Re-examining the relationships among families of Pectinoidea and molecular machinery of the visual cycle in molluscs

by

Garrett Dalton Smedley

A dissertation submitted to the graduate faculty

in partial fulfillment of the for the degree of

DOCTOR OF PHILOSOPHY

Major: Interdepartmental Genetics and Genomics

Program of Study Committee:
Jeanne Serb, Major Professor
Tracy Heath
Dennis Lavrov
Don Sakaguchi
Nicole Valenzuela

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2020

DEDICATION

I would like to dedicate this dissertation to my parents, Gary and Charlene Smedley. Your overwhelming love and support over a very long academic career and everything leading up to this has built me into the person that I am. I cannot thank you both enough for everything you have done for me but vow to one day be able to repay the favor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I owe a great deal of gratitude and my professional success to my mentor, Dr. Jeanne Serb, who took a risk in taking me under her wing. Despite the trials and tribulations the years have presented, Jeanne consistently served as a reliable source of support and a well-versed sounding board to help me reach my professional goals. I also owe great thanks and respect to my lab mates who have come and gone during my tenure in the lab: Dr. Davide Faggionato and Dr. Kyle McElroy. I thank you both for the unending support in intellectual spit-balling, manuscript edits, and much desired breaks from the monotony of protein expression experiments. I would also like to thank my other committee members, Dr. Tracy Heath, Dr. Dennis Lavrov, Dr. Don Sakaguchi, and Dr. Nicole Valenzuela for their guidance and support throughout my research.

I must also thank my wealth of support outside the university: my parents to whom this dissertation is dedicated, my sister Kelsey Smedley, my long list of D&D friends, and my partner in crime Taylor Ebert. Kelsey: from the ups and downs, I cannot express how lucky I am to have a hardheaded yet impossible not to love best friend for a sibling. Thank you for sharing my impeccable sense of humor and laughing with me when the rest of the world could not hope to understand. To all of my D&D friends whose dedications would fill a second dissertation. It is an honor to know all of you, and the stories we have written together, truly, have kept me sane over the years and can only hope we can continue to create together. And most importantly Taylor: I simply cannot imagine fighting through this with anyone else. The understanding, whole-hearted kindness, and caring individual you are has kept me level-headed and calm

even at the worst of times. All too often, you have been my reason for getting out of bed and to continue moving forward. You have earned this arguably more than I have.

ABSTRACT

Biologists have long used eyes as a model to study the evolution of complex traits, and the growing number of molecular datasets have allowed new insights into the conservation and function of the molecular machinery underlying photosensitivity across animals. However, molluscs, the second largest animal phylum, have been largely ignored in these studies despite possessing morphologically diverse eyes. Here, I investigated three levels of biological organization in molluscan eyes: organs, pathways, and proteins. I began by generating a robust phylogenetic hypothesis of the bivalve superfamily Pectinoidea to determine when the mirror-type eye of scallops (Pectinidae) originated. From this study, I propose a novel topology in which Propeamussiidae is non-monophyletic, where a subset of species resolve as sister to the Pectinidae and second species group, including the type species *Propeamussium dalli,* are sister to the Spondylidae. This relationship suggests a single origin of eyes prior to Pectinoidea with multiple instances of loss throughout the superfamily. Next, I assembled and searched available molluscan transcriptomes using known visual cycle proteins of vertebrates and insects to expand on the current hypothesis of the retinoid visual cycle in molluscs. The search results were then divided by protein family and used to develop protein phylogenies with vertebrate and insect anchor sequences to suggest putative homology of function of molluscan blast results. From this study, I was able to propose a new, light independent retinoid visual cycle for molluscs that includes proteins homologous to both vertebrates and insects. Interestingly, I was unable to identify a homologous isomerohydrolase or retinyl storage pathway in molluscs, thus future studies will require experimental work to determine a possible lineage-specific pathway. Finally, I

investigated the relationship between genotype and phenotype in opsins by mutating targeted amino acids of interest and examining how these alternations affected opsin function when expressed *in vitro*. Comparing two closely related scallop retinochromes, I identified and mutated sites lining the binding pocket of the retinochrome that may interact with the chromophore, but were not conserved between the two retinochrome samples. These experiments showed that sites lining the binding pocket may be responsible for fine tuning the spectra of opsin proteins and that size and class of amino acid side chains may be responsible for changes in the λ_{max}. This study highlights the potential of retinochrome as a model in mapping the relationship of genotype and phenotype in opsins which can be used to build more in depth and accurate prediction models of photopigments.

CHAPTER 1. GENERAL INTRODUCTION

Overview

Animals use light to modulate a variety of behaviors in navigation, mate selection, predator avoidance, and detecting environmental cues to enhance survival and fitness (Nilsson 2009, 2013). The structures responsible for light perception within animals are equally varied (Serb and Eernisse 2008; Land and Fernald 1992), and their components often have complex evolutionary histories of co-option and adaptation (Vopalensky and Kozmik 2009; Oakley and Speiser 2015). As a result, eyes have evolved over 60 times across animals (von Salvini-Plawen and Mayr 1977), and while eyes vary in size, shape, and complexity, much of the molecular machinery involved in light perception and resetting phototransduction pathways have been conserved (Speiser et al. 2014; Ramirez et al. 2016), such as opsins (Terakita 2005; Ramirez et al. 2016; Shichida and Matsuyama 2009; Porter et al. 2012) and retinal shuttle proteins (Gonzalez-Fernandez 2002). Much of our understanding of these visual processes are from work in the vertebrates (Lamb, Collin, and Pugh 2007). Despite some of the nuances of eye construction, vertebrates possess a single eye type (e.g., a “camera-type eye”) that originated once in the last common ancestor of lampreys and jawed vertebrates (500 Mya) (Land and Fernald 1992; Lamb, Pugh, and Collin 2008). However, to understand whether the numerous eye types have converged on a limited number of molecular pathway solutions or multiple pathways are utilized by the different eye types to accomplish the same task, we must look at lineages of organisms in which eyes have evolved multiple times.

Mollusca Phylum

The phylum mollusca is the second largest phylum of animals, second to arthropods, composed of eight extant classes (Ponder and Lindberg 2008; Stöger et al. 2013) with some of the greatest diversity of eye types (von Salvini-Plawen and Mayr 1977; Serb and Eernisse 2008). Eyes are described in four of the eight classes (Gastropoda, Bivalvia, Polyplacophora, and Cephalopoda) across at least seven to as many as 11 lineages. Eye structures vary in number, location, and resolution even between closely related species. Molluscan eyes vary greatly in complexity with structures ranging from simple cup/pit eyes to closed lensed eyes analogous to vertebrate eyes, compound eyes similar to insects, pinhole eyes, and eyes with a mirror-like crystalline layer. Pinhole type eyes are found in nautiloids (cephalopoda) and are open to the environment allowing little light to enter without an iris or lens to focus the light. Coleoid cephalopods possess camera-type eyes. Camera-type eyes are analogous to those found in vertebrates, containing an iris, cornea, lens, and vitreous cavities, but cephalopod camera eyes also possess photoreceptor cells which are rhabdomeric as opposed to the ciliary photoreceptors found in vertebrates. Gastropods possess cephalic eyes ranging from simple pit eyes to lensed eyes (reviewed in (Zieger and Meyer-Rochow 2008).

Within molluscs, bivalves have the greatest diversity of eyes (reviewed in (Morton 2001)). Most of the eye types are non-cephalic and are serially repeated along the epithelial tissue lining the shell. For example, compound eyes are found in ark clams (Bivalvia: Arcidae) and share many structural similarities to the compound eyes of arthropods with an independent evolutionary origin (Waller 1980). The compound eyes

of ark clams are thought to perceive movement but the numerous rows of eyes along the mantle increase an ark clam's sensitivity to cues in their environment (Nilsson 1994). Ark clams also possess simple cup eyes composed of photoreceptor cells blocked by pigment cells. Some of these eyes contain lenses, such as ocelli found in a few chitons (Moseley 1885). Scallops (Bivalvia: Pectinidae) possess a unique double retina, mirror eye which contains a lens that focuses light causing it to pass through the retina before being reflected by a guanine crystalline mirror back at the retina allowing perception (Land 1965).

Pectinidae (Rafinesque, 1815) is one of the most diverse families of bivalves with over 250 extant species distributed across a variety of environments. This biological diversity has allowed scallops to serve as a model in morphological (Serb et al. 2011, 2017; Sherratt et al. 2016; Stanley 1970), ecological (Millward and Whyte. 1992; Alejandrino, Puslednik, and Serb 2011; Guderley and Tremblay 2013; Hayami 1991; Tremblay, Samson-Dô, and Guderley 2015) and molecular (Bieler et al. 2014; Matsumoto and Hayami 2000; Steiner and Hammer 2000) studies. The fossil record describing bivalves is quite robust given the wide array of shell structures allowing time calibration of phylogenomic analyses (Waller 2006). However, morphological and molecular studies have led to conflicting hypotheses on the relationships between lineages of scallops making it difficult to understand the evolution of complex traits in scallops (Bieler et al. 2014; Dijkstra and Maestrati 2012; Waller 1991, 2006). Scallop eyes have been an effective model to understand the evolution of complex traits (Piatigorsky 2008; Speiser and Johnsen 2008; Morton 2008). Yet, in order to explain the

context of eye evolution in scallops, such as the number of times eyes have evolved, we need a better understanding of how the Pectinidae fits in the superfamily Pectinoidea.

Phototransduction

Despite the range of morphologically complex photoreceptive systems in animals, all use a photopigment to absorb light and mediate a signaling cascade. In most animals, the photopigment is composed of an opsin, a seven transmembrane G-protein coupled receptor (GPCR), and its retinoid chromophore (a vitamin A derivative). The retinal chromophore binds to a lysine residue in the seventh transmembrane helix (TM7) of the opsin creating an unstable Schiff-base (Hara-Nishimura et al. 1993). To stabilize the photopigment, glutamic acid in either TM3 of vertebrate and invertebrate rhodopsin or the extracellular loop IV-V of squid retinochrome (RTC) acts as a counterion in the binding pocket (Terakita, Yamashita, and Shichida 2000). With the interaction of a photon of light, the opsin undergoes a conformational change initiating the release of specific G-proteins resulting in a signaling cascade. The signaling cascade becomes an action potential leading to interpretation by nervous tissues and ultimately a physiological response. This process is phototransduction (Shichida and Matsuyama 2009).

Phototransduction is composed of two processes. First, the opsin absorbs a photon of a defined wavelength causing the attached retinal chromophore to isomerize from a functional 11-*cis* to an all-*trans* state. This photoisomerization results in a conformational change of the opsin protein allowing release of its coupled G-protein beginning a signaling cascade. The second half of phototransduction is the resetting of the photopigment. The isomerization of the retinal causes a conformational change of

the opsin resulting in an intermediate form of the photopigment, commonly referred to as meta-opsin. Opsins create either bistable or monostable photopigments where bistable photopigments maintain their stability and retinal isomer upon creation of the meta-opsin (Hubbard and St.George 1958). Monostable photopigments, however, become thermodynamically unstable upon creation of the meta-opsin causing release of the retinal isomer (Tsukamoto and Terakita 2010). In order for the photopigment to become functional again, the opsin must re-bind a functional isomer of the retinal, 11-*cis* or all-*trans* retinal, depending on the opsin type. This requires a secondary path to ensure the availability of functional retinal isomers for phototransduction to occur. The isomerization pathway of all-*trans* to 11-*cis* retinal is referred to as the visual cycle (Marmor and Martin 1978; Wald 1935).

The retinoid visual cycle is a complex series of protein interactions that function to regenerate usable retinal chromophores (reviewed in (Gonzalez-Fernandez 2002; Hofmann and Palczewski 2015; von Lintig et al. 2010). In addition to resetting phototransductive pathways, the retinoid visual cycle also functions to prepare phototransductive pathways for the first light of day and alleviate the toxicity of excess and free retinal isomers (Rando 1990; Lamb and Pugh 2004; Gonzalez-Fernandez 2002; Chalovich and Eisenberg 2013; Travis et al. 2007). There are three described pathways identified from vertebrates, insects, and mollusks. The vertebrate retinoid visual cycle is thoroughly described and is made up of a large network of proteins which cross between the photoreceptor cells and the retinal pigment epithelium (RPE) (Fischer et al. 1999; Liou et al. 1982). Insect vision primarily functions on bistable photopigments but recent findings suggest a retinoid visual cycle possessing retinal

dehydrogenases homologous to that of vertebrates and also including retinal isomers between the photoreceptor cells and the pigment cells of the ommatidia (T. Wang and Montell 2005; X. Wang et al. 2010; Arshavsky 2010).

The retinoid visual cycle in molluscs functions via retinochrome, a mollusc specific opsin which functions as a photoisomerase in tandem with a retinal shuttle protein, RALBP (Ozaki et al. 1983; Terakita, Hara, and Hara 1989). This classic view of the retinoid visual cycle in molluscs has not been examined in over 30 years, but the current understanding does not address known issues involved with the use of retinal isomers as a chromophore. Because of the necessity for light to reset the system, it is unclear how the visual cycle of molluscs prepare the visual opsins for the first exposure of light. It is also unclear how retinal isomers, which are metabolically expensive and inherently toxic in relatively small doses, are stored for use by the opsins when needed. In order to address these questions, the visual cycle of molluscs must be reexamined.

Opsin Photopigment Function

The phylogenetic classifications of known opsins are well understood (Porter et al. 2012; Ramirez et al. 2016; Feuda et al. 2012); however, the wavelength at which the photopigment becomes active is less complete. Spectral tuning is the modulation of the absorbance peaks or lambda max (λ_{max}) of a photopigment by alteration of the amino acid side chains which affect the chromophore binding pocket (Lin et al. 1998). Most work on the spectral tuning of opsins have focused on the identification of λ_{max} to answer ecological (Bernard and Remington 1991; Cronin, Marshall, and Caldwell 2000; Hart 2001) or evolutionary (Hart and Hunt 2007; Hunt et al. 2009; Wakakuwa et al. 2010) questions of an organism's sensory system. The retinal chromophore is the light

absorbing molecule of the photopigment, and it is hypothesized that the relationship between the chromophore and the polar environment (Shimono et al. 2001) or shape (Hirayama et al. 1994) of the binding pocket determines the λ_{max} of the photopigment. Mutations in important amino acid sites within the opsin protein, such as the Schiff-base and counterion to the Schiff-base linkage sites or amino acid residues near these sites (Shimono et al. 2000, 2001; Nathans 1990), have been shown to play a role in the spectral tuning of opsins. Yet, these sites tend to be highly conserved particularly in closely related species, making it likely that sites elsewhere in the opsin have an important role in spectral tuning as well. The functional assays to express and identify the lambda max of opsin proteins is cumbersome, particularly with the increase in discovery rate of photosensitive proteins. To counter this problem, computational modeling approaches have been used to identify the relationship between the genotype and phenotype of opsin proteins (Melaccio et al. 2011; Melaccio, Ferré, and Olivucci 2012; Ferré and Olivucci 2003; González-Luque et al. 2000). The accuracy of these models depends on developing larger datasets of mapped relationships between the secondary structure of an opsin proteins and its effect on the absorbance of the protein. Creating three dimensional models requires high-resolution diffraction data from crystallized proteins.

