TPM (transcripts per million), and FPKM (fragments per kilobase of transcript per million mapped reads). Separate gene-level DE analyses were performed with EBSeq (Leng et al. 2013) using the TPM and FPKM dataset. Results were filtered based on the inferred posterior probability that a gene was differentially expressed (PPDE; equal to one minus the false discovery rate: 1-FDR) for a particular expression pattern.

As the number of conditions increase, so do the number of possible expression patterns. With six conditions, there are 203 possible patterns (Table S3). To limit the results to those informative to our question, we identified 44 potentially informative expression patterns that are gonophore and medusa-specific. Transcripts marked by the remaining 159 expression patterns were ignored. Transcripts identified with a PPDE $\geq$ 0.95 along one of these 44 patterns (Table S4), with the highest expression observed in one of the gonophore/medusa containing conditions, were selected as candidates for further study in medusa evolution (Table S5).

Probe Synthesis and *in situ* Hybridization

Sequences of transcripts listed in Table 4 were identified from each assembly. The reading frames of each species copy were aligned and primers (Table S6) were selected to encompass homologous regions of each transcript. These fragments were then amplified from cDNA, cloned using the Invitrogen TOPO-TA Cloning Kit, and DIG labeled riboprobes were synthesized from clones using the Invitrogen T7/T3 Megascript kit. ISH protocol was adapted from Gajewsky et al. (1996). Hybridization was carried out at 50˚C for 16-18 hours with a probe concentration of .1 ng/µl. Hybridization was detected by immunostaining with anti-DIG-Fab-AP (ROCHE) and NBT/BCIP.

Molecular Phylogenetic Analyses

Cnidarian sequences belonging to homeobox and *PDGF/VEGF* gene families were mined from the nr NCBI database and subject to phylogenetic analysis as a means to quickly establish orthology with those genes in our dataset. *Podocoryna carnea* and *H. symbiolongicarpus* amino acid sequences belonging to the gene families of interest were extracted from the assemblies and subject to a clustering analysis using CD-HIT (Li and Godzik 2006; Fu et al. 2012) (under a 90% sequence similarity threshold) to remove redundant gene copies. Alignments were conducted with Mafft (Katoh et al. 2005) under the L-insi alignment algorithm. Maximum likelihood estimates of the gene trees were then inferred using RAxML (Stamatakis et al. 2008) on the CIPRES portal (Miller et al. 2010) using the rapid bootstrapping (-f a) algorithm with 100 bootstrap replicates under the PROTGAMMA+WAG model (Figures S3,S4).

Acknowledgments

We would like to thank Neil Blackstone for lab cultures of *P. carnea*, Clark Bloomer at KUMC-GSF for RNA isolation, library construction, and sequencing, Kirsten Jensen for the illustrations in Figure 1, Susan Boyles for illustrations in Figure 5, Casey Dunn and Felipe Zapata for help implementing the Agalma package, and Justin Blumenstiel and Uri Frank for discussions, and members of the Hileman Lab (University of Kansas) for comments on a previous version.

References

Aerne BL, Baader CD, Schmid V. 1995. Life stage and tissue-specific expression of the homeobox gene cnox1-Pc of the hydrozoan *Podocoryne carnea*. Dev Biol. 169:547-556.

Allman GJ. 1864. On the construction and limitation of genera among the Hydroida. Ann Mag Nat Hist Ser. 13:345–80.

Amiel A, Houliston E. 2009. Three distinct RNA localization mechanisms contribute to oocyte polarity establishment in the cnidarian *Clytia hemisphaerica*. Dev Biol. 327:191-203.

Avset K. 1961. The development of the medusa *Podocoryne carnea*. Nytt Mag Zool. 10:49-56.

Berrill NJ. 1953. Growth and form in gymnoblastic hydroids; polymorphism within the Hydractiniidae. J Morph. 92:241-272.

Bode HR. 2001. The role of Hox genes in axial patterning in *Hydra*. Amer Zool. 41:621–628.

Boero F, Sará M. 1987. Motile sexual stages and evolution of Leptomedusae (Cnidaria). Boll Zool. 54:131-139.

Bouillon J, Medel D, Cantero ALP. 1997. The taxonomic status of the genus *Stylactaria* Stechow, 1921 (Hydroidomedusae, Anthomedusae, Hydractiniidae), with the description of a new species. Sci Mar. 61:471-486.

Boyle AP, Araya CL, Brdlik C, Cayting P, Cheng C, Cheng Y, Gardner K, Hillier LW, Janette J, Jiang L, et al. 2014. Comparative analysis of regulatory information and circuits across distant species. Nature. 512:453-456.

Bunting M. 1894. The origin of the sex-cells in *Hydractinia* and *Podocoryne*; and the development of *Hydractinia*. J Morph. 9:203-236.

Cartwright P, Nawrocki AM. 2010. Character evolution in Hydrozoa (phylum Cnidaria). J Integ Comp Biol. 50:456-472.

Chiori R, Jager M, Denker E, Wincker P, Silva CD, Guyader HL, Manuel M, Quéinnec E. 2009. Are Hox genes ancestrally involved in axial patterning? Evidence from the hydrozoan *Clytia hemisphaerica* (Cnidaria). PLoS One. 4:e4231.

Conesa A, Götz S, Garcia-Gomez JM, Terol J, Talon M, Robels M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 21:3674-3676.

Cornelius PFS. 1992. Medusa loss in leptolid Hydrozoa (Cnidaria), hydroid rafting, and abbreviated life cycles among their remote island faunae: An interim review Sci Mar. 56:245-261.

Cunningham CW, Buss LW. 1993. Molecular evidence for multiple episodes of paedomorphosis in the family Hydractiniidae. Biochem Sys Ecol. 21:57-69.

Duffy DJ. 2011. Modulation of Wnt signaling: A route to speciation? Comm & Integ Biol. 4:59-61.

Duffy DJ, Millane RC, Frank U. 2012. A heat shock protein and Wnt signaling crosstalk during axial patterning and stem cell proliferation. Dev Biol. 362:271-281.

Duffy DJ, Plickert G, Kuenzel T, Tilmann W, Frank U. 2010. Wnt signaling promotes oral but suppresses aboral structures in *Hydractinia* metamorphosis and regeneration. Dev. 137:3057-3066.

Dunn CW, Howison M, Zapata F. 2013a. Agalma: an automated phylogenomics workflow. BMC Bioinformatics. 14:330.

Dunn CW, Luo X, Wi Z. 2013b. Phylogenetic analysis of gene expression. Int Comp Biol. 53:847-856.

Felsenstein J. 1985. Phylogenies and the comparative method. Amer Nat. 125:1-15.

Frey J. 1968. Die Entwicklungsleistungen der Medusenknospen und Medusen von *Podocoryne carnea* M. Sars nach Isolation und Dissoziation. Wilhelm Roux' Arch Entwicklungsmech Org. 168:428-464.

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CH-HIT: accelerated for clustering the next generation sequencing data. Bioinformatics. 28:3150-3152.

Gajewsky M, Leitz T, Schlossherr J, Plickert G. 1996. LWamides from Cnidaria constitute a novel family of neuropeptide with morphogenetic activity. Roux's Arch Dev Biol. 205:232– 242.

Galle S, Yanze N, Seipel K. 2005. The homeobox gene Msx in development and transdifferentiation of jellyfish striated muscle. Int J Dev Biol. 49:961-967.

Galliot B, Schmid V. 2002. Cnidarians as a model system for understanding evolution and regeneration. Int J Dev Biol. 46:39-48.

Gauchat D, Mazet F, Berney C, Schummer M, Kreger S, Pawlockski J, Galliot B. 2000. Evolution of Antp-class genes and differential expression of *Hydra* Hox/paraHox genes in anterior patterning. PNAS. 97:4493–4498.

Gee L, Hartig J, Law L, Wittlieb J, Khalturin K, Bosch TCG, Bode HR. 2010. ß-catenin plays a central role in setting up the head organizer in hydra. Dev Biol. 340:116-124.

Gibbons MJ, Janson LA, Ismail A, Samaai T. 2010. Life cycle strategy, species richness and distribution in marine Hydrozoa (Cnidaria: Medusozoa). J Biogeogr. 37:441-448.

Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucl Acids Res. 36:3420-3435.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al. 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat Biotechnol. 29:644-52.

Guder, C, Pinho S, Nacak TG, Schmidt HA, Hobmayer B, et al. 2005. An ancient Wnt-Dickkopf anatagonism in *Hydra*. Dev. 133:901-911.

Hao DC, Ge G, Xiao P, Zhang Y, Yang L,. 2011. The first insight into the tissue specific *Taxus* transcriptome via Illumina second generation sequencing. PLoS One. 6:e21220.