The first crystallographic analysis of an non-vertebrate opsin was from the retina of the Japanese flying squid, *Todarodes pacificus* (Shimamura et al. 2008). Squid rhodopsin showed a similar seven transmembrane helices structure to that of vertebrate rhodopsin; however, squid rhodopsin possesses a more structured cytoplasmic region to that of vertebrate rhodopsin. Structural differences may highlight the reason behind

functional differences where squid rhodopsin is bistable and activates a phototransduction pathway based on IP_3 signaling cascade different G-protein to that of vertebrate rhodopsin (Terakita et al. 1998). This crystal structures provide important information to model protein folding and chromophore binding pocket of a non-vertebrate opsin allowing us to identify amino acid sites that may play a role in spectral tuning of other opsin classes (Sekharan and Morokuma 2010).

Dissertation Organization

The objective of my dissertation is to investigate visual systems in molluscs. In chapter II, I evaluate the current hypotheses describing the phylogenetic relationships among taxonomic families within the superfamily Pectinoidea and present an expansion of previously published molecular data including taxa from four of five extant families. With these results, I introduce a new hypothesis of the origin of eyes outside Pectinoidea based on presence and absence of eyes in species across Pectinoidea. In chapter III, I challenge the current hypothesis of the retinoid visual cycle within molluscs and present the results of molluscan transcriptome searches using known proteins within the visual cycle of vertebrates and insects as query proteins. Based on the topology of the gene trees produced using identified transcripts, I propose a novel molluscan retinoid visual cycle model that is similar in structure to insects. Furthermore, by using phylogenetic methods, I identified new molluscan proteins that are homologous to vertebrate and insect visual cycle proteins. In chapter IV, I use two closely related scallop retinochromes to begin to examine the relationship between genotype and phenotype of opsin proteins. My results demonstrate that the presence of a hydroxyl group in proximity to the bound retinal chromophore causes a blue shift in the spectral

tuning of RTC. I also suggest the effects of conformational changes of the binding pocket by amino acids secondary to those lining the binding pocket may create a sort of hierarchy of effect in which some amino acids compensate or overpower the effect of others. Finally, I highlight the efficacy of retinochrome as a model to investigate the relationship of genotype to phenotype in opsin proteins allowing further study of this relationship to build more effective prediction models.

References

Alejandrino, Alvin, Louise Puslednik, and Jeanne M. Serb. 2011. “Convergent and Parallel Evolution in Life Habit of the Scallops (Bivalvia: Pectinidae).” *BMC Evolutionary Biology* 11 (1): 164. https://doi.org/10.1186/1471-2148-11-164.

Arshavsky, Vadim Y. 2010. “Vision: The Retinoid Cycle in Drosophila.” *Current Biology* 20 (3): R96–98. https://doi.org/10.1016/j.cub.2009.12.039.

Bernard, Gary D., and Charles L. Remington. 1991. “Color Vision in Lycaena Butterflies: Spectral Tuning of Receptor Arrays in Relation to Behavioral Ecology.” *Proceedings of the National Academy of Sciences of the United States of America* 88 (7): 2783–87. https://doi.org/10.1073/pnas.88.7.2783.

Bieler, Rudiger, Paula M. Mikkelsen, Timothy M. Collins, Emily A. Glover, Vanessa L. Gonzalez, Daniel L. Graf, Elizabeth M. Harper, et al. 2014. “Investigating the Bivalve Tree of Life - An Exemplar-Based Approach Combining Molecular and Novel Morphological Characters.” *Invertebrate Systematics* 28 (1): 32–115. https://doi.org/10.1071/IS13010.

Chalovich, Joseph M., and Evan Eisenberg. 2013. *New Insights into Retinoid Metabolism and Cycling within the Retina*. *Biophysical Chemistry*. Vol. 257. https://doi.org/10.1016/j.immuni.2010.12.017.Two-stage.

Cronin, Thomas W., N. Justin Marshall, and Roy L. Caldwell. 2000. “Spectral Tuning and the Visual Ecology of Mantis Shrimps.” *Philosophical Transactions of the Royal Society of London B* 355: 1263–67.

Dijkstra, Henk H., and Philippe Maestrati. 2012. “Pectinoidea (Mollusca, Bivalvia, Propeamussiidae, Cyclochlamydidae n. Fam., Entoliidae and Pectinidae) from the Vanuatu Archipelago.” *Zoosystema* 34 (2): 389–408. https://doi.org/10.5252/z2012n2a12.

Ferré, Nicolas, and Massimo Olivucci. 2003. “Probing the Rhodopsin Cavity with Reduced Retinal Models at the CASPT2/CASSCF/AMBER Level of Theory.” *Journal of the American Chemical Society* 125 (23): 6868–69. https://doi.org/10.1021/ja035087d.

Feuda, Roberto, Sinead C. Hamilton, James O. McInerney, and Davide Pisani. 2012. “Metazoan Opsin Evolution Reveals a Simple Route to Animal Vision.” *Proceedings of the National Academy of Sciences* 109 (46): 18868–72. https://doi.org/10.1073/pnas.1305990110.

Fischer, Andy J., Josh Wallman, James R. Mertz, and William K. Stell. 1999. “Localization of Retinoid Binding Proteins, Retinoid Receptors, and Retinaldehyde Dehydrogenase in the Chick Eye.” *Journal of Neurocytology* 28 (7): 597–609. https://doi.org/10.1023/A:1007071406746.

Gonzalez-Fernandez, Federico. 2002. “Evolution of the Visual Cycle: The Role of Retinoid-Binding Proteins.” *Journal of Endocrinology* 175: 75–88. https://doi.org/10.1677/joe.0.1750075.

González-Luque, Remedios, Marco Garavelli, Fernando Bernardi, Manuela Merchán, Michael A. Robb, and Massimo Olivucci. 2000. “Computational Evidence in Favor of a Two-State, Two-Mode Model of the Retinal Chromophore Photoisomerization.” *Proceedings of the National Academy of Sciences of the United States of America* 97 (17): 9379–84. https://doi.org/10.1073/pnas.97.17.9379.

Guderley, Helga E, and Isabelle Tremblay. 2013. “Escape Responses by Jet Propulsion in Scallops.” *Canadian Journal of Zoology* 91: 420–30.

Hara-Nishimura, Ikuko, Maki Kondo, Mikio Nishimura, Reiko Hara, and Tomiyuki Hara. 1993. “Amino Acid Sequence Surrounding the Retinal-Binding Site in Retinochrome of the Squid, Todarodes Pacificus.” *FEBS Letters* 335 (1): 94–98. https://doi.org/0014-5793(93)80447-3 [pii] ET - 1993/11/29.

Hart, Nathan S. 2001. “The Visual Ecology of Avian Photoreceptors.” *Progress in Retinal and Eye Research* 20 (5): 675–703. https://doi.org/10.1016/S1350-9462(01)00009-X.

Hart, Nathan S., and David M. Hunt. 2007. “Avian Visual Pigments: Characteristics, Spectral Tuning, and Evolution.” *American Naturalist* 169 (SUPPL.). https://doi.org/10.1086/510141.

Hayami, Itaru. 1991. “Living and Fossil Scallop Shells as Airfoils: An Experimental Study.” *Paleobiology* 17: 1–18.

Hirayama, Junichi, Yasushi Imamoto, Yoshinori Shichida, Tŏru Yoshizawa, Alfred E. Asato, Robert S.H. Liu, and Naoki Kamo. 1994. “SHAPE OF THE CHROMOPHOIW BINDING SITE IN Pharaonis PHOBORHODOPSIN FROM A STUDY USING RETINAL ANALOGS.” *Photochemistry and Photobiology* 60 (4): 388–93. https://doi.org/10.1111/j.1751-1097.1994.tb05121.x.

Hofmann, Lukas, and Krzysztof Palczewski. 2015. “Advances in Understanding the Molecular Basis of the First Steps in Color Vision.” *Prog Retin Eye Res*, 46–66. https://doi.org/10.1016/j.preteyeres.2015.07.004.Advances.

Hubbard, Ruth, and Robert C.C. St.George. 1958. “The Rhodopsin System of the Squid.” *The Journal of General Physiology* 41 (3): 501–28. https://doi.org/10.1085/jgp.41.3.501.

Hunt, David M., Livia S. Carvalho, Jill A. Cowing, and Wayne L. Davies. 2009. “Evolution and Spectral Tuning of Visual Pigments in Birds and Mammals.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1531): 2941–55. https://doi.org/10.1098/rstb.2009.0044.

Lamb, Trevor D., Shaun P. Collin, and Edward N. Pugh. 2007. “Evolution of the Vertebrate Eye: Opsins, Photoreceptors, Retina and Eye Cup.” *Nature* 8 (12): 960–76. https://doi.org/10.1038/jid.2014.371.

Lamb, Trevor D., and Edward N. Pugh. 2004. “Dark Adaptation and the Retinoid Cycle of Vision.” *Progress in Retinal and Eye Research* 23 (3): 307–80. https://doi.org/10.1016/j.preteyeres.2004.03.001.

Lamb, Trevor D., Edward N. Pugh, and Shaun P. Collin. 2008. “The Origin of the Vertebrate Eye.” *Evolution: Education and Outreach* 1 (4): 415–26. https://doi.org/10.1007/s12052-008-0091-2.

Land, Michael F. 1965. “Image Formation by a Concave Reflector in the Eye of the Scallop, Pecten Maximus.” *Journal of Physiology* 179: 138–53.

Land, Michael F, and Russell D. Fernald. 1992. “The Evolution of Eyes.” *Annual Review of Neuroscience* 15 (1990): 1–29. https://doi.org/10.1146/annurev.ne.15.030192.000245.

Lin, Steven W., Gerd G. Kochendoerfer, Kate S. Carroll, Dorothy Wang, Richard A. Mathies, and Thomas P. Sakmar. 1998. “Mechanisms of Spectral Tuning in Blue Cone Visual Pigments: Visible and Raman Spectroscopy of Blue-Shifted Rhodopsin Mutants.” *Journal of Biological Chemistry* 273 (38): 24583–91. https://doi.org/10.1074/jbc.273.38.24583.

Lintig, Johannes von, Philip D. Kiser, Marcin Golczak, and Krzysztof Palczewski. 2010. “The Biochemical and Structural Basis for Trans-to-Cis Isomerization of Retinoids in the Chemistry of Vision.” *Trends in Biochemical Sciences* 35 (7): 400–410. https://doi.org/10.1016/j.tibs.2010.01.005.

Liou, Gregory I., C. David B. Bridges, Shaoling Fong, Richard A. Alvarez, and Federico Gonzalez-Fernandez. 1982. “Vitamin a Transport between Retina and Pigment Epithelium-an Interstitial Protein Carrying Endogenous Retinol (Interstitial Retinol-Binding Protein).” *Vision Research* 22 (12): 1457–67. https://doi.org/10.1016/0042-6989(82)90210-3.

Marmor, Michael F., and Lizbeth J. Martin. 1978. “100 Years of the Visual Cycle.” *Survey of Ophthalmology* 22 (4): 279–85. https://doi.org/10.1016/0039-6257(78)90074-7.

Matsumoto, Masahiro, and Itaru Hayami. 2000. “Phylogenetic Analysis of the Family Pectinidae (Bivalvia) Based on Mitochondrial Cytochrome C Oxidase Subunit I.” *Journal Molluscan Studies* 66 (4): 477–88. https://doi.org/10.1093/mollus/66.4.477.

Melaccio, Federico, Nicolas Ferré, and Massimo Olivucci. 2012. “Quantum Chemical Modeling of Rhodopsin Mutants Displaying Switchable Colors.” *Physical Chemistry Chemical Physics* 14 (36): 12485–95. https://doi.org/10.1039/c2cp40940b.

Melaccio, Federico, Massimo Olivucci, Roland Lindh, and Nicolas Ferre. 2011. “Unique QM/MM Potential Energy Surface Exploration Using Microiterations.” *International Journal of Quantum Chemistry* 111: 4020–29. https://doi.org/10.1002/qua.

Millward, A., and M.A. Whyte. 1992. “The Hydrodynamic Characteristics of Six Scallops of the Superfamily Pectinacea, Class Bivalvia. Journal of Zoology (London) 227:547-566.,” no. 1 992: 547–66.

Morton, Brian. 2001. “The Evolution of Eyes in the Bivalvia.” *Oceanography and Marine Biology, an Annual Review* 39: 165–205.

———. 2008. “The Evolution of Eyes in the Bivalvia: New Insights*.” *American Malacological Bulletin* 26 (1–2): 35–45. https://doi.org/10.4003/006.026.0205.

Moseley, H.N. 1885. “On the Presence of Eyes in the Shells of Certain Chitonidæ, and on the Structure of These Organs.” *Journal of Cell Science* s2-25 (97): 37–60.

Nathans, Jeremy. 1990. “Determinants of Visual Pigment Absorbance: Identification of the Retinylidene Schiff’s Base Counterion in Bovine Rhodopsin.” *Biochemistry* 29 (41): 9746–52. https://doi.org/10.1021/bi00493a034.

Nilsson, Dan-E. 1994. “Eyes as Optical Alarm Systems in Fan Worms and Ark Clams.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 346 (1316): 195–212. https://doi.org/10.1098/rstb.1994.0141.

———. 2009. “The Evolution of Eyes and Visually Guided Behaviour.” *Philosophical Transactions of the Royal Society B-Biological Sciences* 364: 2833–47.

———. 2013. “Eye Evolution and Its Functional Basis.” *Visual Neuroscience* 30 (1–2): 5–20. https://doi.org/10.1017/S0952523813000035.

Oakley, Todd H., and Daniel I. Speiser. 2015. “How Complexity Originates: The Evolution of Animal Eyes.” *Annual Review of Ecology, Evolution, and Systematics* 46 (1): 237–60. https://doi.org/10.1146/annurev-ecolsys-110512-135907.

Ozaki, Koichi, Reiko Hara, Tomiyuki Hara, and Toshiaki Kakitani. 1983. “Squid Retinochrome. Configurational Changes of the Retinal Chromophore.” *Biophysical Journal* 44 (1): 127–37. https://doi.org/10.1016/S0006-3495(83)84285-4.

Piatigorsky, Joram. 2008. “Evolution of Mollusc Lens Crystalline: Glutathione S-Transferase/S- Crystallins and Aldehyde Dehydrogenase/Î©-Crystallins.” *American Malacological Bulletin* 26 (1–2): 73–81. http://www.scopus.com/inward/record.url?eid=2-s2.0-58149393227&partnerID=40.

Ponder, W.F., and D.R. Lindberg. 2008. *Phylogeny and Evolution of the Mollusca*. Berkeley: University of California Press.

Porter, Megan L., Joseph R. Blasic, Michael J. Bok, Evan G. Cameron, Thomas Pringle, Thomas W. Cronin, and Phyllis R. Robinson. 2012. “Shedding New Light on Opsin Evolution.” *Proc Biol Sci* 279 (1726): 3–14. https://doi.org/10.1098/rspb.2011.1819.

Ramirez, M. Desmond, Autum N. Pairett, M. Sabrina Pankey, Jeanne M. Serb, Daniel I. Speiser, Andrew J. Swafford, and Todd H. Oakley. 2016. “The Last Common Ancestor of Most Bilaterian Animals Possessed at Least Nine Opsins.” *Genome Biology and Evolution* 8 (12): 3640–52. https://doi.org/10.1093/gbe/evw248.

Rando, Robert R. 1990. “The Chemistry of Vitamin A and Vision.” *Angewandte Chemie-International Edition in English* 29 (5): 461–80. https://doi.org/10.1002/anie.199004611.

Salvini-Plawen, L. von, and Ernst Mayr. 1977. “On the Evolution of Photoreceptors and Eyes.” *Journal of Evolutionary Biology* Vol. 10: 207–63. https://doi.org/10.1007/978-1-4615-6953-4_4.

Sekharan, Sivakumar, and Keiji Morokuma. 2010. “Drawing the Retinal out of Its Comfort Zone: An ONIOM(QM/MM) Study of Mutant Squid Rhodopsin.” *Journal of Physical Chemistry Letters* 1 (3): 668–72. https://doi.org/10.1021/jz100026k.

Serb, Jeanne M., Alvin Alejandrino, Erik Otárola-Castillo, and Dean C. Adams. 2011. “Morphological Convergence of Shell Shape in Distantly Related Scallop Species (Mollusca: Pectinidae).” *Zoological Journal of the Linnean Society* 163 (2): 571–84. https://doi.org/10.1111/j.1096-3642.2011.00707.x.

Serb, Jeanne M., and Douglas J. Eernisse. 2008. “Charting Evolution’s Trajectory: Using Molluscan Eye Diversity to Understand Parallel and Convergent Evolution.” *Evolution: Education and Outreach* 1 (4): 439–47. https://doi.org/10.1007/s12052-008-0084-1.

Serb, Jeanne M., Emma Sherratt, Alvin Alejandrino, and Dean C. Adams. 2017. “Phylogenetic Convergence and Multiple Shell Shape Optima for Gliding Scallops (Bivalvia: Pectinidae).” *Journal of Evolutionary Biology* 30 (9): 1736–47. https://doi.org/10.1111/jeb.13137.

Sherratt, Emma, Alvin Alejandrino, Andrew C. Kraemer, Jeanne M. Serb, and Dean C. Adams. 2016. “Trends in the Sand: Directional Evolution in the Shell Shape of Recessing Scallops (Bivalvia: Pectinidae).” *Evolution* Early View: doi:10.1111/evo.12995. https://doi.org/10.1111/evo.12995.

Shichida, Yoshinori, and Take Matsuyama. 2009. “Evolution of Opsins and Phototransduction.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 364 (1531): 2881–95. https://doi.org/10.1098/rstb.2009.0051.

Shimamura, Tatsuro, Kenji Hiraki, Naoko Takahashi, Tetsuya Hori, Hideo Ago, Katsuyoshi Masuda, Koji Takio, Masaji Ishiguro, and Masashi Miyano. 2008. “Crystal Structure of Squid Rhodopsin with Intracellularly Extended Cytoplasmic Region.” *Journal of Biological Chemistry* 283 (26): 17753–56. https://doi.org/10.1074/jbc.C800040200.

Shimono, Kazumi, Yukako Ikeura, Yuki Sudo, Masayuki Iwamoto, and Naoki Kamo. 2001. “Environment around the Chromophore in Pharaonis Phoborhodopsin: Mutation Analysis of the Retinal Binding Site.” *Biochimica et Biophysica Acta - Biomembranes* 1515 (2): 92–100. https://doi.org/10.1016/S0005-2736(01)00394-7.

Shimono, Kazumi, Masayuki Iwamoto, Masato Sumi, and Naoki Kamo. 2000. “Effects of Three Characteristic Amino Acid Residues of Pharaonis Phoborhodopsin on the Absorption Maximum¶.” *Photochemistry and Photobiology* 72 (1): 141. https://doi.org/10.1562/0031-8655(2000)072<0141:eotcaa>2.0.co;2.

Speiser, Daniel I., and Sönke Johnsen. 2008. “Comparative Morphology of the Concave Mirror Eyes of Scallops (Pectinoidea)*.” *American Malacological Bulletin* 26 (July 2007): 27–33. https://doi.org/10.4003/006.026.0204.

Speiser, Daniel I., M. Sabrina Pankey, Alexander K. Zaharoff, Barbara A. Battelle, Heather D. Bracken-Grissom, Jesse W. Breinholt, Seth M. Bybee, et al. 2014. “Using Phylogenetically-Informed Annotation (PIA) to Search for Light-Interacting Genes in Transcriptomes from Non-Model Organisms.” *BMC Bioinformatics* 15 (1): 350. https://doi.org/10.1186/s12859-014-0350-x.

Stanley, Steven M. 1970. “Relation of Shell Form to Life Habits of the Bivalvia (Mollusca).” *Geological Society of America, Memoir* 125: 1–296.

Steiner, Gerhard, and Sabine Hammer. 2000. “Molecular Phylogeny of the Bivalvia Inferred from 18S RDNA Sequences with Particular Reference to the Pteriomorphia.” *Geological Society, London, Special Publications* 177 (1999): 11–29. https://doi.org/10.1144/GSL.SP.2000.177.01.02.

Stöger, I., J.D. Sigwart, Y. Kano, T. Knebelsberger, B.A. Marshall, E. Schwabe, and M. Schrödl. 2013. “The Continuing Debate on Deep Molluscan Phylogeny: Evidence for Serialia (Mollusca, Monoplacophora + Polyplacophora).” *BioMed Research International* 2013 (December): 10–13. https://doi.org/10.1155/2013/407072.

Terakita, Akihisa. 2005. “The Opsins.” *Genome Biology* 6 (3): 213. https://doi.org/10.1186/gb-2005-6-3-213.

Terakita, Akihisa, Reiko Hara, and Tomiyuki Hara. 1989. “Retinal-Binding Protein as a Shuttle for Retinal in the Rhodopsin-Retinochrome System of the Squid Visual Cells.” *Vision Research* 29 (6): 639–52. https://doi.org/10.1016/0042-6989(89)90026-6.

Terakita, Akihisa, Takahiro Yamashita, and Yoshinori Shichida. 2000. “Highly Conserved Glutamic Acid in the Extracellular IV-V Loop in Rhodopsins Acts as the Counterion in Retinochrome, a Member of the Rhodopsin Family.” *Proceedings of the National Academy of Sciences of the United States of America* 97 (26): 14263–67. https://doi.org/10.1073/pnas.260349597.

Terakita, Akihisa, Takahiro Yamashita, Shuji Tachibanaki, and Yoshinori Shichida. 1998. “Selective Activation of G-Protein Subtypes by Vertebrate and Invertebrate Rhodopsins.” *FEBS Letters* 439 (1–2): 110–14. https://doi.org/10.1016/S0014-5793(98)01340-4.

Travis, Gabriel H., Marcin Golczak, Alexander R. Moise, and Krzysztof Palczewski. 2007. “Diseases Caused by Defects in the Visual Cycle: Retinoids as Potential Therapeutic Agents.” *Annual Review of Pharmacology and Toxicology* 47 (1): 469–512. https://doi.org/10.1146/annurev.pharmtox.47.120505.105225.

Tremblay, Isabelle, Myriam Samson-Dô, and Helga E. Guderley. 2015. “When Behavior and Mechanics Meet: Scallop Swimming Capacities and Their Hinge Ligament.” *Journal of Shellfish Research* 34 (2): 203–12. https://doi.org/10.2983/035.034.0201.

Tsukamoto, Hisao, and Akihisa Terakita. 2010. “Diversity and Functional Propertis of Bistable Pigments.” *Photochem. Photobiol. Sci.* 9 (11): 1418. https://doi.org/10.1039/c0pp00126k.

Vopalensky, Pavel, and Zbynek Kozmik. 2009. “Eye Evolution: Common Use and Independent Recruitment of Genetic Components.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1531): 2819–32. https://doi.org/10.1098/rstb.2009.0079.

Wakakuwa, Motohiro, Akihisa Terakita, Mitsumasa Koyanagi, Doekele G. Stavenga, Yoshinori Shichida, and Kentaro Arikawa. 2010. “Evolution and Mechanism of Spectral Tuning of Blue-Absorbing Visual Pigments in Butterflies.” *PLoS ONE* 5 (11): 1–8. https://doi.org/10.1371/journal.pone.0015015.

Wald, George. 1935. “Carotenoids and the Visual Cycle.”

Waller, Thomas R. 1980. “Scanning Electron Microscopy of Shell and Mantle in the Order Arcoida (Mollusca: Bivalvia).” *Smithsonian Contributions to Zoology* 313: 1–58.

———. 1991. “Evolutionary Relationships among Commercial Scallops (Mollusca: Bivalvia: Pectinidae).” In *Scallops: Biology, Ecology and Aquaculture*, edited by S E Shumway, 1–73. New York: Elsevier.

———. 2006. “Phylogeny of Families in the Pectinoidea (Mollusca: Bivalvia): Importance of the Fossil Record.” *Zoological Journal of the Linnean Society* 148 (3): 313–42. https://doi.org/10.1111/j.1096-3642.2006.00258.x.

Wang, Tao, and Craig Montell. 2005. “Rhodopsin Formation in Drosophila Is Dependent on the PINTA Retinoid-Binding Protein.” *Journal of Neuroscience* 25 (21): 5187–94. https://doi.org/10.1523/JNEUROSCI.0995-05.2005.

Wang, Xiaoyue, Tao Wang, Yuchen Jiao, Johannes von Lintig, and Craig Montell. 2010. “Requirement for an Enzymatic Visual Cycle in Drosophila.” *Current Biology* 20 (2): 93–102. https://doi.org/10.1016/j.cub.2009.12.022.

Zieger, Marina V., and Victor Benno Meyer-Rochow. 2008. “Understanding the Cephalic Eyes of Pulmonate Gastropods: A Review*.” *American Malacological Bulletin* 26 (1–2): 47–66. https://doi.org/10.4003/006.026.0206.

CHAPTER 2. MOLECULAR PHYLOGENY OF THE PECTINOIDEA (BIVALVIA) INDICATES PROPEAMUSSIIDAE TO BE A NON-MONOPHYLETIC FAMILY WITH ONE CLADE SISTER TO THE SCALLOPS (PECTINIDAE).

Authors: G. Dalton Smedley[a], Jorge A. Audino[b], Courtney Grula[a1], Anita Porath-Krause[a2], Autum N. Pairett[a], Alvin Alejandrino[a3], Latayshia Lacey[a], Felicity Masters[c], Peter F. Duncan[c], Ellen E. Strong[d], Jeanne M. Serb[a*]

[a] *Department of Evolution, Ecology, and Organismal Biology, Iowa State University, 2200 Osborn Dr, 251 Bessey Hall, Ames, IA 50011, USA*

[b] *Department of Zoology, University of São Paulo, Rua do Matão, Travessa 14, n. 101, 05508-090 São Paulo, SP, Brazil*

[c] *Faculty of Science, Health, Education, and Engineering, University of the Sunshine Coast, Maroochydore DC, Queensland 4558, Australia*

[d] *Department of Invertebrate Zoology, National Museum of National History, Smithsonian Institution, 10th and Constitution Ave NW, Washington DC, 20560, USA*

Modified from a manuscript published in *Molecular Phylogenetics and Evolution*

Abstract

Scallops (Pectinidae) are one of the most diverse families of bivalves and have been a model system in evolutionary biology. However, in order to understand phenotypic evolution, the Pectinidae needs to be placed in a deeper phylogenetic framework within the superfamily Pectinoidea. We reconstructed a molecular phylogeny for 60 species from four of the five extant families within the Pectinoidea using a five gene dataset

(12S, 16S, 18S, 28S rRNAs and histone H3). Our analyses give consistent support for the non-monophyly of the Propeamussiidae, with a subset of species as the sister group to the Pectinidae, the Propeamussiidae type species as sister to the Spondylidae, and the majority of propeamussiid taxa sister to the Spondylidae + *Pr. dalli*. This topology represents a previously undescribed relationship of pectinoidean families. Our results suggest a single origin for eyes within the superfamily and likely multiple instances of loss for these characters. However, it is now evident that reconstructing the evolutionary relationships of Pectinoidea will require a more comprehensive taxonomic sampling of the Propeamussiidae *sensu lato.*

Introduction

Scallops Pectinidae Rafinesque, 1815 are one of the most ecologically and morphologically diverse families in the class Bivalvia. With over 250 extant species currently considered valid, they are distributed across polar, temperate, and tropical marine ecosystems of shallow sublittoral reefs, sandy bays, sea grass beds and coarse substrates of the continental shelves, with a smaller number of species restricted to deeper water (Serb, 2016). Pectinidae is an ideal model to study the evolution of complex traits due to the number and biological diversity of extant species, the link between shell morphology and habitat use (Stanley, 1970), and their high preservability in the paleontological record (Valentine et al., 2006). Researchers have investigated the evolution of traits such as shell shape (Serb et al., 2011, 2017; Sherratt et al., 2016; Stanley, 1970), behavior (Alejandrino et al., 2011), swimming mechanics (Guderley and Tremblay, 2013; Hayami, 1991; Millward and Whyte, 1992; Tremblay et al., 2015), and phototransduction (Faggionato and Serb, 2017; Gomez et al., 2011; Kingston et al.,

2017; Porath-Krause et al., 2016; Serb et al., 2013). One compelling set of phenotypes is the complex sensory systems, including eyes, found in this superfamily (Audino et al., 2015a, 2015b, 2015c; Land, 1965; Speiser et al., 2011, 2016; Speiser and Johnsen, 2008). Most work has concentrated on the eyes of scallops, which were first described in 1791 (Poli, 1791). Subsequent research focused on the anatomy and optics of these eyes to understand how the eyes capture light and focus images (Land, 1965; Palmer et al., 2017; Speiser et al., 2016; Speiser and Wilkens, 2016). Recent molecular approaches have provided insights into the evolution of gene families involved in scallop photoreception (Gomez et al., 2011; Kojima et al., 1997; Pairett and Serb, 2013; Piatigorsky et al., 2000; Porath-Krause et al., 2016; Serb et al., 2013). However, in order to understand the origin and evolution of these and other traits, the family Pectinidae needs to be placed in a deeper phylogenetic framework within the superfamily Pectinoidea.