Helm RR, Siebert S, Tulin S, Smith J, Dunn CW. 2013. Characterization of differential transcript abundance through time during *Nematostella vectensis* development. BMC Genomics. 14:266.

Hemmrich G, Khalturin K, Boehm A, Puchert M, Anton-Erxleben F, et al. 2012. Molecular signatures of the three stem cell lineages in *Hydra* and the emergence of stem cell function at the base of multicellularity. Mol Biol Evol. 29:3267-3280.

Jakob W, Schierwater B. 2007. Changing hydrozoan bauplans by silencing Hox-like genes. PLoS One. 8:e694.

Käsbauer T, Towb P, Alexandrova O, David CN, Dall'Armi E, et al. 2007. The Notch signaling pathway in the cnidarian *Hydra*. Dev Biol. 303:376-390.

Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in the accuracy of multiple sequence alignment. Nucl Acids Res. 33:511-518.

Leclère L, Schuchert P, Cruaud C, Couloux A, Manuel M. 2009. Molecular

phylogenetics of Thecata (Hydrozoa, Cnidaria) reveals long-term maintenance of life history traits despite high frequency of recent character changes. Sys Biol. 58:509-526.

Leclère L, Schuchert P, Manuel M. 2007. Phylogeny of the Plumularioidea (Hydrozoa, Leptothecata): evolution of colonial organisation and life cycle. Zool Scripta. 36:371-394.

Leng N, Dawson JA, Thomson HA, Ruotti V, Rissman AI, Smits BMG, Haag JD, Gould MN, Stewart RM, Kendziorski C. 2013. EBSeq: an emipirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics. 29:1035-1043.

Lengfeld T, Watanabe H, Simakov O, Lindgens D, Gee L, Law L, Schmidt HA, Özbek S, Bode H, Holstein TW. 2009. Multiple Wnts are involved in *Hydra* organizer formation and regeneration. Dev Biol. 330:186-199.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. BMC Bioinformatics. 12:323.

Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics. 22:1658-1659.

Marlow H, Roettinger R, Boekhout M, Martindale MQ. 2012. Functional roles of Notch signaling in the cnidarian *Nematostella vectensis*. Dev Biol. 362:295-308.

Marques AC, Migotto AE. 2001. Cladistic analysis and new classification of the family Tubulariidae (Hydrozoa, Anthomedusae). Pap Avuls Zool. 41:465–88.

Miglietta MP, Cunningham CW,. 2012. Evolution of life cycle, colony morphology, and host-specificity in the Family Hydractiniidae (Hydrozoa, Cnidaria). Evolution. doi:10.1111/j.1558-5646.2012.01717.x.

Miglietta MP, McNally L, Cunningham CW. 2010. Evolution of calcium-carbonate skeletons in the Hydractiniidae. J Integ Comp Biol. Symposia:1-8.

Miglietta MP, Schuchert P, Cunningham CW. 2009. Reconciling genealogical and morphological species in a worldwide study of the Family Hydractiniidae (Cnidaria, Hydrozoa). Zool Scripta. 38:403-430.

Millane RC, Kansa J, Duffy DJ, Seoighe C, Cunningham S, Plickert G, Frank U. 2011. Induced stem cell neoplasia in a cnidarian by ectopic expression of a POU domain transcription factor. Dev. 138:2429-2439.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. GCE. 1-8.

Momose T, Derelle R, Houliston E. 2008. A maternally localized Wnt ligand required for axial patterning in the cnidarian *Clytia hemisphaerica*. Dev. 135:2105-2113.

Momose T, Houliston E. 2007. Two oppositely localized frizzled RNAs as axis determinants in cnidarian embryo. PLoS Biol. 5:e70.

Müller W. 1964. Experimentelle Untersuchungen Über Stockentwicklung, Polypendifferenzierung, und sexualchimären bei *Hydractinia echinata*. Wilhelm Roux' Arch Entwicklungsmech Org. 155:181-268.

Müller W, Frank U, Teo R, Mokady O, Guette C, Plickert G. 2007. Wnt signaling in hydroid development: ectopic heads and giant buds induced by GSK-3ß inhibitors. Int J Dev Biol. 51:211-220.

Nawrocki AM, Cartwright P. 2013. Expression of Wnt pathway genes in polyps and medusa-like structures of *Ectopleura larynx* (Cnidaria: Hydrozoa). Evol & Dev. 15:373-384.

Panáková D, Sprong H, Marois E, Thiele C, Eaton S. 2005. Lipoprotein particles are required for Hedgehog and Wingless signaling. Nature. 435:58-65.

Pankey MS, Mini VN, Imholte GC, Suchard MA, Oakley TH. 2014. Predictable transcriptome evolution in the convergent and complex bioluminescent organs of squid. PNAS. doi:10.1073/pnas.141657411.

Plickert G, Jacoby V, Frank U, Müller WA, Mokady O. 2006. Wnt signaling in hydroid development: Formation of the primary body axis in embryogenesis and its subsequent patterning. Dev Biol. 298:368-378.

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, et al. 2012. The Pfam protein families database. Nucl Acid Res. Database Issue:doi:10.1093/nar/gkr1065.

Sanders SM, Shcheglovitova M, Cartwright P. 2014. Differential gene expression between functionally specialized polyps of the colonial hydrozoan *Hydractinia symbiolongicarpus* (Phylum Cnidaria). BMC Genomics. 15:406.

Schummer M, Scheurlen I, Schaller C, Galliot B. 1992. HOM/Hox homeobox genes are present in hydra (*Chlorohydra viridissima*) and are differentially expressed during regeneration. EMBO J. 11:1815–1823.

Schunter C, Vollmer SV, Macpherson E, Pascual M. 2014. Transcriptome analyses and differential gene expression in a non-model fish species with alternative mating tactics. BMC Genomics. 15:167.

Seipel K, Yanze N, Müller P, Streltwolf R, Schmid V. 2004. Basic leucine sipper transcripton factors C/EBP and Mafl in the hydrozoan jellyfish *Podocoryne carnea*. Dev Dynamics. 230:392-402.

Shenk MA, Bode HR, Steele RE. 1993a. Expression of Cnox-2, a HOM/Hox homeobox gene in *Hydra*, is correlated with axial pattern formation. Dev. 117:657–667.

Shenk MA, Gee L, Steele RE, Bode HR. 1993b. Expression of Cnox-2, a HOM/Hox gene, is suppressed during head formation in *Hydra*. Dev Biol. 160:108–118.

Siebert S, Robinson MD, Tintori SC, Goetz F, Helm RR, Smith SA, Shaner N, Haddock SHD, Dunn CW. 2011. Differential gene expression in the siphonophore *Nanomia bijuga* (Cnidaria) assessed with multiple next-generation sequencing workflows. PLoS One. 6:e22953.

Smith KM, Gee L, Bode HR. 2000. HyAlx, an aristaless-related gene, is involved in tentacle formation in hydra. Dev. 127:4743–4752.

Spring J, Yanze N, Jösch C, Middel AM, Winninger B, Schmid V. 2002. Conservation of Brachyury, Mef2, and Snail in the myogenic lineage of jellyfish: A connection to the mesoderm of Bilateria. Dev Biol. 244:372-384.

Spring J, Yanze N, Middel AM, Stierwald M, Gröger H, Schmid V. 2000. The mesoderm specification factor Twist in the life cycle of jellyfish. Dev Biol. 228:363-375.

Stamatakis A, Hoover P, Rougemont J. 2008. A fast bootstrapping alogorithm for the RAxML web-servers. Sys Biol. 57:758-771.

Teo R, Möhrlen F, Plickert G, Müller WA, Frank U. 2006. An evolutionary conserved role of Wnt-signaling in stem cell fate decision. Dev Biol. 289:91-99.

Yang R, Wang X. 2013. Organ evolution in angiosperms driven by correlated divergences of gene sequences and expressed patterns. Plant Cell. 25:71-82.

Yanze N, Spring J, Schmidli C, Schmid V. 2001. Conservation of Hox/ParaHox-related genes in the early development of a cnidarian. Dev Biol. 236:89-98.

Supplementary Materials

Figure S1. Number DE transcripts vs FDR. Traces the total number of differentially expressed as the FDR changes from 0.10 to 0.00. FPKM.tot = total number of DE transcripts from the FPKM dataset; TPM.tot = total number of DE transcripts from the TPM dataset; Shared = total number of DE transcripts identified in both datasets; FPKM.uniq = number of DE transcripts unique to the FPKM dataset (equal to the difference between FPKM.tot and Shared); TPM.uniq = number of DE transcripts unique to the TPM dataset.

Figure S2. ISH of *APLP* in polyps. A) Expression in *H. symbiolongicarpus* gastrozooid. B) Expression in *H. symbiolongicarpus* female gonozooid. C) Expression in a *P. carnea* gastrozooid. D) Expression in *P. carnea* gastrozooid with focus on the hypostome.

Figure S3. Homeobox gene tree. Phylogeny of cnidarian *Hox* genes. Highlighted in red are sequences belonging to *H. symbiolongicarpus*' and *P. carnea*'s filtered assemblies (corresponding to prefixes Hs and Pc, respectively). '*' = sequences analyzed with ISH. Maximum likelihood bootstrap support values given at the base of each node.

Figure S4. *PDGF/VEGF* gene tree. Phylogeny of cnidarian *PDGF/VEGF* genes. Highlighted in red are sequences belonging to *H. symbiolongicarpus*' and *P. carnea*'s filtered assemblies (corresponding to prefixes Hs and Pc, respectively). '*' = sequences analyzed with ISH. Maximum likelihood bootstrap support values given at the base of each node.

Table S1. Reciprocal blast best hits text file. An excel file giving the transcript ID from each species filtered transcriptome that were reciprocal blast best hits and the name of the putative ortholog (RBBH_#) they were assigned.

Table S2. Reciprocal blast best hits summary table. An excel file summarizing the annotations and DE results for each ortholog. For each ortholog, the species top blast hit, gene ontology terms, and HMM families are given. Also, the patterns and PPDE given from each DE analysis are given.

Table S3. List of all possible expression patterns. An excel file summarizing all of the possible expression patterns when assessing DE between six conditions. Numbers in each cell do not indicate magnitude of expression, but statistically equivalent or different levels of expression. 'Gast' = *H. symbiolongicarpus* gastrozooid. 'F_Gono' = *H. symbiolongicarpus* female sporosac. 'M_Gono' = *H. symbiolongicarpus* male sporosac. 'NRP' = *P. carnea* polyp without medusa buds. 'RP' = *P. carnea* polyp with medusa buds. 'Med' = *P. carnea* adult, liberated medusa.

Table S4. List of potentially informative patterns. The filtered EBSeq pattern list following our biological constraints (described above) to identify DE patterns informative to our question. 'Gast' = *H. symbiolongicarpus* gastrozooid. 'F_Gono' = *H. symbiolongicarpus* female sporosac. 'M_Gono' = *H. symbiolongicarpus* male sporosac. 'NRP' = *P. carnea* polyp without medusa buds. 'RP' = *P. carnea* polyp with medusa buds. 'Med' = *P. carnea* adult, liberated medusa.

Table S5. Summary of DE orthologs specific to gonophore containing conditions. Excel file containing summary information for each of the 945 orthologs DE between *H. symbiolongicarpus* sporosacs and *P. carnea* developing and adult medusae along one of the 44 potentially informative patterns in either dataset. Summary information contains the top blast hit for each species ortholog sequence, as well as the pattern, PPDE, and tissue specificity assigned in both the FPKM and TPM datasets. In the '*.Specificity' columns: 'Go' = *H. symbiolongicarpus* sporosacs, 'RP' = *P. carnea* medusa buds, 'Med' = *P.carnea* adult, liberated medusa. Given the pattern assigned, the ortholog can be specific to a combination of these tissues (i.e. 'Med_RP' = up-regulated expression in both adult and developing medusa relative to all other tissues sampled).

Table S6. Primers for each candidate gene validated with ISH. Excel table containing the name, top blast hit, and primer sequences for gene validated with ISH and which dataset that gene was DE in. 'NS' = not significant. 'NA' = not tested.

CHAPTER 4:

Frizzled3, a membrane-bound Wnt receptor specific to colonial hydrozoans and its implications on the origins of coloniality

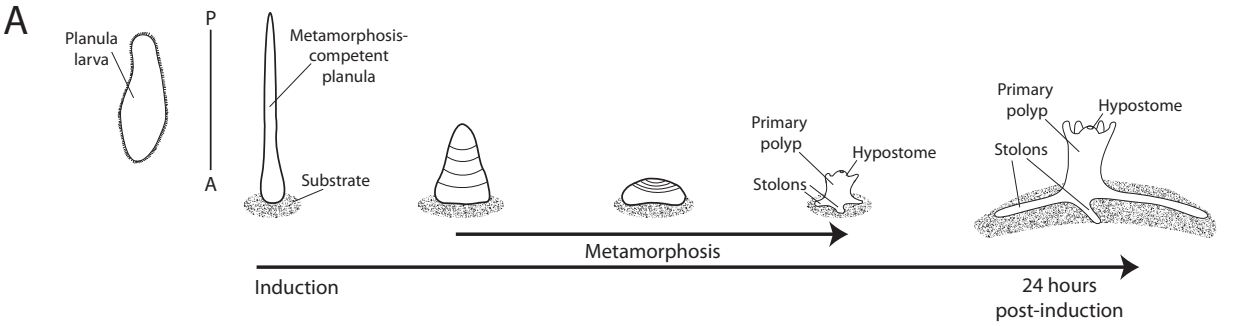
Abstract

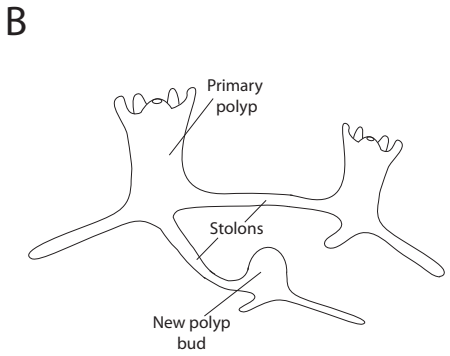
In Hydrozoa, species are either solitary or colonial. These two life history strategies characterize the two major lineages of Hydrozoa: Trachylina comprise mostly solitary species whereas Hydroidolina comprise mostly colonial species. Individual polyps of a colony are connected by the radial growth of epithelial mat or peripheral tube-like structures called stolons. Stolons elongate, branch, and asexually bud new polyps. Mat and stolons function to integrate the colony through continuous epithelia and shared gastrovascular cavity. Although mechanisms of patterning polyps have been well studied, little is known about the signaling processes governing the patterning of colonies. Here we identify a signaling receptor that appears to be specific to colonial hydrozoans and is expressed in tissues specific to colonies. Although the Wnt pathway has been well characterized in the development of the hydrozoan polyp, little is known about its role in colony development. *In situ* hybridization (ISH) of the Wnt membrane-bound receptor *frizzled3* in the colonial hydrozoans *Hydractinia symbiolongicarpus* and *Podocoryna carnea* reveal spatially restricted patterns of expression in the ectoderm of mat and stolons. Phylogenetic analysis of cnidarian Frizzled genes reveal five well supported cnidarian orthologs, *frizzledA*, *frizzled1*, *frizzled2*, *frizzled3*, and *frizzled4*. *Frizzled3* is specific to colonial hydrozoans. *Frizzled3* expression patterns recovered in the colonies of *H. symbiolongicarpus* and *P. carnea* are consistent with it playing a role in regulating the growth of stolons and mat tissue. Our gene tree analysis and whole mount *in situ* hybridization studies, in conjunction with recent phylogenetic studies of hydrozoans, implicate *frizzled3* as playing a role in the evolutionary transition from solitary to colonial hydrozoans.

Introduction

Hydrozoans are known for their complex life cycles that include a benthic colonial stage. However, many hydrozoan species are solitary. Recent phylogenetic studies of hydrozoans have revealed that these two life history strategies characterize two reciprocally monophyletic groups: Trachylina which comprise mostly solitary taxa and the more speciose group, Hydroidolina that comprise mostly colonial taxa [1-3]. Colonial hydrozoan species are composed of individual polyps connected by continuous epithelia, forming the shared gastrovascular cavity. The structures connecting the polyps are sheet-like mat tissue or elongating tube-like stolons. These stolons grow outward from the primary polyp (Figure 1A) and bud more polyps (Figure 1B) to form a colony (Figure 1C).

In addition to members of Trachylina, solitary hydrozoans primarily fall into the hydroidolina clade Aplanulata, which includes the model system *Hydra*. Although Aplanulata was thought to possess colonial species, a recent study of coloniality in the Aplanulata species *Ectopleura larynx* revealed that it is not a true colony as "coloniality" it is not achieved through asexual budding, but through fusion of the offspring to the parent [4]. Previous phylogenetic studies have recovered Aplanulata as derived within Hydroidolina, rendering colonial lineages as paraphyletic and coloniality as ancestral for Hydroidolina (Figure 2)[1-3]. A recent phylogenomic study recovered Aplanulata as monophyletic and sister to the rest of Hydroidolina [5], such that true colonial hydrozoan lineages are monophyletic and coloniality evolved after the divergence of Aplanulata from the rest of the Hydroidolina lineage. This study however had limited taxon sampling with several colonial clades missing and thus this interpretation should be viewed as tentative (Figure 2).