The relationship of the Pectinidae to the other families in the Pectinoidea has been highly contentious due to high levels of homoplasy in shell characters (Dijkstra and Maestrati, 2012; Hertlein, 1969) and alternative interpretations of the fossil record (Waller, 2006, 1991, 1978) (Figure 2-1). As a result, three families (Propeamussiidae, Spondylidae, Entoliidae) singly or in combination have been proposed to be the sister taxon to the Pectinidae by different authors at different times. The prevailing view has been that the Propeamussiidae Abbott, 1954, or glass scallops (~200 species), represent the closest relatives of the Pectinidae. Propeamussiids possess very thin, often translucent shells and inhabit the marine epipelagic (80 m) to the abyssal (4000 m) zones. They appear to be a lineage of relict species that survived severe

environmental changes at the end of the Cretaceous by inhabiting deep and/or cold-water refugia (Waller, 1991) where most modern propeamussiids and the oldest extant lineage of Pectinidae (Camptonectinae: *Delectopecten*) are still found. Additionally, propeamussiids and some pectinid lineages have a similar shell shape. These data suggest a possible sister relationship between the two families, which has been supported by other studies which include molecular data for their phylogenetic analyses (Bieler et al., 2014, Figure 30; Matsumoto and Hayami, 2000) (Figure 2-1A). Recently, one lineage of micro glass scallops (1.5 - 6 mm as adults) was elevated to its own family, the Cyclochlamydidae Dijkstra and Maestrati, 2012; however its phylogenetic relationship to the Pectinidae is unknown. There remains two other pectinoidean families: the Entoliidae Teppner, 1922, a mostly extinct family with only two extant monotypic genera (*Entolium*, *Pectinella*) (Waller, 2006), and the Spondylidae Gray, 1826 or thorny oysters (68 species), a cementing family with finger-like protrusions on the shell. These less-studied families have been hypothesized to be the sister group to the Pectinidae, either separately (Waller, 2006, 1991) (Figures 2-1B vs 2-1C) or together as the sister clade (Waller, 1978) (Figure 2-1D). New data on the age of first known fossil occurrences in conjunction with morphological characteristics are the basis of a revised phylogenetic hypothesis supporting the Spondylidae as the sister group to the Pectinidae, with Entoliidae + Propeamussiidae forming a second clade (Waller, 2006) (Figure 2-1B). A molecular phylogeny based on a single mitochondrial gene also supports the Pectinidae + Spondylidae relationship, but differs in that the Propeamussiidae was recovered as the sister group to Pectinidae + Spondylidae (Matsumoto, 2003). However, that study did not include the Entoliidae. In contrast,

three- and four-gene datasets of mitochondrial and nuclear markers have recovered Spondylidae + Propeamussiidae as the sister group to the Pectinidae (Alejandrino et al., 2011; Puslednik and Serb, 2008). To date, no molecular phylogenetic analysis has included more than three pectinoidean families (Plazzi et al., 2011; Plazzi and Passamonti, 2010; Sharma et al., 2012; Sun and Gao, 2017), which has prevented more definitive resolution.

We generated a 18S rDNA dataset for 60 pectinoidean species and five species of Limidae to complement an existing multigene dataset (Sherratt et al., 2016) and broadened the taxonomic representation to include four of the five extant families of Pectinoidea. We then calibrated the multi-locus phylogenetic hypothesis using fossil data from three families. Using this framework, our goal was to clarify the phylogenetic relationships between Pectinidae and other families within Pectinoidea.

Materials and methods

Specimens and Samples

We assembled 60 taxa from four of the five extant families in the superfamily Pectinoidea plus five species of Limidae to serve as the outgroup. We sampled 18 species from the Propeamussiidae, 37 species of Pectinidae, four species of Spondylidae, and a single extant species of Entoliidae (Supplementary Table 2-S1). Due to the challenges of acquiring samples, we were unable to include taxa from the newly described family Cyclochlamydidae. Samples used in this study were obtained from colleagues and museum collections (see Supplementary Table 2-S1 and Acknowledgments). The majority of Indo-Pacific specimens included in this study were

obtained during expeditions organized by the MNHN and Pro-Natura International as part of the *Our Planet Reviewed* program, and by the MNHN and the Institut de Recherche pour le Développement as part of the *Tropical Deep-Sea Benthos* program. Species identifications of the Indo-Pacific specimens were determined by Henk H. Dijkstra at the Naturalis Biodiversity Center (Netherlands). All tissues were preserved in ethanol and shell voucher specimens are available from museum collections listed in Supplementary Table 2-S1.

Molecular Laboratory Methods

Total genomic DNA (gDNA) was extracted from either mantle or adductor tissues following the manufacturer's protocol of the Qiagen DNeasy Blood and Tissue kit. A portion of the nuclear gene 18S ribosomal RNA (~700bp) was amplified using the 18S a2.0 forward (5'-ATGGTTGCAAAGCTGAAAC-3') and 18S 9R reverse (5'-GATCCTTCCGCAGGTTCACCTAC-3') primers (Giribet et al., 1996; Whiting et al., 1997). PCR reactions were carried out in 25μL total volume reactions containing 12.5μl 2x MyTaq Red Mix (Bioline), 1μl of 10μM 18S rRNA forward and reverse primers (18s a2.0 and 18s 9R, respectively), 9.5μl double distilled water, and 1μL of template. Reactions underwent one round of PCR consisting of an initial denaturation step (2 minutes at 95^{o} C) followed by 30 cycles of chain denaturation (15 seconds at 95^{o} C), primer annealing (15 seconds at 50^{o} C), and elongation (10 to 60 seconds at 72^{o} C). Roughly 5μl of the amplification products were visualized on a 2% agarose gel using a 1 kb size standard. Samples with the expected band size (~700bp) were sent to Iowa State University DNA Facility for Sanger sequencing using Applied Biosystems 3730xl. In total, 18S rRNA sequences for 60 taxa (16 Propeamussiidae species, 35 Pectinidae

species, three Spondylidae species, one Entoliidae species, and five Limidae species) were successfully generated.

The 18S rRNA sequences were added to a multigene dataset consisting of two mitochondrial genes (12S and 16S rRNAs) and two nuclear genes (28S rRNA and histone H3) from previously published work from our lab (Alejandrino et al., 2011) (Supplementary Table 2-S1). To complete the dataset, we generated DNA sequences for 28S rDNA and histone H3 of *Pectinella aequoris* (Entoliidae) using the methods in Alejandrino *et al.* (2011).

Phylogenetic Analyses

DNA sequences for each gene portion were aligned separately in MAFFT v7.222 (Katoh and Standley, 2013) using the automatic algorithm to select the best alignment method and remaining settings/options set as default. Ambiguously aligned nucleotides due to large insertion-deletions (indels) in 12S, 16S, and 28S rRNA genes were removed using settings for a less stringent selection on the Gblock 0.91b server (Castresana, 2000; Dereeper et al., 2008; Talavera and Castresana, 2007). Individual gene alignments were concatenated in Geneious v4.7.6 (Kearse et al., 2012) to produce a final dataset of five gene regions: 12S rRNA (1-315 bp), 16S rRNA (316-674 bp), 18S rRNA (675-1161 bp), 28S rRNA (1162-1937 bp), and histone H3 (1938-2276 bp). Mitochondrial-only (12S and 16S rRNAs) and nuclear-only (18S rRNA, 28S rRNA and histone H3) datasets were also produced.

Phylogenetic analyses were carried out under maximum likelihood (ML: (Felsenstein, 1981)) and Bayesian inference (BI: (Mau et al., 1999)). Nucleotide substitution model was determined using PartitionFinder2 (Lanfear et al., 2016). For this

analysis, the datablock was defined by gene, as above, with branch lengths unlinked. All evolution models and schemes were investigated using Akaike Information Criterion with sample size correction (AICc) metric. ML analyses were conducted using RAxML-HPC v8.2.9 on XSEDE (Stamatakis, 2014) as implemented on the CIPRES Scientific Gateway v3.3 (Miller et al., 2010). Branch support was determined with 500 bootstrap iterations for best-scoring ML tree. All other parameters were set at the program's default. BI analyses were conducted using MrBayes v3.2.6 (Ronquist et al., 2012) as implemented on the CIPRES Scientific Gateway v3.3. We ran three independent analyses, each with eight Markov chain Monte Carlo (MCMC) chains sampling every 100 generations and the temperature for heated chains set at 0.15. The MCMC analysis was set to run for 50 million generations or until a standard deviation of split frequency value of 0.01 was reached signifying convergence following the stop rule after 4.2 million generations. The post-run analyses were set with a 50% burn-in and all other parameters not mentioned above were left at the program's default. We then visually inspected the combined trace files to confirm acceptable mixing and high ESS (effective sampling size) across all parameters (>300) in Tracer v1.6 (Rambaut et al., 2014). Post-burn-in trees were used to construct the 50% majority rule consensus tree and to estimate posterior probabilities.

We used the Approximately Unbiased (AU) test (Shimodaira, 2002) to compare our results to six alternative phylogenetic hypotheses. These alternative topologies were generated via ML in RAxML to constrain either 1) a monophyletic Propeamussiidae or 2) a clade of Propeamussiidae that excluded *Parvamussium ina*. In addition, four hypotheses from previous studies (Figure 2-1) were compared. Site-wise likelihoods

were calculated in RAxML for the unconstrained and constrained ML topologies and analyzed in CONSEL (Shimodaira and Hasegawa, 2001) using default parameters for *p*-values.

Divergence time estimation was conducted using RevBayes version 1.0.9 under the Fossilized Birth-Death model (Hohna, et al. 2016). A relaxed molecular clock model was defined assuming an uncorrelated exponential model on branch rates. Posterior probabilities were sampled by Markov Chain Monte Carlo process (MCMC) for 500,000 iterations. Maximum clade credibility tree, with a burn-in of 10%, was generated after pruning the five fossil taxa used to calibrate internal nodes. Fossil ages were incorporated based on available data in Waller (2006) and in the Paleobiology Database (https://paleobiodb.org/). Priors for fossil ages were drawn from uniform distributions and the root (Pectinoidea + Limidae) was constrained between 485.4–419.2 MYA (million year ago). The age of Pectinidae was constrained around 251.3–247.2 MYA based on the fossil of *Praechlamys spp*., also considering the fossil record of *Argopecten spp.* (15.99–2.61 MYA), an extant genera. The Spondylidae was constrained around 171.6–168.3 while Entoliidae was calibrated based on the fossil of *Pectinella spp*. (251.3–247.2 MYA). Finally, the Limidae was also constrained between 330.9–323.2 MYA, based on *Paleolima spp*.

Results

A total of 111 sequences were generated in this study and 196 sequences were obtained from previous work (Sherratt et al., 2016) for 60 species across four families of Pectinoidea with five species of Limidae serving as the outgroup. The lengths of each gene region after alignment were: 12S rRNA: 315 bp; 16S rRNA: 359 bp; 18S rRNA:

487 bp; 28S rRNA: 776 bp; histone H3: 339 bp. DNA sequences were deposited in GenBank (NCBI accession numbers MH MH463998- MH464109; Table S1). Our concatenated five-gene dataset had a total aligned length of 2276 bp. The molecular dataset was complete for 54 of the 65 taxa, while the remaining 11 taxa lacked at least one gene. Incomplete gene sets occurred in some species from all four families of Pectinoidea, but there was no pattern based on taxonomic membership (Supplementary Table 2-1). PartitionFinder 2 suggested a four partition scheme. A GTR+G evolution model was suggested for 12S and 16S partitions and a GTR+I+G evolution model for 28S and 18S+H3 partitions. However, after 200 million generations, the MrBayes analyses still had not reached convergence suggesting the PartitionFinder scheme too complicated given the dataset, requiring us to use a less complicated substitution model. A general time reversible (GTR) model with gamma-distributed rates across nucleotide sites was applied to ML and both BI analyses using the gene partitions described above.

ML and BI analyses of the concatenated five gene dataset reconstructed the same five lineages of pectinoidean taxa and produced similar topologies (Figure 2-2 for ML; Supplementary Figure 2-S1 for BI phylogram). The only difference between the two topologies was that the Bayesian analysis was unable to resolve the relationships among the five pectinoidean clades. Interestingly, the relationships among these clades in the ML topology did not match any of the proposed phylogenetic hypotheses for Pectinoidea (Figure 2-1). The single representative of Entoliidae (*Pectinella aequoris*) was recovered as sister to the remaining pectinoideans in the ML tree with high support (100% BS). The Propeamussiidae is not monophyletic, with the majority of the species

(n = 13) forming a clade with low support (64% BS, 78 PP). The type species, *Propeamussium dalli* was not a member of this clade, but rather the sister group (55% BS, 72 PP) to a well-supported monophyletic Spondylidae (100% BS; 100 PP). A third propeamussiid lineage of three species was a moderately supported clade that was the sister group to the Pectinidae (64% BS, 86 PP), and a fourth was represented by *Parvamussium ina* nested within the Pectinidae (66% BS, 89 PP). Thus, the Pectinidae as currently conceived is paraphyletic in our analyses, and the Propeamussiidae polyphyletic. Non-monophyly of the Propeamussiidae was also supported in ML and BI analyses of the mitochondrial-only and nuclear-only datasets (Supplementary Figure 2-S2-S5).

Using the best tree from each ML analysis, AU tests were performed to statistically compare our results against competing hypotheses that constrain the Propeamussiidae as monophyletic and that constrain the Propeamussiidae as monophyletic to the exclusion of *Parvamussium ina*. Additionally, we compared our results with four alternative sister groups for the Pectinidae described in previous studies (Figure 2-1). The AU test significantly rejected (p-values < 0.01) the hypotheses with a monophyletic Propeamussiidae + Pectinidae (Figure 2-1A), Spondylidae + Pectinidae (Figure 2-1B), and Entoliidae + Pectinidae (Figure 2-1C) (Table 2-1).

We estimated divergence dates among extant taxa using five fossil calibration points with horizontal bars representing the 95% highest posterior density (HPD) intervals for each node (Figure 2-3). The resulting topology recovered similar relationships to topologies derived from the RAxML (Figure 2-2) and MrBayes analyses (Supplementary Fig 2-S1) with two exceptions. First, the time-calibrated phylogeny

recovered the Entoliidae taxon as sister to largest Propeamussiidae clade with an inferred divergence time of approximately 300.3 MYA (Late Carboniferous). Second, the Pectinidae is monophyletic in the time-calibrated phylogeny as *Parvamussium ina* was the sister lineage to the family. Divergence times place the origin of the Pectinidae in the Permian (284.25 MYA) and the Spondylidae in the Late Jurassic (154.33 MYA). The divergence time of the superfamily Pectinoidea was estimated to be approximately 395 MYA (Devonian).