A
Planula larva
P
A
Metamorphosis-competent planula
Substrate
Primary polyp
Hypostome
Stolons
Primary polyp
Hypostome
Stolons
Metamorphosis
Induction
24 hours post-induction


B
Primary polyp
Stolons
New polyp bud

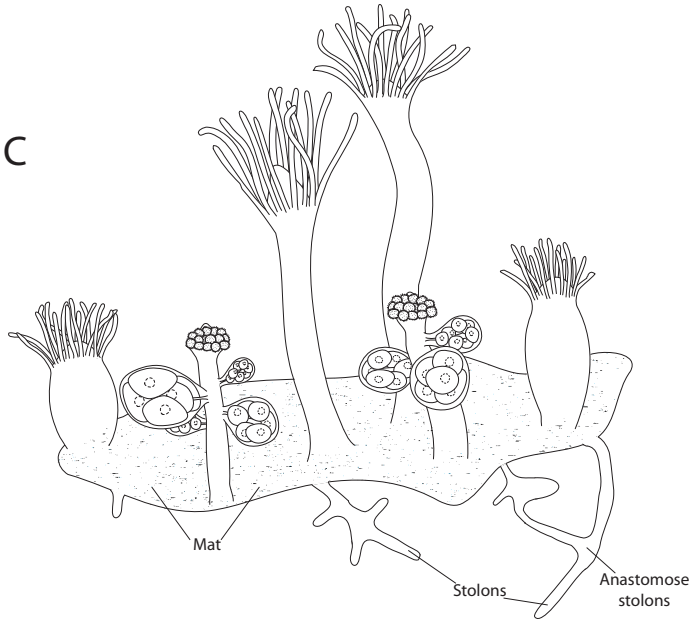

C
Mat
Stolons
Anastomose stolons

Figure 1. *Hydractinia symbiolongicarpus* life cycle.

A) Illustration of metamorphosis in *H. symbiolongicarpus.* About three days after fertilization, the anterior end (labeled 'A') of a metamorphosis-competent planula will attach to the substrate and is able to metamorphose into a polyp. Metamorphosis is induced with a three hour incubation in 116mM CsCl. During metamorphosis, the posterior end (labeled 'P') of the planula compresses and forms concentric rings that will become the hypostome of the polyp. The oral end will then begin to expand and form the hypostome, mouth and tentacles. As this occurs, stolons also begin to branch from the polyp base and by 24 hours the primary polyp is fully developed with branching stolons. B) As stolons continue to branch and elongate, new polyps are produced and form a colony. C). Over time the colony matures and a central continuous epithelial mat forms, from which peripheral stolons can grow, branch, and anastomose.

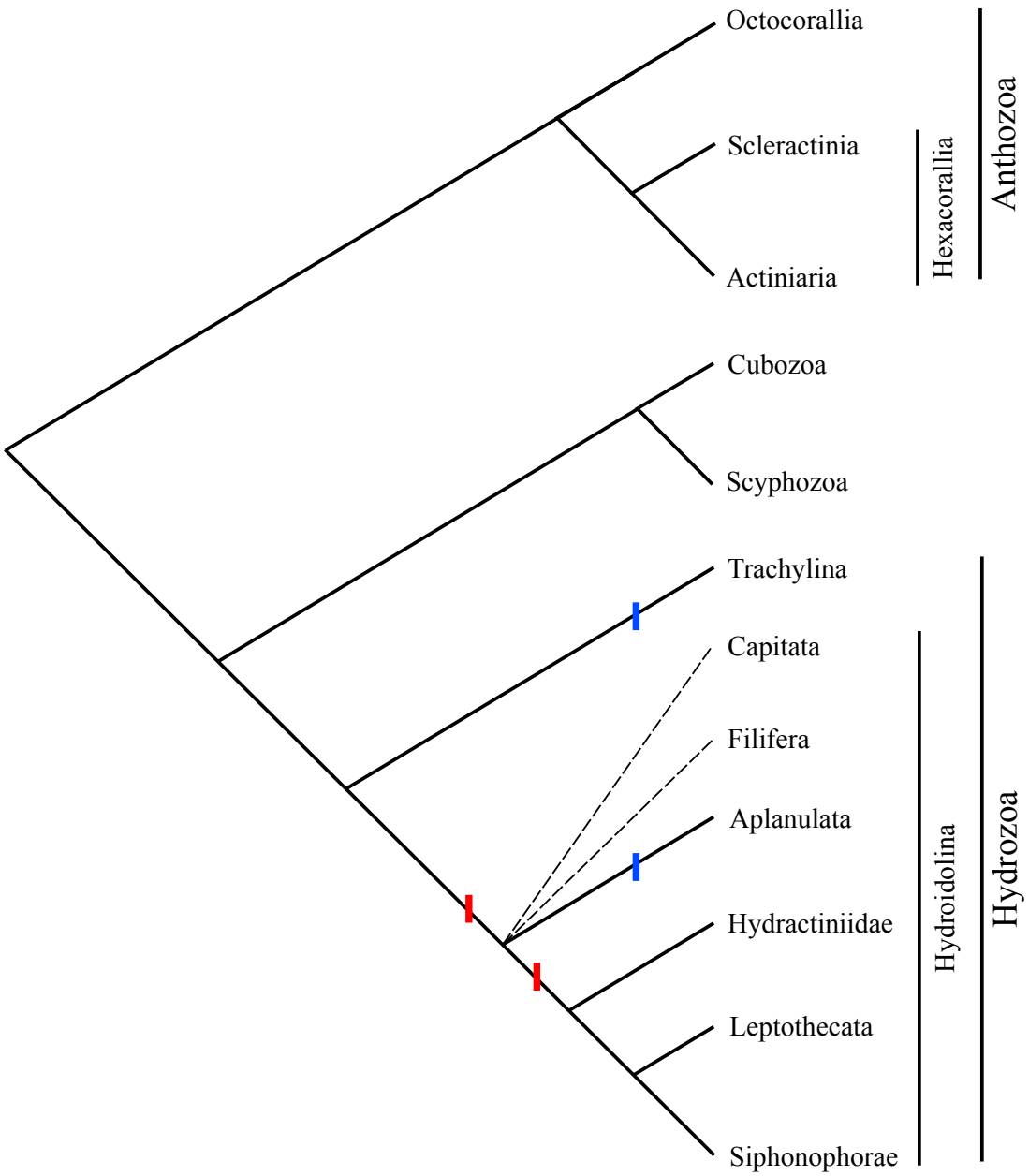


Figure 2. Cnidarian phylogeny.

Cladogram of Cnidaria. Dashed lines represent lineages not sampled in by Felipe et al. [5]. Blue ticks = solitary lineages. Red ticks = possible origins of coloniality.

Although considerable insight has been gained in the phylogenetic patterns of coloniality in Hydrozoa, little is known about the developmental changes that accompanied this innovation (although see Cartwright et al. [6]). The Wnt pathway is a good candidate given that it plays a prominant role in patterning the polyp [7-21] and medusa [9,16,19,22]. Here we implicate the Wnt receptor *frizzled3* in the development of colony specific structures and find that this gene appears to be specific to colonial hydrozoans.

Materials and Methods

Animals

Transplanted colonies of *H. symbiolongicarpus* and *P. carnea* were grown on microscope slides, placed in slide racks and kept in seawater (REEF CRYSTALS, Aquarium Systems) aquaria at room temperature (~21°C) with a salinity of 32 and 29 ppt, respectively. Colonies were fed 2-3-day-old nauplii of *Artemia* three times a week. For *H. symbiolongicarpus*, spawning was induced by light after keeping animals in the dark for ~48 hours. For *P. carnea*, freshly liberated medusae were collected and fed one-day-old nauplii of *Artemia* and then spawning was induced by light after keeping medusae in the dark for ~24 hours. Eggs and sperm were collected and kept in a Petri dish for ~72 hours. By day three, metamorphosis-competent larvae were then incubated for three hours in 116mM CsCl in seawater [23,24]. *P. carnea* stolon regeneration was induced by excising gastrozooids from the colony and kept in Petri dishes on glass cover slips. Prior to tissue collection for whole mount *in situ* hybridization, specimen were starved for three days, anesthetized with menthol crystals, and fixed in 4% paraformaldehyde overnight at 4°C. Fixed tissue was rinsed with and stored in 100% methanol at -20°C.