Discussion

Four different sister group relationships to the scallops have been hypothesized based on morphological evidence spanning the paleo- and neontological record or from molecular data. A traditional interpretation of shell similarity between the Propeamussiidae and some scallop taxa led some to conclude a sister group relationship between the two families (Figure 2-1A). However, morphological comparison of fossil and Recent taxa and re-interpretation of first occurances in the fossil record have been the basis of three other possible topologies. Waller (1978) proposed Spondylidae + Entoliidae (= Syncyclonemidae) to be the sister taxon of the Pectinidae based on a single synapomorphy of lip morphology, but noted that these taxa have many primitive features and resemble the fossil precursors to the Pectinidae more than the extant members (Figure 2-1D). Subsequently, Waller (1991) presented a revised hypothesis with the Entoliidae alone as the sister to the Pectinidae (Figure 2-1C). Most recently, fossil evidence from the Mesozoic appears to bridge morphological gaps among pectinoidean lineages (Waller, 2006). This and recognition of a “pectiniform” in an early stage of spondyliid growth led Waller (2006) to propose the

Spondylidae as the sister lineage to the Pectinidae (Figure 2-1B). Two of these hypotheses (Figure 2-1C, 2-1D) place the Propeamussiidae as sister to all other Pectinoidea. Interestingly, molecular phylogenetics has largely supported a fifth relationship, with the Propeamussiidae + Spondylidae as the sister group to the Pectinidae. Our estimated phylogenies show both a propeamussiid clade sister to the Pectinidae as well as a second propeamussiid lineage that shares a common ancestor with a monophyletic Spondylidae (Figures 2-2 and 2-3). Thus, our data support the traditional hypothesis, in part, but highlights two important future directions. First, the non-monophyly of the Propeamussiidae suggests that the characteristics that have been used as synapomorphies for the family should be re-examined. Second, if the relationship between Spondylidae and *Propeamussium dalli* (the type species) holds, a taxonomic revision of the Propeamussiidae will be necessary.

Few published time-calibrated phylogenies have included the Pectinoidea, and those that do have been inferred from a small subset of pectinoidean taxa (e.g., Bieler et al. 2014). In constrast, our estimation of divergence times for the Pectinoidea is based on a larger taxonomic sampling that includes four of the five families and fossil taxa from three of these families (Entoliidae, Pectinidae, and Spondylidae). Through this sampling strategy, we were able to independently estimate age of the superfamily. Interestingly, our time-calibrated phylogeny supports a somewhat earlier origin of the Pectinoidea (Late Devonian, 395 MYA) than currently accepted date of the Early Carboniferous period (358.9 MYA) when †Pernopectinidae is regarded as the stem group of the superfamily (see Waller 2006). Future inclusion of fossil taxa in

phylogenetically informed macroevoutionary analyses will be critical for interpreting patterns of diversification and extinction for the group.

Understanding relationships among the families of Pectinoidea could give an interesting context to the evolution of eyes within the superfamily. Eyes occur ventrally and often serially repeated on both left and right mantle lobes, located at the end of short stalks on the middle fold (Dakin, 1910). Scallops possess many single chambered eyes with a mirror-like reflector lining the back of the eye which focuses light back onto a double-retina system in the middle of the eye (Land, 1965; Palmer et al., 2017). Pectinidae and Spondylidae are known to have this unique eye structure, while Propeamussiidae were thought to lack eyes (Waller, 1972); however, the absence of eyes in propeamussiids may reflect their distribution in dysphotic (200 - 1000 m) or aphotic (>1000 m) depths (Waller, 2006). There has been some debate regarding the presence or absence of eyes in extinct entoliids. Eyes may be present in the extant genus *Pectinella* [(Waller, 2006) images of the eyes were not illustrated], but with only two extant species, fluid-preserved specimens are rare (e.g., no specimens in the largest US collection USNM, co-author EE Strong) and we have been unable to secure a specimen for examination. If eyes are present in the Entoliidae (Waller, 2006) and our phylogenetic hypothesis is corroborated with future analyses, its placement as the sister taxon to the remaining Pectinoidea suggests a single origin of eyes in the common ancestor of the superfamily. However, patterns of eye loss in the Propeamussiidae *sensu lato* need to be examined from both historical and habitat perspectives.

Conclusion

The results of these current analyses suggest a novel topology for relationships within the superfamily Pectinoidea. Our results tentatively indicate the Propeamussiidae may be polyphyletic, but the AU test results do not reject all alternative hypotheses in which the family is constrained to be monophyletic. The inclusion of molecular data for a species of Entoliidae for the first time provides the first test of its phylogenetic placement as the sister to all other Pectinoidea. Our phylogenetic hypothesis also impacts the interpretation of trait histories in the superfamily with implications to phenotypic evolution. For instance, our data tentatively supports the hypothesis for a single origin of eyes in the superfamily. Future work should focus on bolstering support for this scenario through the examination of a more comprehensive molecular dataset. However, if the relationships recovered here hold, a taxonomic revision of the Propeamussiidae is warranted.

Acknowledgements

The majority of Indo-Pacific specimens were obtained during expeditions (PI Philippe Bouchet) organized by the MNHN and Pro-Natura International as part of the *Our Planet Reviewed* program, and by the MNHN and the Institut de Recherche pour le Développement as part of the *Tropical Deep-Sea Benthos* program. Ship time was programmed on R.V. Alis by the French Oceanographic Fleet, on R.V. Vizconde de Eza by the Instituto Español de Oceanografia, on F.V. DA-BFAR by the Philippines Bureau of Fisheries and Aquatic Resources, and on chartered local fishing boats through funding from the Total Foundation and Prince Albert II of Monaco Foundation. The loans

of molecular samples from MNHN were arranged by Nicolas Puillandre and Barbara Buge. We are very grateful for the assistance and loans provided by the staff of museums and research institutions, especially G. Paulay and J. Slapcinsky [UF]; S. Morrison and C. Whisson [WAM]; P. Greenhall and C. Walter [NMNH]; A. Baldinger [MCZ]; M. Siddall and S. Lodhi [AMNH]; R. Bieler and J. Gerber [FMNH]; R. Kawamoto [BPBM]; E. Kools [CAS]. Members of the D.C. Adams, T.A. Heath, and N. Valenzuela labs provided valuable comments on an earlier version of this manuscript. The authors are grateful to two anonymous reviewers for helpful input. Authors declare no conflict of interest.

Funding: Financial support was provided by the United States National Science Foundation [DEB-1118884 to JMS]. Funding for specimen collections was provided by the Total Foundation, Prince Albert II of Monaco Foundation, Stavros Niarchos Foundation, Richard Lounsbery Foundation, the French Ministry of Foreign Affairs, and the Philippines Bureau of Fisheries and Aquatic Resources. The funding bodies did not play a role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

References

Alejandrino, Alvin, Loiuse Puslednik, and Jeanne M Serb. 2011. "Convergent and parallel evolution in life habit of the scallops (Bivalvia: Pectinidae)." *BMC Evolutionary Biology* 11 (1): 164.

Audino, J.A., J.E.A. Marian, A. Wanninger, and S.G. Lopes. 2015. "Anatomy of the pallial tentacular organs of the scallop Nodipecten nodosus (Linnaeus, 1758) (Bivalvia: Pectinidae)." *Zool. Anz.* 258: 39-46.

Audino, J.A., J.E.A. Marian, A. Wanninger, and S.G. Lopes. 2015. "Development of the pallial eye in Nodipecten nodosus (Mollusca: Bivalvia): insights into early visual performance in scallops." *Zoomorphology* 134: 403-415.

Audino, J.A., J.E.A. Marian, A. Wanninger, and S.G. Lopes. 2015. "Mantle margin morphogenesis in Nodipecten nodosus (Mollusca: Bivalvia): New insights into the development and the roles of bivalve pallial folds." *BMC Dev. Biol.* 15 (22). https://doi.org/10.1186/s12861-015-0074-9.

Bieler, Rudiger, Paula M. Mikkelsen, Timothy M. Collins, Emily A. Glover, Vanessa L. Gonzalez, Daniel L. Graf, Elizabeth M. Harper, et al. 2014. "Investigating the Bivalve Tree of Life - an exemplar-based approach combining molecular and novel morphological characters." *Invertebrate Systematics* 28: 32-115.

Castresana, J. 2000. "Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis." *Molecular Biology and Evolution* 17: 540-552.

Dakin, William J. 1910. "The eye of Pecten." *Quarterly Journal of Microscopical Science* 55: 49-112.

Dereeper, A, V Guignon, G Blanc, S Audic, F Chevenet, J.F. Dufayard, S Guindon, et al. 2008. *Phylogeny.fr: robust phylogenetic analysis for the non-specialist.* July 1. http://phylogeny.lirmm.fr/phylo_cgi/index.cgi.

Dijkstra, Henk H., and Philippe Maestrati. 2012. "Pectinoidea (Mollusca, Bivalvia, Propeamussiidae, Cyclochlamydidae n. fam., Entoliidae and Pectinidae) from the Vanuatu Archipelago." *Zoosystema* 34 (2): 389-408.

Edgar, Robert C. 2004. "MUSCLE: multiple sequence alignment with high accuracy and high throughput." *Nucleic Acids Research* 32 (5): 1792-1797.

Faggionato, Davide, and Jeanne M Serb. 2017. "Strategy to identify and test putative light-sensitive non-opsin G-protein-coupled receptors: A case study." *Biological Bulletin* 233: 70-82.

Felsenstein, J. 1981. "Evolutionary trees from DNA sequences: a maximum likelihood approach." *Journal of Molecular Evolution* 17: 368-376.

Giribet, Gonzalo, Salvador Carranza, Jaume Baguna, Marta Riutort, and Carles Ribera. 1996. "First Molecular Evidence for the Existence of a Tardigrada + Arthropoda Clade." *Molecular Biology and Evolution* 13 (1): 76-84.

Gollapalli, D., P. Maiti, and R. Rando. 2004. "RPE65 operates in the vertebrate visual cycle by stereospecifically binding all-trans-retinyl esters." *Biochemistry* 11824-11830.

Gomez, M.D.P., Lady Espinosa, Nelson Ramirez, and Enrico Nasi. 2011. "Arrestin in ciliary invertebrate photoreceptors: molecular identification and functional analysis in vivo." *Journal of neuroscience* 31 (5): 1811-1819.

Guderley, H.E., and I. Tremblay. 2013. "Escape responses by jet propulsion in scallops." *Canadian Journal of Zoology* 91: 420-430.

Hayami, I. 1991. "Living and fossil scallop shells as airfoils: an experimental study." *Paleobiology* 17: 1-18.

Hertlein, Leo G. 1969. *Fossiliferous Boulder of Early Tertiary Age from Ross-Island, Antarctica.* Vol. 4, in *Antarctic Journal of the United States*, 199.

Hohna, S., M. J. Landis, T. A. Heath, B. Boussau, N. Lartillot, B. R. Moore, J. P. Hueslsenbeck, and F. Ronquist. 2016. "RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specific language." *Systematic Biology* 65: 726-736.

Katoh, Kazutaka, and Daron M. Standley. 2013. "MAFFT Multiple Sequence Alignment Software Version7: Improvements in Performance and Usability." *Molecular Biology and Evolution* 30 (4): 772-780.

Kearse, M, R Moir, A Wilson, S Stones-Havas, M Cheung, S Sturrock, S Buxton, et al. 2012. "Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data." *Bioinformatics* 28 (12): 1647-1649.

Kingston, Alexandra C N, Alan M Kuzirian, Roger T Hanlon, and Thomas D Cronin. 2015. "Visual phototransduction components in cephalopod chromatophores suggest dermal photoreception." *The Company of Biologists* 1596-1602.

Kojima, D., A. Terakita, T. Ishikawa, Y. Tsukahara, A. Maeda, and Y. Shichida. 1997. "A novel Go-mediated phototransduction cascade in scallop visual cells." *J. Biol. Chem.* 272: 22979-22982.

Land, M.F. 1965. "Image formation by a concave reflector in the eye of the scallop, Pecten maximus." *J. Physiol.* 179: 138-153.

Lanfear, R., p. Frandsen, A. Wright, T. Senfield, and B. Calcott. 2016. "PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morhpological phylogenetic analyses." *Mol. Biol. Evol.* 34: 772-773.

Matsumoto, Masahiro. 2003. "Phylogenetic analysis of the subclass Pteriomorphia (Bivalvia) from mtDNA COI sequences." *Molecular Phylogenetics and Evolution* 27 (3): 429-440.

Matsumoto, Masahiro, and Itaru Hayami. 2000. "Phylogenetic Analysis of the Family Pectinidae (Bivalvia) Based on Mitochondrial Cytochrome C Oxidase Subunit I." *Journal Molluscan Studies* 66 (4): 477-488.

Mau, B., M. Newton, and B. Larget. 1999. "Bayesian phylogenetic inference via Markov chain Monte Carlo methods." *Biometrics* 55: 1-12.

Miller, M.A., W. Pfeiffer, and T. Schwartz. 2010. "Creating the CIPRES Science Gateway for inference of large phylogenetic trees." New Orleans, LA, Nov. 14. 1-8.

Morton, B., and M.H. Thurston. 1989. "The functional morphology of Propeamussium lucidum (Bivalvia, Pectinacea), a deep-sea predatory scallop." *J. Zool.* 218: 471-496.

Pairett, A. N., and J. M. Serb. 2013. "De Novo assembyl and characterization of two transcriptomes reveal multiple light-mediated functions in the scallop eye (Bivalvia: Pectinidae)." *PLoS One* 8: e69852. https://doi.org/10.1371/journal.pone.0069852.

Palmer, B. A., G. J. Taylor, V. Brumfeld, D. Gur, M. Shemesh, N. Elad, A. Osherov, D. Oron, S. Weiner, and L. Addadi. 2017. "The image-forming mirror in the eye of the scallop." *Science* 358 (80): 1172-1175.

Piatigorsky, J., Z. Kozmik, J. Horwitz, L. Ding, E. Carosa, W. G. Robison, P. J. Steinbach, and E. R. Tamm. 2000. "Omega-crystallin of the scallop lens: a dimeric aldehyde dehydrogenase class1/2 enzyme-crystallin." *J. Biol. Chem.* 275: 41064-41073.

Plazzi, F., A. Ceregato, M. Taviani, and M. Passamonti. 2011. "A molecular phylogeny of bivalve mollusks: ancient radiations and divergences as revealed by mitochondrial genes." *PLoS One* 6.

Plazzi, F., and M Passamonti. 2010. "Towards a molecular phylogeny of Mollusks: bivalves' early evolution as revealed by mitochondrial genes." *Mol. Phylogenet. Evol.* 57: 641-657.

Poli, G. 1791. "Testacea utriusque siciliae eorumque historia et anatome tabulis aeneis illustrata." *Parmae: Ex regio typographeio* 2.

Porath-Krause, Anita J., Autumn N. Pairett, Davide Faggionato, Bhagyashree S. Birla, Kannan Sankar, and Jeanne M. Serb. 2016. "Structural differences and differential expression among rhabdomeric opsins reveal functional change after gene duplication in the bay scallop, Argopecten irradians (Pectinidae)." *BMC Evolutionary Biology* 16 (250).

Puslednik, Louise, and Jeanne M. Serb. 2008. "Molecular phylogenetics of the Pectinidae (Mollusca: Bivalvia) and effect of increased taxon sampling outgroup selection on tree topology." *Molecular Phylogenetics and Evolution* 48 (3): 1178-1188.

Rambaut, A., A. Drummond, D. Xie, and M. Suchard. 2018. "Posterior summarization in Bayesian phylogenetics using Tracer 1.7." *Systematic Biology* 67 (5): 901-904.

Ronquist, F., S. Klopfstein, L. Vilhelmsen, S. Schulmeister, D.L. Murray, and A. P. Rasnitsyn. 2012. "A total-evidence approach to dating with fossils, applied to the early radiation of the hymenoptera." *Syst. Biol.* 61: 973-999.

Serb, Jeanne M. 2016. "Reconciling Morphological and Molecular Approaches in Developing a Phylogeny for the Pectinidae (Mollusca: Bivalvia)." *Developments in Aquculture and Fisheries Science*, Third ed.

Serb, Jeanne M., A. Alejandrino, E. Otarola-Castillo, and D. C. Adams. 2011. "Shell shape quantification using geometric morphometrics reveals morphological convergence of distantly related scallop species (Pectinidae)." *Zool. J. Linn. Soc.* 163: 571-584.

Serb, Jeanne M., Anita J. Porath-Krause, and Autum N. Pairett. 2013. "Uncovering a gene duplication of the photoreceptive protein, opsin, in scallops (Bivalvia: Pectinidae)." *Intergrative and comparative Biology* 53 (1): 68-77.

Serb, Jeanne M., Emma Sherratt, Alvin Alejandrino, and Dean C. Adams. 2017. "Phylogenetic convergence and multiple shell shape optima for gliding scallops (Bivalvia: Pectinidae)." *Journal of evolutionary biology* 30 (9): 1736-1747.

Sharma, P. P., V. L. Gonzalez, G. Y. Kawauchi, S.C.S. Andrade, A. Guzman, T. M. Collins, E. A. Glover, et al. 2012. "Phylogenetic analysis of four nuclear protein-encoding genes largely corroborates the traditional classification of Bivalvia (Mollusca)." *Mol. Phylogenet. Evol.* 65: 64-74.

Sherratt, Emma, Alvin Alejandrino, Andrew C. Kraemer, Jeanne M. Serb, and Dean C. Adams. 2016. "Trends in the sand: Directional evolution in the shell shape of recessing scallops (Bivalvia: Pectinidae)." *Evolution international journal of organic evolution* 70 (9): 2061-2073.

Shimodaira, H, and M Hasegawa. 2001. "CONSEL: for assessing the confidence of phylogenetic tree selection." *Bioinformatics* 17: 1246-1247.

Shimodaira, H. 2002. "An approximately unbiased test of phylogenetic tree selection." *Syst. Biol.* 51: 492-508.

Speiser, Daniel I., and L. A. Wilkens. 2016. "Neurobiology and behaviour of the scallop." In *Scallops: Biology, Ecology, Aquaculture, and Fisheries*, edited by S. E. Shumway and G. J. Parsons, 219-300. Elseiver.

Speiser, Daniel I., and Sonke Johnsen. 2008. "Comparative Morphology of the Concave Mirror Eyes of Scallops (Pectinoidea)." *American Malacological Bulletin* 26: 27-33.

Speiser, Daniel I., E. R. Loew, and S. Johnsen. 2011. "Spectral sensitivity of the concave mirror eyes of scallops: potential influences of habitat, self-screening and longitudinal chromatic aberration." *J. Exp. Biol.* 214: 422-431.

Speiser, Daniel I., Y. L. Gagnon, R. K. Chhetri, A. L. Oldenburg, and S. Johnsen. 2016. "Examining the effects of chromatic aberration, object distance, and eye shape on image-formation in the mirror-based eyes of the bay scallop, Argopecten irradians." *Integr. Comp Biol.* 56: 796-808.

Stamatakis, A. 2014. "RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies." *Bioinformatics* Open Access.

Stanley, S. M. 1970. "Relation of shell form to life habits of the Bivalvia (Mollusca)." *Geol. Soc. Am. Memoirs* 125: 1-296.

Sun, W., and L. Gao. 2017. "Phylogeny and comparative genomic analysis of Pteriomorphia (Mollusca: Bivalvia) based on complete mitochondrial genomes." *Mar. Biol. Res.* 13: 255-268.

Talavera, G, and J Castresana. 2007. "Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments." *Systematic Biology* 56: 564-577.

Tremblay, Isabelle, Myriam Samson-Do, and Helga E. Guderley. 2015. "When Behavior and Mechanics Meet: Scallop Swimming Capacities and Their Hinge Ligament." *Journal of Shellfish Research* 34 (2): 203-212.

Valentine, J. W., D. Jablonski, S. Kidwell, and K. Roy. 2006. "Assessing the fidelity of the fossil record by using marine bivalves." *Proc. Natl. Acad. Sci. USA* 103: 6599-6604.

Waller, Thomas. 1991. "Evolutionary relationships among commercial scallops (Mollusca: Bivalvia: Pectinidae)." *Scallops: Biology, Ecology and Aquaculture* 1-73.

Waller, Thomas. 1978. "Morphology, morphoclines and a new classification of the Pteriomorphia (Mollusca: Bivalvia)." *Philosophical Transaction of the Royal Society B1* 284 (1001): 345-365.

Waller, Thomas R. 2006. "Phylogeny of families in the Pectinoidea (Mollusca: bivalvia): importance of the fossil record." *Zoological Journal of the Linnean Society* 148: 313-342.

Waller, Thomas R. 1972. "The functional significance of some shell microstructures in the Pectinacea (Mollusca: Bivalvia)." *Int. Geol. Congr. 24th Sess. Sect.* 7: 48-56.

Whiting, Michael F., James C. Carpenter, Quentin D. Wheeler, and Ward C. Wheeler. 1997. "The Strepsiptera Problem: Phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology." *Systematic Biology* 46 (1): 1-68.

Figures and Tables

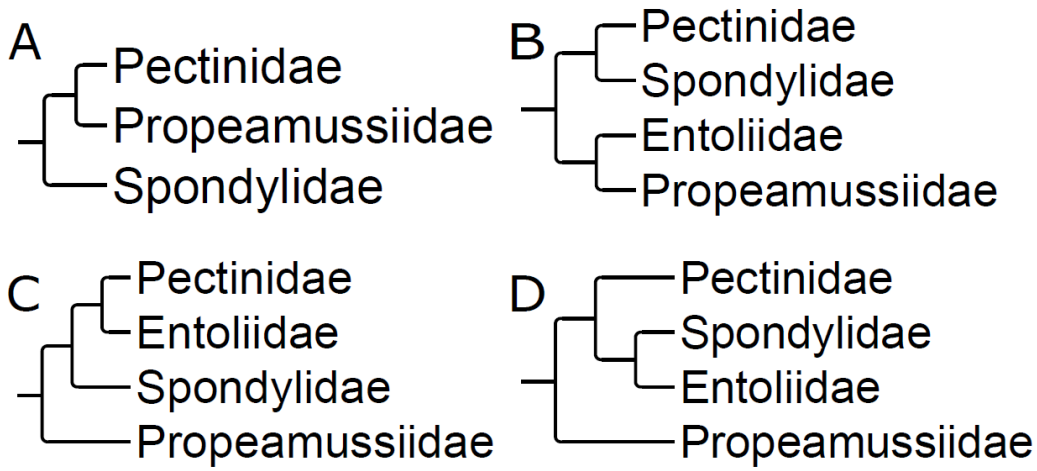


Figure 2-1 - Existing hypotheses of relationships among pectinoidean families. A, molecular data (Bieler et al., 2014; Matsumoto and Hayami, 2000); B, paleontological and morphological data (Waller, 2006); C, paleontological and morphological data (Waller, 1991); D, morphological data (Waller, 1978).

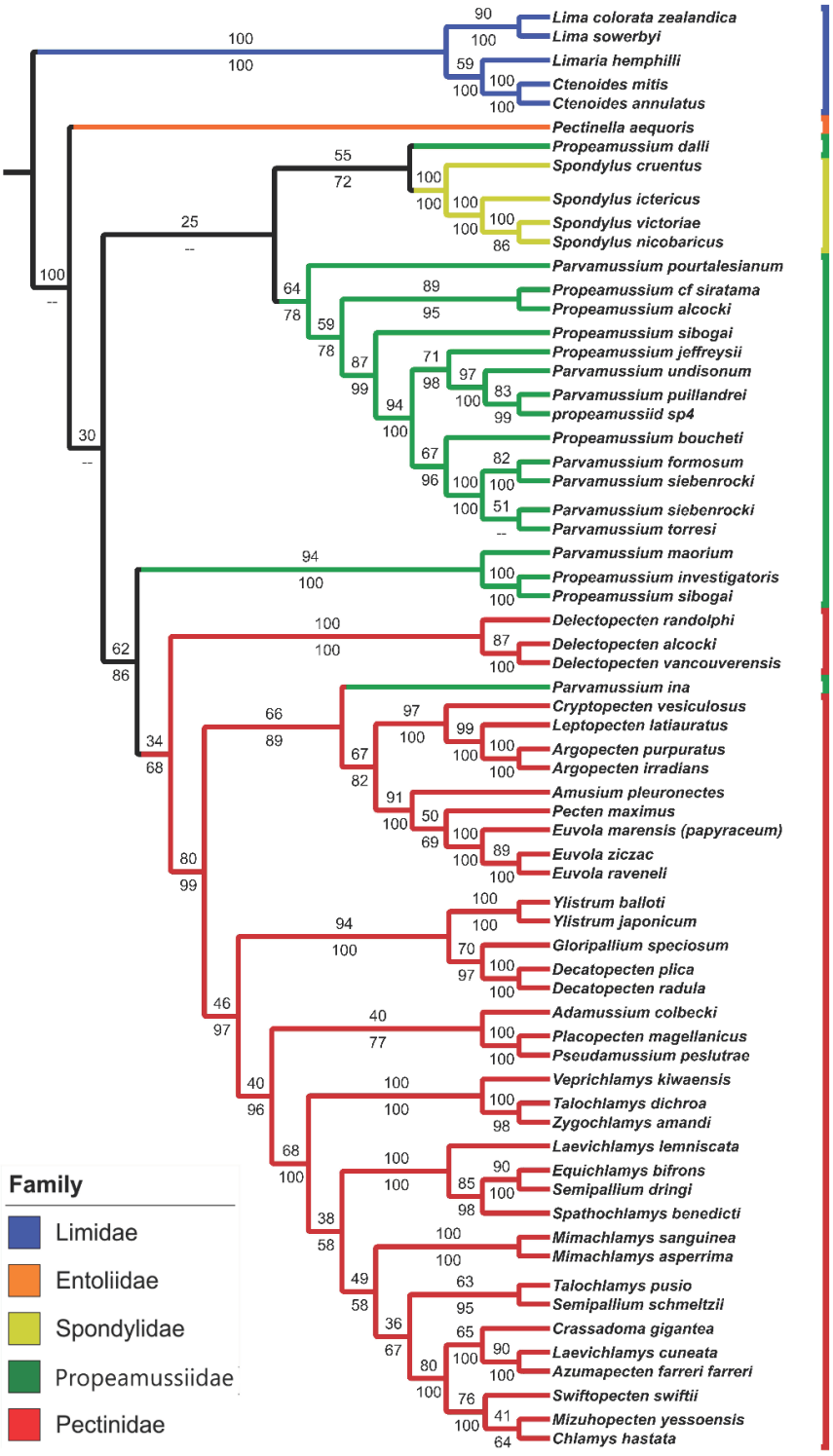


Figure 2-2. Maximum likelihood phylogeny of pectinoidean families (lnL=−25647.23) based on combined 12S, 16S, 18S, 28S and histone H3 sequences. Taxa are color coded by family. Numbers above the branches indicate bootstrap support; numbers below branches are Bayesian posterior probabilities. A dash (–) indicates no support for that node.

Figure 2-3 - Divergence time estimation analysis of Pectinoidea inferred via Bayesian Inference under the Fossilized Birth-Death model. Bottom axis represents millions of years with tree branches colored by family as previously noted. Fossil samples used to calibrate internal nodes are represented by red circle with estimated node ages also in red. 95% HPD are reported as grey bars and Bayesian Posterior Probabilities are reported for each node. For the ingroup, morphological data supporting the presence or absence of eyes are indictated by a black square or a square with an "x", respectively. White boxes indicate lack of data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2-1 - Summary of AU tests of alternative pectinoidean topologies

Topology	AU test *p*-value*
Unconstrained ML	0.599
Monophyletic Propeamussiidae	0.323
Monophyletic Propeamussiidae, excluding *Pa. ina*	0.574
Figure 2-1A Hypothesis	**0.011**
Figure 2-1B Hypothesis	**0.014**
Figure 2-1C Hypothesis	**0.005**
Figure 2-1D Hypothesis	0.069

* RAxML constraint analyses and corresponding *p*-values of AU tests implemented in CONSEL. Siginficantly different topologies are in bold.

CHAPTER 3. MOLLUSCAN TRANSCRIPTOMES SUGGESTS THE PRESENCE OF A LIGHT INDEPENDENT RETINOID VISUAL CYCLE HOMOLOGOUS TO VERTEBRATE AND INSECT VISUAL CYCLES.