Probe Synthesis and *in situ* Hybridization

Sequences for *frizzled3* transcripts were identified from previously published transcriptome assemblies of *H. symbiolongicarpus* [25] and *P. carnea* [26]. *Frizzled3* was amplified from each species cDNA using the following PCR primers: *H. symbiolongicarpus* forward 5′-TTTGTCGCACTTCCTCTGCT-3′ and reverse 5′-TCCGCTAGTCACACCTACGA-3′ to obtain a 351bp fragment; *P. carnea* forward 5′-TGGTATGGCATCCGCACTTT-3′ and reverse 5′-CCAACAACAACCGAAGCTGG-3′ to obtain a 420bp fragment. These products

were cloned using the Invitrogen TOPO-TA Cloning Kit and DIG labeled riboprobes were synthesized from clones using the Invitrogen T7/T3 Megascript kit. ISH protocol was adapted from Gajewsky et al. [27]. Hybridization was carried out at 50˚C for 16-18 hours with a probe concentration of .1 ng/µl. Hybridization was detected by immunostaining with anti-DIG-Fab-AP (ROCHE) and NBT/BCIP.

Molecular Phylogenetic Analyses

Cnidarian sequences belonging to the Frizzled gene family with a seven transmembrane domain (Figure 3A) were mined from NCBI's non-redundant (nr) protein database and several published and unpublished cnidarian transcriptomes and genomes. Redundant gene copies were removed with CD-HIT [28,29] and remaining genes were aligned with Mafft [30] using the L-insi alignment algorithm. Ambiguously aligned regions were removed from the alignment with Gblocks [31] under least stringent settings. A maximum likelihood estimate of the Frizzled gene tree was inferred using RAxML [32] on the CIPRES portal [33]. Support was assessed using the rapid bootstrapping algorithm (-f a) with 1,000 bootstrap replicates under the PROTGAMMA+WAG model.

Results and Discussion

The Evolution of the Frizzled Gene Family

A molecular phylogeny of membrane-bound Frizzled genes was generated from 22 cnidarian taxa in order to investigate the evolution of the Frizzled gene family. Frizzled genes included in our analysis were limited to those with a seven-pass transmembrane domain (Figure 3A), excluding secreted Frizzled-related proteins, which have the Wnt binding, cysteine-rich domain (CRD) but lack the seven-pass transmembrane domain [34], which have been implicated as Wnt antagonists during development [22]. Phylogenetic analysis of cnidarian membrane-bound Frizzled genes reveal five well supported cnidarian orthologs, *frizzledA*, *frizzled1*, *frizzled2*, *frizzled3*, and *frizzled4* (Figure 3B). While *frizzled1*, *frizzled2*, and *frizzled4* are represented in all major cnidarian taxonomic clades sampled, *frizzledA* is specific to anthozoans and *frizzled3* is specific to colonial hydrozoans (Figures 2,3B). Phylogenetic placement of two colonial hydrozoan lineages Capitata and non-hydractiniid filiferans is unresolved (Figure 2)[1,2], and genomic and transcriptomic data was not available to search for Frizzled genes in these lineages. While discovery of *frizzled3* in these lineages would further confirm that *frizzled3* evolved alongside coloniality, resolving whether *frizzled3* evolved after Aplanulata diverged or was subsequently lost as Aplanulata transitioned back to a solitary form awaits further phylogenetic resolution within Hydroidolina.

Spatial Expression of *frizzled3*

As *frizzled3* appears to be a synapomorphy of colonial hydrozoans (Figure 3B), we chose to analyze the spatial expression of *frizzled3* during early colony development of *H. symbiolongicarpus* and *P. carnea* with ISH. During metamorphosis, the posterior end of the

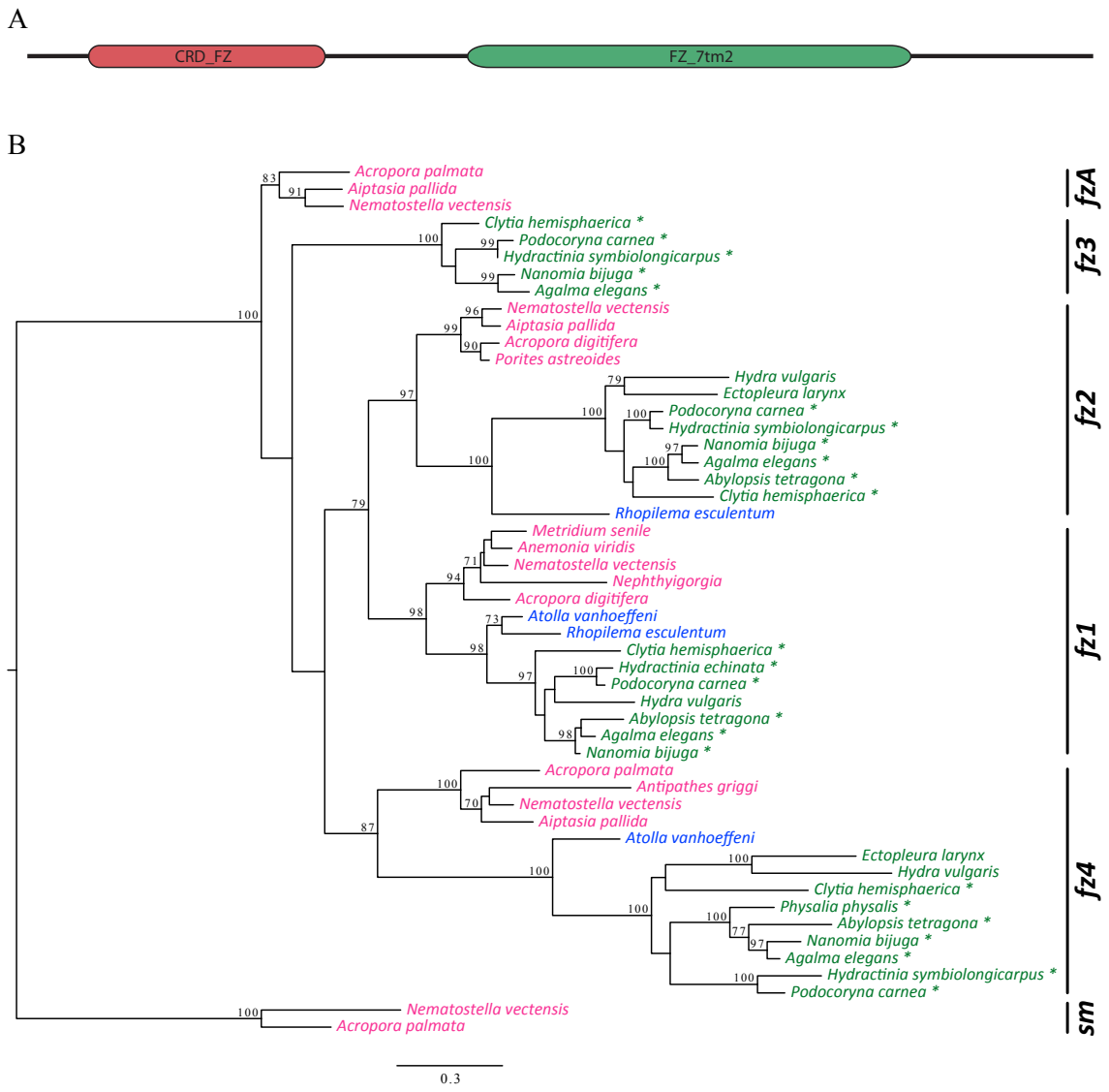


Figure 3. Frizzled phylogeny.

A) Membrane-bound Frizzled schematic, showing the cysteine-rich Frizzled (CRD_FZ) and seven-pass transmembrane (FZ_7tm2) domains. B) ML estimate of cnidarian membrane bound Frizzled (*fz*) gene tree. ML bootstrap support values not shown for nodes with less than 70% support. Tree is rooted on two anthozoan copies of Smoothened (*sm*). Green = hydrozoans. Blue = scyphozoans. Pink = anthozoans. '*' = colonial hydrozoans.

planula compresses, telescoping down and forming concentric rings that will become the hypostome (structure that houses the mouth) of the polyp (Figure 1A). As metamorphosis continues, the oral end will begin to expand, forming tentacles and the mouth (Figure 1A). As the oral end elongates, stolons begin to branch from the polyp base and by 24 hours the primary polyp is fully developed with branching stolons (Figure 1A). In *H. symbiolongicarpus*, *frizzled3* expression was detected in stolon tissues 24 hours after the induction of metamorphosis (Figure 4A-C). *Frizzled3* expression at this stage was specific to the ectoderm of the more proximal portions of the newly formed stolons, and was excluded from the stolon tips and polyp base (Figure 4A,B). This pattern continues throughout colony ontogeny, and in mature colonies of *H. symbiolongicarpus*, *frizzled3* expression remains primarily restricted to the stolon tissue (Figure 4D,E). As colony expansion continues, branched stolons anastomose and the ectoderm begins to merge, forming a continuous ectodermal mat and eliminating the *frizzled3* expression boundary along the ectodermal edge of the stolons (Figure 4D). More distal portions of the stolons show a more discontiguous expression pattern (Figure 4E), in what appears to be stem cells called interstitial cells (*i*-cells) that are migrating along the stolons. In non-stoloniferous colonies, accumulated expression can be seen around the periphery of the mat (Figure 4F,G). Furthermore, there are pockets of expression around *i*-cell clusters within the mat tissues near the edge of the colony (Figure 4G).

Similar expression patterns were observed in the stolons emanating from the primary polyp (Figure 5A) and mature, stoloniferous colonies of *P. carnea* (not shown). During stolon regeneration in *P. carnea*, *frizzled3* expression was examined in polyps at 24, 48, and 72 hours after excision from the colony. At 24 hours post excision, stolons had not formed and *frizzled3* expression was not detected (not shown). By 48 hours, stolons had regenerated and ISH revealed

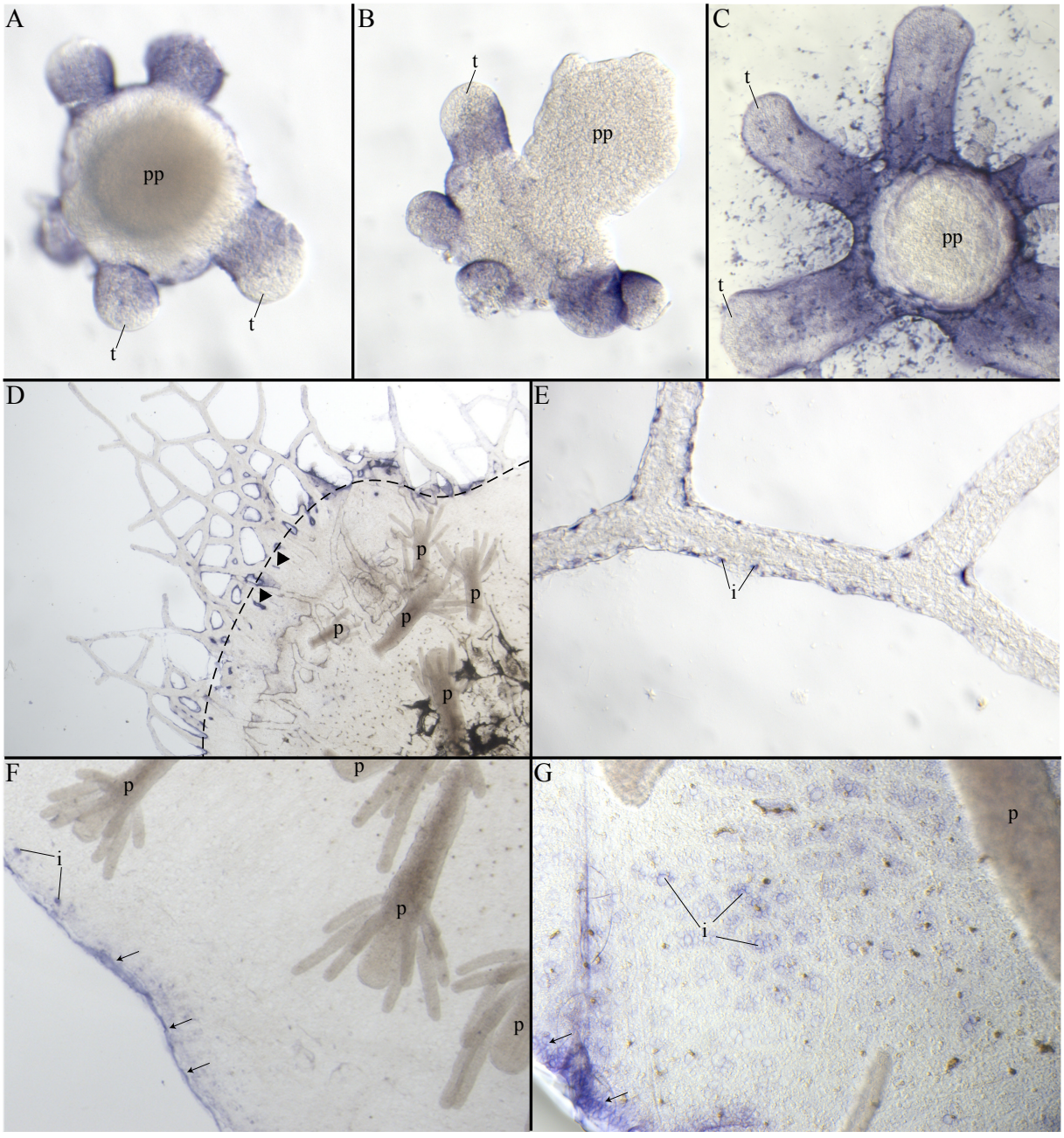

A
pp
t
t
B
t
pp
C
t
pp
t
D
p
p
p
p
p
E
i
F
p
p
i
p
p
p
G
p
i

Figure 4. Whole mount *in situ* hybridization of *frizzled3* in *H. symbiolongicarpus*.

A) Aboral view of primary polyp and newly budded stolons 24 hours after induction of metamorphosis showing *frizzled3* expression in proximal region of stolons (in blue). B) Lateral view of primary polyp and newly budded stolons 24 hours after induction of metamorphosis. C) Aboral view of a primary polyp and stolons 48 hours after induction of metamorphosis showing *frizzled3* expression throughout stolons. D) Stoloniferous colony showing *frizzled3* expression at the edges of the mat and out ectoderm of the stolons. E) Higher magnification of stolons more distal to the colony. F) Mat tissue showing expression at the distal ends. G) Higher magnification of colony mat. Legend: i – *i*-cells; pp- primary polyp; p – polyps; t – stolon tip. Arrows indicate areas of accumulated expression along the edge of the mat. Arrowheads mark regions where the ectoderm is fusing over stolons, form a continuous ectodermal mat. Dashed line in panel D roughly marks the boundary between the mat and peripheral stolons.

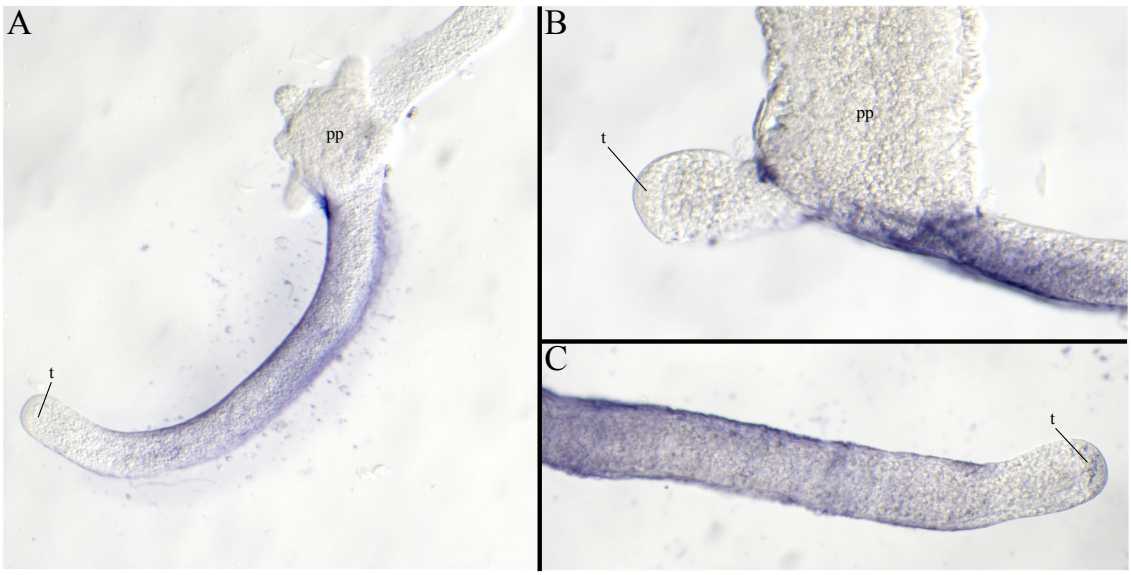


Figure 5. Whole mount *in situ* hybridization of *frizzled3* in *P. carnea*.