G. Dalton Smedley[1], Kyle McElroy[1], Jeanne M. Serb[1]

[1]Department of Ecology, Evolution, and Organismal Biology; Iowa State University; Ames, Iowa, U.S.A. 50011

Abstract

Vision is a two-step process where light is absorbed by a photopigment initiating a signaling cascade, followed by the resetting of the photopigment. This second pathway is the retinoid visual cycle in which all-*trans* retinal undergoes isomerization back to a functional 11-*cis* isomer and the photopigment is reconstituted. There are currently three described retinoid visual cycles for vertebrates, insects, and molluscs. However, the under-studied molluscan visual cycle does not explain how the photopigments become prepared for exposure to the first light of day nor how the cellular toxicity inherent to the retinal chromophore is managed. To better understand the components in the molluscan visual cycle, we assembled and queried publicly available transcriptomes from 50 molluscan and 14 other protostome transcriptomes using known vertebrate and insect visual cycle proteins. We then used phylogenetically-informed annotation methods to test homology and predict putative function of these candidate transcripts. We constructed three gene family trees. Our results indicate a lack of invertebrate retinyl storage proteins, supporting the hypothesis that LRAT and CRBP are novel vertebrate proteins. The topologies of the short-chain dehydrogenases/reductases and retinal shuttle proteins gene trees suggest the

presence of homologous proteins between the three visual cycles, while the carotenoid oxygenase phylogeny showed no evidence of a homologous RPE65 isomerohydrolase protein in either molluscs or arthropods. With these results we propose a new retinoid visual cycle model for molluscs that is most similar in structure to that described in insects. Interestingly, the putative homology between insect and mollusc proteins suggests the operation of some visual cycle processes outside the photoreceptor cells. Further work is required to functionally understand the presence and activity of a light-independent visual cycle in molluscs as well as localization studies to determine the cell specificity.

Introduction

Eyes have evolved over 60 times across animals (von Salvini-Plawen and Mayr 1977), and despite the diversity in morphology, the molecules responsible for light perception have been largely conserved (Findlay and Pappin 1986; Feuda et al. 2012). Phototransduction is the process by which photons of light are converted to a nerve impulse ultimately leading to light perception. The functional unit of phototransduction is the photopigment, composed of an opsin apoprotein bound to a retinal chromophore. Opsins are seven transmembrane G-protein coupled receptors (GPCR), a large gene family that can be subdivided into nine clades, defined, in part, by the G-protein they bind (Ramirez et al. 2016), and the retinal chromophore is a derivative of vitamin A, which acts as the light absorbing molecule of the photopigment.

Vision is a two-step process. The first step is phototransduction and begins when a photon of light is absorbed by the 11-*cis* retinal chromophore bound to the opsin. This absorption causes isomerization of the retinal chromophore from an 11-*cis* to all-*trans*

retinal and a conformational change of the opsin protein called meta-opsin. The conformational change allows release of the G-protein bound to the opsin, initiating the signaling cascade, ultimately leading to the nerve impulse and thus light perception. The all-*trans* retinal isomer makes the photopigment inert, and therefore a new functional 11-*cis* isomer is required to initiate another signaling cascade. The second step of phototransduction is the resetting of the photopigment by uptake of a new 11-*cis* retinal isomer or recycling the all-*trans* isomer. This process of resetting the photopigment is called the retinoid visual cycle (Wald 1935).

The retinoid visual cycle serves multiple purposes. In addition to resetting the phototransduction pathway, the molecular machinery of the visual cycle also helps to alleviate some of the inherent physiological problems tied with the use of retinal as a chromophore. Due to the oxidative properties of retinal, it is relatively insoluble and prone to degradation, making it toxic to photoreceptive cells in relatively small doses (Gonzalez-Fernandez, 2002). Dehydrogenases and shuttle proteins within the visual cycle are used to control concentrations of free retinal isomers and transport these isomers across cell membranes. The visual cycle also primes the photoreceptive pathways for first light of day. A dark regeneration pathway of the visual cycle ensures that there is a functional 11-*cis* retinal isomer bound to opsins, creating a functional photopigment, so that phototransduction can be induced after periods of darkness. Finally, the visual cycle is important in the isomerization of retinal produced from metabolic pathways. Three retinoid visual cycles in vertebrates, insects, and molluscs have been described to solve these problems and reset the phototransduction pathway.

The vertebrate visual cycle functions via a thoroughly described complex of enzymes and shuttle proteins (Figure 3-1a) (Lamb and Pugh 2004). Vertebrates use a monostable visual photopigment, which becomes thermodynamically unstable upon creation of the meta-rhodopsin, causing the release of the all-*trans* retinal. The opsin must bind to a new 11-*cis* retinal in order to reconstitute the photopigment. Following isomerization of the 11-*cis* retinal to all-*trans*, retinal dehydrogenase 8/12 (RDH8/RDH12) reduces the all-*trans* retinal to an all-*trans* retinol (ROL) (Maeda et al. 2007) before the retinol becomes bound to interstitial retinol binding protein (IRBP). This shuttle protein transports the all-*trans* retinol between the photoreceptor cell and the retinal pigment epithelium (RPE). The retinol is then passed to a four-protein complex consisting of lecithin-retinol acyltransferase (LRAT), an isomerohydrolase (RPE65), cellular retinol binding protein (CRBP), and retinol dehydrogenase 5 (RDH5). LRAT adds a palmitate group to the retinol, creating all-*trans* retinyl and allowing the retinyl to either continue through the visual cycle pathway for recycling or enter a retinyl ester storage pathway (MacDonald and Ong 1988; Ong, MacDonald, and Gubitosi 1988; Saari, Bredberg, and Farrell 1993). Unlike other retinal isomers involved in the visual cycle pathway, the all-*trans* retinyl is a stable, non-toxic form of vitamin A, allowing it to function as a storage molecule for retinal isomers (Travis et al. 2007). If the retinoid stays within the cycle, RPE65 acts as the isomerohydrolase within the complex converting all-*trans* retinyl to 11-*cis* retinol (Gollapalli and Rando 2003; Mata et al. 2004; Jin et al. 2005). Finally, RDH5 converts 11-*cis* retinol to an 11-*cis* retinal (Jang et al. 2001) allowing IRBP to bind the retinal and transport it back into the photoreceptor cell to bind to the opsin and begin the cycle anew.

Unlike monostable photopigments of vertebrates, insects utilize bistable photopigments that remain thermodynamically stable after the isomerization to the all-*trans* 3-hydroxyretinal (3-OH-all-*trans* RAL) chromophore. Bistability allows the opsin protein to retain the chromophore upon absorption of light (Hamdorf and Schwemer 1975; Hillman, Hochstein, and Minke 1983). The chromophore can then be isomerized back to an 11-*cis* 3-hydroxyretinal (3-OH-11-*cis* RAL) by the absorption of a second photon of light at a different wavelength, resetting the system without the need for the exchange of chromophores. However, recent studies of visual systems in insects identified of a dark regeneration visual cycle that takes place primarily within the retinal pigment cells (RPC) (Figure 3-1b). The retinoid visual cycle in insects is hypothesized to function in the event of opsin degradation causing release of the all-*trans* 3-hydroxyretinal or from the production of all-*trans* 3-hydroxyretinol as a result of metabolic pathways (X. Wang et al. 2010, 2012; Montell 2012). Regardless of the origin of the all-*trans* 3-hydroxyretinal, the isomer is reduced to all-*trans* 3-hydroxyretinol (3-OH-all-*trans* ROL) by photoreceptor dehydrogenase (PDH) (X. Wang et al. 2010). Prolonged depolarization afterpotential is not apparent (PINTA) is a CRAL-TRIO domain shuttle protein and has been shown to bind and transport the 3-hydroxyretinols within the RPC. While PINTA is believed to function in the transport of 3-hydroxyretinol isomers during the visual cycle, the extent of its function is still not completely known (T. Wang and Montell 2005). The all-*trans* 3-hydroxyretinol then binds to an isomerase and isomerizes the all-*trans* 3-hydroxyretinol to an 11-*cis* 3-hydroxyretinol. This step is thought to be light-dependent, and although a photoisomerase phylogenetically related to the molluscan retinochrome has been found in insects, it was only isolated from the

photoreceptor cells and has yet to be fully characterized (Pepe and Cugnoli 1992; Smith and Goldsmith 1991; Macias-Muñoz, Rangel Olguin, and Briscoe 2019). The newly formed 11-*cis* 3-hydroxyretinol is oxidized to an 11-*cis* 3-hydroxyretinal by retinol dehydrogenase B (RDHB) and delivered back to the rhodopsin completing the cycle (X. Wang et al. 2012). There are still unknowns surrounding the function of this newly described pathway including the identity of the isomerase or possible isomerohydrolase and the identity of the transport system of the 3-hydroxyretinal isomers between the retinal pigment and photoreceptor cells.

The third system is found in molluscs (Figure 3-1c). Molluscs possess both monostable and bistable visual photopigments and the visual cycle was thought to function via a photoisomerase opsin protein, retinochrome (RTC) (Sperling and Hubbard 1975). In molluscs, all-*trans* retinal is shuttled from a visual opsin to RTC via a retinaldehyde binding protein (RALBP) where RTC binds the all-*trans* retinal. The RTC photoisomerase absorbs a photon of light isomerizing the chromophore to an 11-*cis* isomer. The shuttle protein then rebinds the 11-*cis* retinal and transports it back to the visual opsin, resetting the phototransduction system. Unlike the vertebrate and insect visual systems, the current molluscan visual cycle does not require the transport of these retinal isomers between photoreceptor and pigment cell types. Instead, visual opsins are imbedded in the outer segments of the photoreceptor cells and RTC is imbedded in the inner segment (Hara and Hara 1972, 1976). The lack of clarity for how molluscs handle the physiological dangers of retinal chromophore coupled with new insights into the insect visual cycle (X. Wang et al. 2010; Arshavsky 2010; Montell 2012) calls into question the simplicity of the currently described molluscan visual cycle.

Here, we hypothesize that molluscs use a more complex and light-independent visual cycle similar to the described cycles in vertebrates and insects. We predict that proteins within molluscan photoreceptive tissues would be homologous to dehydrogenase, shuttle, and storage proteins found in vertebrate and insect visual cycles. To test this hypothesis, we used experimentally determined visual cycle proteins from vertebrates and insects to query available transcriptomes from 50 molluscan species and 14 other protostome transcriptomes. Phylogenetic reconstructions of CRAL-TRIO shuttle proteins, retinol dehydrogenases, and carotenoid oxygenase protein families showed homology among vertebrates, insects, and molluscs. These results provide two important insights about the molluscan visual cycle: 1) the presence of dehydrogenases homologous to insect and vertebrate proteins, which suggests a more complicated visual cycle that includes a light-independent path, and 2) the absence of LRAT and CRBP which indicates the lack of a retinyl ester storage pathway in molluscs.

Methods

Transcriptome Assemblies and Quality Scoring

Raw RNA sequencing data was downloaded from National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) for 64 species including 50 molluscan species, three lophotrochozoan species outside molluscs, 10 species from ecdysozoa, and one cnidarian species listed in Supplemental Table 3-S1. Transcriptomic datasets were selected based on two primary criteria: 1) breadth of taxonomic representation and 2) photosensitive specificity of tissues. For each taxonomic lineage, available datasets were prioritized based on tissue type, where

photosensitive tissues such as eyes and mantle were the highest priority, followed by head or nervous tissues, and then whole organism. Only datasets with pair-ended reads were included due to the greater accuracy of read alignment with these data relative to single-ended reads. SRA files were downloaded, and the quality of the raw reads was analyzed prior to assembly using FastQC v0.11.5 (Andrews, 2010). Raw sequences were subjected to the trimmomatic package of Trinity v2.6.6 (Grabherr et al. 2011; Haas et al. 2013) to perform trimming of raw sequences under default parameters (ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 SLIDINGWINDOW:4:20 MINLEN:25) with additional trimming of sequence heads and tails based on the FastQC results. Clean reads were then assembled using Trinity de novo assembly package given with normalization based on read set and default parameters. To reduce the redundancy of transcripts, assembled transcripts were clustered using CD-HIT v4.6.8 (Fu et al. 2012; W. Li and Godzik 2006). Nucleotide transcripts were clustered using a 95% identity clustering threshold with a word size of five.

The quality of transcript assemblies was measured via three methods: mapping RNA-seq reads back to the assembled transcriptome, measuring gene completeness, and measuring sequence length distribution (Supplemental Table 3-S1). Input reads were mapped to the assembled transcripts with Bowtie2 v2.3.4.1 (Langmead and Salzberg 2012), using default parameters to quantify the percent of reads corresponding to assembled transcripts. BUSCO (Benchmarking Universal Single-Copy Orthologs) v3.0.1 (Seppey, Manni, and Zdobnov 2019) was used, with default parameters (MODE: trans), to assess transcriptome completeness based on the metazoan lineage database (*odb9*). The N50 score is also reported from the Trinity

RNAseq output; however, ExN50 is considered a more informative statistic because it is limited to the more highly expressed transcripts reducing the bias inherent do to lesser expressed transcripts (Geniza and Jaiswal 2017). To calculate ExN50, RSEM (RNA-Seq by Expectation Maximization) (B. Li and Dewey 2014) was used to calculate transcript abundance estimates and automatically generate gene count matrix given Bowtie2 as the alignment method.

Blast and Transcript Homology Assessment

In order to identify transcripts homologous to known visual cycle proteins, Transdecoder (Haas et al. 2013) was used to translate transcripts and identify open reading frames with a length of >100 amino acids. Transcripts blasted with blastp via ncbi-blast+ v2.7.1 (Altschul et al. 1990) against query proteins (Supplemental Table 3-S1). Blast results were stringent, retaining only transcripts with at least an e-value of 1e-20. Selected transcripts were then collected using the blastdbcmd command in blast+ to extract transcripts as fasta sequences. These sequences were imported into Geneious v4.7.6 (Kearse et al. 2012) and separated by protein family based on motif identification via Uniprot database (Table 3-1).

Gene tree reconstruction and reconciliation

Reconciliation methods are used to explain topology differences between a species tree and a gene tree when evolutionary events other than speciation have occurred. We assembled a species supertree from a combination of published phylogenetic trees (Zapata et al. 2014; Kocot et al. 2011, 2019, 2017; Chiari et al. 2012; Goodheart et al. 2017; Lindgren et al. 2012; González et al. 2015; Borner et al. 2014).

Then, using blast results, we built three gene trees to test for homology between transcripts and query proteins from the respective retinoid visual cycles. Sequences were aligned by protein family using MAFFT v7.221 (Katoh and Standley 2013) to create three protein family datasets: 1) carotenoid oxygenases (RPE65, BCOI, BCOII, and ninaB); 2) CRAL-TRIO (CRALBP, PINTA, and RALBP), and 3) short-chain dehydrogenases/reductases (PDH, RDH5, RDH8, RDH11, and RDH12). Blasting similar or closely related gene sequences often resulted in the identification of a single transcript for multiple queried proteins. Because of this ambiguity, duplicates of transcripts from a given transcriptome were deleted. TrimAl v1.2 (Capella-Gutiérrez, Silla-Martínez, and Gabaldón 2009) was used to trim sequence ends and remove large gaps within the alignment sequences under a heuristic method to identify the best automated method. We analyzed all gene family alignments under Maximum Likelihood (ML) using IQ-Tree v1.6.10 (Nguyen et al. 2015) on the Cipres Science Gateway (Miller, Pfeiffer, and Schwartz 2010). A phylogeny was built for each protein family with IQ-Tree using 10,000 Ultrafast bootstrap replicates as a search parameter and allowing IQ-Tree to select the substitution model using Bayesian Information Criterion (BIC). A series of vertebrate amino acid sequences were pulled from GenBank to use as anchor sequences within each gene phylogeny and suggest putative homology of identified transcripts. NOTUNG v2.9.1.3 (Darby et al. 2016; Stolzer et al. 2012; Durand, Halldórsson, and Vernot 2006; Vernot et al. 2008; Chen, Durand, and Farach-Colton. 2000; Zmasek and Eddy 2001) was used to reconciled the gene trees against a summarized species tree. The same species tree was used for the three complete gene trees.