A) Oral view of a primary polyp and newly budded stolons 24 hours after induction of metamorphosis showing *frizzled3* expression (in blue) at the proximal region of the stolons. B) Regenerating stolons 48 hours after removal from the colony. C) Regenerating stolon 72 hours after removal from the colony. Legend: pp- primary polyp; t – stolon tip.

frizzled3 expression in the ectoderm of the regenerated stolons (Figure 5B) and continued in 72 hour regenerated stolons (Figure 5C). Similar to expression in stolons of *H. symbiolongicarpus*, *frizzled3* expression was restricted to the proximal portion of the stolons and excluded from the stolon tips throughout all stages of stolon regeneration (Figure 5B). Similar expression of *frizzled3* during normal development and regeneration suggests the same pathways are regulating both of these processes.

Wnt signaling and *frizzled3* in other life cycle stages

While *frizzled3* expression has been observed in other hydrozoans, it had not yet been explored in the context of colony ontogeny. Previous studies have reported *frizzled3*'s role in embryology [16,35,36], polyp patterning [16,22], and medusae development [16,22]. In the leptothecate hydrozoan *Clytia hemisphaerica*, *frizzled3* expression appears to maintain the aboral identity of the embryo [16,35,36], yet is expressed at the oral ends of both the polyp and medusae [16]. Similarly, in *P. carnea*, *frizzled3* is expressed at the tip of the hypostome (mouth) of the polyp and at the oral end of the medusae through development (see Chapter 2). These pleiotropic expression patterns of *frizzled3* suggest repeated co-option of this gene in different developmental contexts to mediate different Wnt ligands. Given that there are 10 hydrozoan Wnt orthologs [37] and only four membrane-bound Wnt receptors (Frizzleds) to transduce the Wnt signal into the cell, it is not surprising that Frizzled receptors are acting in different contexts.

While there has been no study showing a specific Wnt ligand's role in stolon regeneration and development, *Wnt16*, *Wnt11a*, and *Wnt2* are expressed in the stolons of *H. echinata* [37]. Similar to *frizzled3*, *Wnt11a* was only recovered in colonial hydrozoans by Hensel et al. (2014), although this could be an artifact of low sampling. Ectopic induction of canonical Wnt signaling

with azakenpaullone treatments shows that canonical Wnt signaling up-regulates *Wnt2* expression while down-regulating *Wnt11a* expression in the stolons [37]. Likewise, throughout embryonic development in *C. hemisphaerica*, *Wnt3* (the canonical Wnt ligand) and *frizzled1* are co-expressed at the oral pole of the embryo and actively maintain the oral/aboral boundary by inhibiting *frizzled3* [16,35,36]. Furthermore, in *H. echinata* ectopic induction of canonical Wnt signaling during regeneration does prevent stolon development during metamorphosis [19]. If this same mechanism is also active during stolon development and regeneration, *frizzled3* is the receptor of *Wnt11a* and these two genes are responsible for the development colony growth.

Conclusion

The evolution of coloniality within Hydrozoa was a key innovation that enabled them to better compete for substrate in the benthos [38,39], elaborate their colonial organization through a division of labor [40-42], and even enter the pelagic realm as complex integrated colonies displayed by siphonophores and porpitiids [43]. *Frizzled3* expression patterns recovered in the colonies of *H. symbiolongicarpus* and *P. carnea* are consistent with it playing a role in regulating colony patterning and growth as they are expressed in a spatially restricted manner in colony specific tissues. Moreover, a recent study of Wnt gene expression in *H. echinata* found compelling evidence that *Wnt11a* might be a Wnt ligand regulating stolon growth [37] and that these act antagonistically to the canonical Wnt pathway patterning the polyp.

Our phylogenetic analyses of the Frizzled gene family suggests that there are three orthologs in cnidarians and a fourth, *frizzled3*, is specific to colonial hydrozoans. Our study, in conjunction with recent phylogenomic evidence suggests that coloniality evolved once in Hydrozoa, and that the transition to coloniality was accompanied by a gene duplication and co-option of the Frizzled receptor for patterning and grown of colony specific tissues.

Acknowledgments

We would like to thank Felipe Zapata and Casey Dunn for the phylogeny in Figure 2 and participants of the Cnidarian Tree of Life project, where many of the Frizzled sequences came from. We thank Neil Blackstone and Leo Buss for lab cultures of *P. carnea*, Uri Frank for discussion and help with the ISH protocol, and Kirsten Jensen illustrations in Figure 1. This work was supported by NSF grants EF-0531779 and DEB‐095357 to PC.

References

1. Collins AG, Winkelmann S, Hadrys H, Schierwater B: **Phylogeny of Capitata and Corynidae (Cnidaria, Hydrozoa) in light of mitochondrial 16S rDNA data.** *Zool. Scripta* 2005, **34**:91-99.
2. Cartwright P, Evans NM, Dunn CW, Marques AC, Miglietta MP, Schuchert P, Collins AG: **Phylogenetics of Hydroidolina (Hydrozoa: Cnidaria).** *J. Mar. Biol. Assoc.* 2008, **88**:1163-1172.
3. Cartwright P, Nawrocki AM: **Character evolution in Hydrozoa (phylum Cnidaria).** *J. Integ. Comp. Biol.* 2010, **50**:456-472.
4. Nawrocki AM, Cartwright P: **A novel mode of colony formation in a hydrozoan through fusion of sexually generated individuals.** *Curr. Biol.* 2012, **22**:1-5.
5. Zapata F, Goetz F, Smith S, Church S, Sanders SM, Ames CL, McFadden CS, France S, Daly M, Collins AG, Haddock SHD, Dunn CW, Cartwright P: **Phylogenomic analyses of Cnidaria.** *In preparation*.
6. Cartwright P, Schierwater B, Buss LW: **Expression of a *Gsx* parahox gene, *Cnox-2*, in colony ontogeny in *Hydractinia* (Cnidaria: Hydrozoa).** *J. Exp. Zool.* 2006, **306B**:460-469.
7. Hobmayer B, Rentzsch F, Kuhn K, Happel CM, von Laue CC, Snyde P, Rothbacher U, Holstein TW: **WNT signalling molecules act in axis formation in the diploblastic metazoan Hydra.** *Nature* 2000, 407:186–189.
8. Müller WA, Teo R, Möhrlen F: **Patterning a multi-headed mutant in *Hydractinia*: enhancement of head formation and its phenotypic normalization.** *Int. J. Dev. Biol.* 2004, **48**:9-15.
9. Müller W, Frank U, Teo R, Mokady O, Guette C, Plickert G: **Wnt signaling in hydroid**

development: ectopic heads and giant buds induced by GSK-3*ß* inhibitors. *Int. J. Dev. Biol.* 2007, **51**:211-220.

10. Broun M, Gee L, Reinhardt B, Bode HR: **Formation of the head organizer in hydra involves the canonical Wnt pathway.** *Dev*. 2005, **132**:2907-2916.
11. Guder, C, Pinho S, Nacak TG, Schmidt HA, Hobmayer B, Niehrs C, Holstein TW: **An ancient Wnt-Dickkopf anatagonism in *Hydra*.** *Dev*. 2005, **133**:901-911.
12. Guder C, Philipp L, Lengfeld T, Watanabe H, Hobmayer B, Holstein TW: **The Wnt code: cnidarians signal the way.** *Oncogene* 2006, **25**:7450–7460.
13. Augustin R, Franke A, Khalturin K, Kiko R, Siebert S, Hemmrich G, Bosch TCG: **Dickkopf related genes are compnents of the positional value gradient in Hydra.** *Dev. Biol.* 2006, **296**:62-70.
14. Plickert G, Jacoby V, Frank U, Müller WA, Mokady O: **Wnt signaling in hydroid development: Formation of the primary body axis in embryogenesis and its subsequent patterning.** *Dev. Biol.* 2006, **298**:368-378.
15. Khalturin K, Anton-Erxleben F, Milde S, Plötx C, Wittlieb J, Hemmrich G, Bosch TCG: **Transgenic stem cells in *Hydra* reveal an early evolutionary origin for key elements controlling self-renewal and differentiation.** *Dev. Biol.* 2007, **309**:32-44.
16. Momose T, Houliston E: **Two oppositely localized frizzled RNAs as axis determinants in cnidarian embryo.** *PLoS Biol.* 2007, **5**:e70.
17. Lengfeld T, Watanabe H, Simakov O, Lindgens D, Gee L, Law L, Schmidt HA, Özbek S, Bode H, Holstein TW: **Multiple Wnts are involved in *Hydra* organizer formation and regeneration.** *Dev. Biol.* 2009, **330**:186-199.
18. Philipp I, Aufschnaiter R, Özbek S, Pontasch S, Jenewein M, Watanabe H, Rentzsch F,

Holstein TW, Hobmayer B: **Wnt/*ß*-catenin and noncanonical Wnt signaling interact in tissue evagination in the simple eumetazoan *Hydra*.** *PNAS* 2009, **106**:4290-4295.

19. Duffy DJ, Plickert G, Kuenzel T, Tilmann W, Frank U: **Wnt signaling promotes oral but suppresses aboral structures in *Hydractinia* metamorphosis and regeneration.** *Dev.* 2010, **137**:3057-3066.
20. Gee L, Hartig J, Law L, Wittlieb J, Khalturin K, Bosch TCG, Bode HR: ***ß*-catenin plays a central role in setting up the head organizer in hydra.** *Dev. Biol.* 2010, **340**:116-124.
21. Nakamura Y, Tsiairis CD, Özbek S, Holstein TW: **Autoregulatory and repressive inputs localize *Hydra Wnt3* to the head organizer.** *PNAS* 2011, **108**:9137-9142.
22. Nawrocki AM, Cartwright P: **Expression of Wnt pathway genes in polyps and medusa-like structures of *Ectopleura larynx* (Cnidaria: Hydrozoa).** *Evol. & Dev.* 2013, **15**:373-384.
23. Müller WA, Buchal G: **Metamorphose-Induktion bei Planulalarven. II. Induktion durch monovalente Kationen: Die Bedeutung des Gibbs-Donnan-Verhältnisses und der Na^+/K^+-ATPase.** *Wilhelm Roux's Arch.* 1973, **173**:122-135.
24. Frank U, Leitz T, Müller WA: **The hydroid *Hydractinia*: a versatile, informative Cnidarian representative.** *BioEssays* 2001, **23**:963-971.
25. Sanders SM, Shcheglovitova M, Cartwright P: **Differential gene expression between functionally specialized polyps of the colonial hydrozoan *Hydractinia symbiolongicarpus* (Phylum Cnidaria).** *BMC Genomics* 2014, **15**:406.
26. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GBEH00000000.

27. Gajewsky M, Leitz T, Schlossherr J, Plickert G: **LWamides from Cnidaria constitute a novel family of neuropeptide with morphogenetic activity.** *Roux's Arch. Dev. Biol.* 1996, **205**:232– 242.

28. Li W, Godzik A: **Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences.** *Bioinformatics* 2006, **22**:1658-1659.

29. Fu L, Niu B, Zhu Z, Wu S, Li W: **CH-HIT: accelerated for clustering the next generation sequencing data.** *Bioinformatics* 2012, **28**:3150-3152.

30. Katoh K, Kuma K, Toh H, Miyata T: **MAFFT version 5: improvement in the accuracy of multiple sequence alignment**. *Nucl. Acids Res.* 2005, **33**:511-518.

31. Castresana J: **Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis.** *Mol. Biol. Evol.* 2000, **17**:540-552.

32. Stamatakis A, Hoover P, Rougemont J: **A fast bootstrapping alogorithm for the RAxML web-servers.** *Sys. Biol.* 2008, **57**:758-771.

33. Miller MA, Pfeiffer W, Schwartz T: **Creating the CIPRES Science Gateway for inference of large phylogenetic trees.** *GCE* 2010, 1-8.

34. Leclere L, Rentzsch F: **Repeated evolution of identical domain architecture in metazoan netrin domain-containing proteins.** *Genome Biol. Evol.* 2012, **4**:883-899.

35. Momose T, Derelle R, Houliston E: **A maternally localized Wnt ligand required for axial patterning in the cnidarian *Clytia hemisphaerica*.** *Dev.* 2008, **135**:2105-2113.

36. Amiel A, Houliston E: **Three distinct RNA localization mechanisms contribute to oocyte polarity establishment in the cnidarian *Clytia hemisphaerica*.** *Dev. Biol.* 2009, **327**:191-203.

37. Hensel K, Lotan T, Sanders SM, Cartwright P, Frank U: **Lineage-specific evolution of cnidarian Wnt ligands.** *Evol. & Dev.* 2014, **16**:259-269.

38. Jackson JBC: **Competition on marine hard substrate: the adaptive significance of solitary and colonial strategies.** *Am. Nat.* 1977, **111**:743-767.

39. Coates AG, Jackson JBC: **Morphological themes in the evolution of clonal and aclonoal marine invertebrates.** In *Population Biology and Evolution of Clonal Organisms*, J.B.C. Jackson, L.W. Buss, and R.E. Cook, eds. (New Haven, CT: Yale University Press), pp.67-106.

40. Berrill NJ: **Growth and form in gymnoblastic hydroids; polymorphism within the Hydractiniidae.** *J. Morph.* 1953, **92**:241-272.

41. Muller W: **Experimentelle Untersuchungen Über Stockentwicklung, Polypendifferenzierung, und sexualchimären bei *Hydractinia echinata*.** *Wilhelm Roux' Arch Entwicklungsmech. Org.* 1964, **155**:181-268.

42. Burnett AL, Sindelar W, Diehl N: **An examination of polymorphism in the hydroid *Hydractinia echinata*.** *J. Mar. Biol. Assoc. U.K.* 1967, **47**:645-658.

43. Mackie, GO: **Siphonophores, bud colonies, and superorganisms.** In *The Lower Metazoa: Comparative Biology and Phylogeny*, E.C. Dougherty, Z.N. Brown, E.D. Hanson, and W.D. Hartman, eds. (Berkley, CA: University of California Press), pp. 329-337.

CONCLUSION

In this dissertation, I have developed workflows for analyzing RNA-Seq data in both an intra- and interspecific comparative context. I have done this using data generated from various tissues and life cycle stages of two emerging model hydrozoan species, *H. symbiolongicarpus* and *P. carnea*. Throughout the chapters of this dissertation, I have characterized the transcriptomes of each species and used unbiased differential expression analyses to identify candidate genes and gene pathways that are potentially responsible for key transitions in hydrozoan coloniality, polyp polymorphism (complex coloniality), and medusae evolution. Furthermore, using whole mount *in situ* hybridization to characterize the spatial expression of various candidates genes, I validated each approach showing expression consistent with their role in the development of a particular tissue or life cycle stage.

The results presented in this dissertation reveal the power of these unbiased genomic/transcriptomic methods over traditional candidate gene approaches to address longstanding questions of hydrozoan morphology and evolution. Identification of *APLP* in the interspecific DE analysis of Chapter 3 is a prominent example of this. While many prior gene expression studies have implicated Wnt signaling in medusae evolution, none of the major Wnt signaling components were recovered in the analysis. As noted in Chapter 3, *APLP* plays a critical role in regulating *hedgehog* and Wnt signaling during wing development in *Drosophila* by aiding in the long-range dispersal of these morphogens. Although not functionally tied to Wnt signaling in cnidarians, recovery of *APLP* further implicates Wnt signaling in medusae evolution. As seen in *P. carnea*, *APLP* was highly expressed in the endodermally derived gastric structures throughout medusae ontogeny while no expression was detected in the sporosac of *H. symbiolongicarpus*. If *APLP* and Wnt signaling are connected in hydrozoans, down-regulation of

APLP could be involved in the initial truncation of medusae development, limiting the dispersal of the Wnt ligand. This pattern is consistent considering the reduced expression of Wnt signaling components, *Wnt3* and *frizzled*, in reduced gonophores of *Ectopleura* and *Hydractinia*.

Of further note, while phylogenies are important for recognizing evolutionary patterns of character transitions, understanding these complex patterns of character evolution will ultimately come from insight into their development. As noted throughout this dissertation, these features have been extensively researched using phylogenetic methods. Phylogenies and the methods used to analyze them are important for not only setting up these questions, but also lend the framework with which to test our hypotheses and expand on the role of development in evolution. It is through the integration of phylogenomic, transcriptomic, and developmental research that I have been able to explore these topics in more detail.

Future research into hydrozoan life cycle evolution should focus testing the function of the genes identified in this dissertation. Current functional methods established in cnidarian developmental research include RNAi (dsRNA and morpholinos), transgenics (microinjection, TALENs, and CRISPR/Cas9), and pharmaceutical misexpression (e.g. ectopic induction of canonical Wnt signaling via paullone and LiCl treatments). Furthermore, increased taxonomic sampling for phylogenomic studies is another important endeavor and one that is already in progress. Expanded sampling of key cnidarian clades will further uncover novel gene gain and loss that can be attributed to evolution of fundamental life history characters. Moreover, expansion of the interspecific differential expression methods outlined in Chapter 3 to accommodate more than two taxa will greatly advance the results presented here and begin to address polarity of these evolutionary transitions (e.g. whether up-regulation of a gene coincides with character gain or down-regulation likely responsible for character loss). As noted in Chapter

3, this lends several more challenges including independent contrasts and the effects of paralogous gene expression.