The protein phylogenies of CRAL-TRIO, carotenoid oxygenase, and short-chain dehydrogenase/reductase protein families proved cumbersome to analyze due to very large trees and clades composed of transcripts from the same taxa. Because the hypothesis asks presence or absence, we individually blasted the transcripts from each respective tree using the NCBI GenBank database. Transcripts which were less than half the length of the total alignment after the alignment was trimmed or whose blast results did not identify the transcript as a close relative to one of the query proteins with strong support (e-value < 1e-50) were removed from the final tree.

Results

Transcriptome Assemblies

In total, 64 sequence libraries were downloaded from NCBI GenBank, including 50 molluscan species covering seven classes and 14 outgroup species. The outgroup was composed of three non-molluscan species of Lophotrochozoa, nine arthropods, one tardigrade, and one cnidarian species (Supplemental Table 3-S2). Filtered data was used for *de novo* transcriptome assembly. The Bowtie2, BUSCO, and Ex90N50 scores of these three analyses vary across transcriptomes, with most transcriptomes having high scores (Bowtie2: >90% and BUSCO: >90% complete). The aplacophoran mollusc *Chaetoderma sp*. (Bowtie2: 46.07% and BUSCO: 21.8% complete), gastropod *Microhedyle glandulifera* (Bowtie2: 56.77% and BUSCO: 31.4% complete), and jumping spider *Habronattus sansoni* (Bowtie2: 67.79% and BUSCO: 11.0% complete) low assembly scores. Four of the fourteen bivalve transcriptomes had complete BUSCO scores of less than 35% complete. While these datasets were still queried, the few blast

hits from these transcriptomes were removed due to short length and high e-values thus not included in the subsequent phyogenetic analyses.

Blast results and phylogenetic reconstruction

We used sixteen vertebrate and insect genes as query sequences to identify homologs within molluscan and other invertebrate transcriptomes. The resulting transcripts were used as datasets within the respective gene families to reconstruct gene trees. A query sequence of known function reconciling within a clade of blasted transcripts can be used to infer homology by putative function of the newly identified sequence (Speiser et al. 2014). We used maximum likelihood analyses to construct gene trees and reconcile these gene trees with a species supertree using NOTUNG to identify putative homologs. Our blast results identified many transcripts for the short-chain dehydrogenase/reductase, carotenoid oxygenase, and CRAL-TRIO gene families. However, the transcriptome searches did not identify any transcripts for intercellular retinal binding protein (IRBP) or lecithin: retinol acyltransferase (LRAT).

Lipocalin Blast Results

Lipocalin shuttle proteins serve an important role as shuttle proteins of retinoids within the vertebrate visual cycle. From the 64 searched transcriptomes, 26 transcripts were identified based on the lipocalin query sequences; some species, such as the common bay scallop, *Argopecten irradians,* and Suminoe oyster, *Crassostrea ariakensis*, had multiple transcripts. We used these putative lipocalin transcripts in a phylogenetic reconstruction with other shuttle proteins CRBP and CRABP from mouse and human, respectively. The resulting lipocalin gene tree had the two vertebrate

anchor sequences resolving in a clade containing sequences of *A. irradians* and *C. ariakensis* with this clade sister to 24 remaining lipocalin transcripts (Bootstrap, BS: 98). We selected lipocalin transcripts from protostomes from this tree and blasted them against the NCBI GenBank database to further support their identity as CRBP or CRABP proteins. However, these subsequent blast results showed the identity of the transcripts to be more similar to fatty-acid binding proteins. From this analysis, we cannot confidently claim presence of homologous lipocalin proteins from the invertebrate transcriptomes searched.

CRAL-TRIO homology

CRAL-TRIO domain proteins are a family of shuttle proteins described to transport lipophilic molecules between intercellular membranes, named from vertebrate CRALBP and TRIO guanine exchange factor (UniProt database). Three query proteins belong to the protein family of CRAL-TRIO: mouse CRALBP, *Drosophila* PINTA, and *Todarodes pacificus* RALBP. The blastp analysis identified 353 transcripts, which were used to generate a protein phylogeny of the CRAL-TRIO domain protein family using a general matrix for amino acid exchange rates (LG; (Le and Gascuel 2008)) and an empirical frequency of amino acids based on the data (+F) with a free rate model given nine categories (Yang 1995; Soubrier et al. 2012) (Supplemental Figure 3-S1). The final tree contained 147 transcripts and highlights the presence of both a RALBP shuttle protein and a putative PINTA-like protein in molluscs.

Vertebrate anchor sequences of CRALBP and TTPA were recovered in two monophyletic clades. Vertebrate TTPA is a member of the CRAL-TRIO protein family

and is a shuttle protein described to transport vitamin E (UniProt database). TTPA was not a query protein, but was included after preliminary blast results suggested identified transcripts to be TTPA. The vertebrate CRALBP clade resolved sister to a clade of lophotrochozoan transcripts (BS: 100). Both vertebrate CRALBP and molluscan RALBP function in the cellular transport of retinal isomers within the photoreceptor cells. Vertebrate TTPA sequences form a sister clade to a larger clade of lophotrochozoan transcripts with low supports (bootstrap, BS: 44). This clade of lophotrochozoa transcripts resolves with sequences of molluscan RALBP included to suggest putative function of transcripts. One final superclade of interest is composed of two distinct clades sister to one another (BS: 71) composed of lophotrochozoan transcripts and ecdysozoan transcripts, respectively. Within the ecdysozoan clade rests the *Drosophila melanogaster* PINTA sequence; however blasting transcripts within either clade identifies mostly TTPA-like and clavesin-like results. The relationship between the ecdysozoan and lophotrochozoan clades may indicate the presence of a PINTA-like protein within Lophotrochozoa meaning the origin of this protein predates the split between Lophotrochozoa and Ecdysozoa. NOTUNG analysis of the CRAL-TRIO protein tree shows 62 duplications and 291 losses.

Carotenoid oxygenase homology

Carotenoid oxygenases are a family of enzymes that cleave carotenoids to produce retinol in the visual cycle. We used four query proteins belong to the protein family of carotenoid oxygenases: human BCOI and BCOII, mouse RPE65, and *Drosophila* ninaB. BCOI and BCOII are used in the metabolic cleavage of vitamin A to two molecules of retinal; RPE65 functions as the isomerohydrolase in the vertebrate

visual cycle, isomerizing all-*trans* retinyl to 11-*cis* retinol. NinaB has been shown to function in both capacities, as a cleavage enzyme of vitamin A and an isomerohydrolase in the metabolism of retinals. All 109 transcripts identified from the blastp analysis of the invertebrate transcriptomes used to generate a gene phylogeny of the carotenoid oxygenase protein family. After removal of short sequences and those with low secondary blast scores, the final tree contained 63 transcripts and suggests a lack of a sequences in molluscs and other searched transcriptomes homologous to vertebrate RPE65-like (Supplemental Figure 3-S2).

A vertebrate-only BCOII clade resolved sister to all other transcripts within this phylogeny. A reciprocally monophyletic vertebrate BCOI and vertebrate RPE65 clade was the sister group to a superclade of lophotrochozoan and ecdysozoan transcripts (lopho + (ecdys + lopho)). The ecdysozoan clade contains a *Drosophila busckii* ninaB and *D. melanogaster* BCOI sequence and support values within these clades are generally high (BS>75), but low (BS: 44) at nodes separating the clades. The topology of this tree suggests multiple duplications of the carotenoid oxygenase BCOI after the protostome-deuterostrome split. NOTUNG reconciliation reports 72 instances of duplication and 125 losses in the carotenoid oxygenase protein tree.

Short-chain dehydrogenase homology

Short-chain dehydrogenases/reductases (SRD) are a large family of proteins which function in the visual cycle to oxidize or reduce retinal isomers. Five query proteins belong to the large protein family of SRD: mouse RDH5, RDH8 and RDH12, human RDH11, and *Drosophila* PDH. While not included in the original blast searches of the transcriptomes, vertebrate sequences for RDH2, RDH3, RDH5, RDH7, RDH8,

RDH10, RDH11, RDH12, RDH13, RDH14, and RDH16 and *Drosophila* RDHB were included in the phylogenetic analyses. All these dehydrogenases are closely related and many overlap in function within the visual cycle in their respective organisms. They were included to clarify the previously unlabeled clades of transcripts isolated from the blast analyses. In total, 1331 transcripts identified from the blastp analysis of the five query proteins were used to create a protein phylogeny of the short-chain dehydrogenase/reductase protein family. After pruning low quality sequences, the final phylogeny was composed of 322 transcripts, and while the relationships of the tree are complicated, the topology of the resulting phylogeny suggests putative function of lophotrochozoan and ecdysozoan transcripts homologous to that of described vertebrate retinol dehydrogenases involved in the visual (Supplemental Figure 3-S3). NOTUNG reconciliation reports 169 duplications and 1273 losses.

Unlike the other visual cycle gene families, there are three large clades of vertebrate retinol dehydrogenases which resolve sister to clades of protostomes (Supplemental Figure 3-S3). Support values at nodes separating these vertebrate clades from other parts of the tree are generally low (BS <50) while support values within these clades tend to be higher (BS>70). One clade is composed of vertebrate RDH8 and RDH10, RDH7 and RDH7-like sequences from *Mizuhopecten yessoensis*, a scallop species, and PDH and RDHB sequences from *Drosophila melanogaster* and *D. busckii*, respectively. The described function of these retinol dehydrogenases varies with RDH7, RDH8 and PDH reported to function in the oxidation of all-*trans* retinol to retinal, and RDH10 is reported to function in the reduction of all-*trans* retinal to retinol as described on the UniProt database. This clade resolves sister to the remainder of the

tree, and although it contained transcripts from the protostome transcriptomes prior to the trimming of this tree, the clade is solely anchor sequences after pruning.

A second large clade of vertebrate retinol dehydrogenase sequences contains RDH2, RDH3, RDH5, RDH7, and RDH16 (BS: 56; Supplemental Figure 3-S3). Based on descriptions from the UniProt database these proteins serve the same function in the oxidation of isomers of retinol to retinal and several of these are synonyms for the same proteins (RDH2/16 and RDH3/7, for instance) (Su et al. 1998; X. Chai et al. 1995). This clade resolved sister to a large clade composed of molluscan transcripts with many transcripts from individual species. Transcripts in this superclade were identified as a combination of RDH2 and RDH7-like sequences when blasted against the NCBI GenBank database. The topology of this clade suggests a putative function of these potential transcripts in the oxidation of retinol to retinal.

A third superclade is divided into a vertebrate clade of RDH11 and RDH12 includes four arthropod transcripts (Supplemental Figure 3-S3, BS: 33 to 100) and a second clade of vertebrate RDH13 sequences sister to protostome transcripts (BS: 38). Tested RDH11 functions in oxidation of retinol to retinal (Haeseleer et al. 2002) while RDH13 functions in the reduction of retinal to retinol (Belyaeva et al. 2008). Subsequent blasting of transcripts within this clade identifies these transcripts as RDH11-like or RDH12-like. The node supports vary greatly throughout this clade but are generally low (BS ranges 9 to 85). The relationship of these protostome transcripts to the vertebrate anchor sequences suggests a putative function similar to that of RDH13 in the reduction of retinal to retinol.

Finally, there is one large clade composed of molluscan transcripts sister to three vertebrate RDH14 sequences. RDH14 is described in the function in the oxidation of retinol to retinal isomers within vertebrates (Haeseleer et al. 2002). Upon blasting molluscan transcripts from this clade against NCBI database, they are described as RDH12 or RDH13-like. The node support between the vertebrate and mollusc clades is low (BS: 37) with node supports ranging drastically within the molluscan clade.

Discussion

In this study, we attempted to identify and expand gene membership of the molluscan retinoid visual cycle. We analyzed tissue-specific transcriptomes under a phylogenetic framework to better infer homology and putative function. This study represents the first wide-use of readily available molluscan transcriptomes to examine this understudied visual pathway across the phylum of Mollusca. From our searches, we were able to identify transcripts that formed clades with proteins from the vertebrate and insect visual cycles suggesting putative homology. These hypotheses that can be tested in future experimental studies that examine function of these proteins. The results of our study show evidence of a dark regeneration visual cycle in molluscs, a pathway previously undescribed. Here, we propose a new hypothesis for the retinoid visual cycle in molluscs to act as a starting point for future research into this pathway (Figure 3-3).

The retinoid visual cycle functions to reset the photopigments allowing them to function given exposure to light (Kusakabe et al. 2009) as well as to manage the cytotoxicity inherent to retinal isomers involved in the visual cycle pathway (Gonzalez-Fernandez 2002). In our analyses, we were unable to identify transcripts homologous to either lecithin: retinol acyltransferase (LRAT) or cellular retinol binding protein (CRBP).

In vertebrates, these two proteins function in tandem to create the retinyl esters used as storage molecules to prevent the toxicity inherent with excess retinoids (MacDonald and Ong 1988; Ong, MacDonald, and Gubitosi 1988; Saari, Bredberg, and Farrell 1993). Our findings support previous claims that the LRAT/CRBP storage pathway is a novel vertebrate pathway (Albalat 2009); but, despite the absence of these transcripts, the toxicity of the retinoids remains an issue in the photoreceptor cells of molluscs. Retinochrome has been hypothesized to serve as a storage molecule for retinal isomers given its inherent ability to isomerize retinal isomers 11-*cis* and 13-*cis* back to all-*trans* during dark incubation (Ozaki et al. 1983). Without evidence for a secondary pathway, we hypothesize that retinochrome may function as a storage molecule in place of the vertebrate retinyl storage pathway while acting as a photoisomerase only when the demand for 11-*cis* retinal is high and in the presence of light.

Our results support the presence of the previously described RALBP shuttle protein in molluscs. Retinaldehyde binding protein (RALBP) acts in the transport of all-*trans* and 11-*cis* retinal between rhodopsin of the outer segment, and retinochrome of the inner segment of the photoreceptor cells in squid (Terakita, Hara, and Hara 1989). RALBP has only been reported in the squid photoreceptor cells; however, our study suggests the presence of RALBP throughout molluscs. A second superclade within the CRAL-TRIO protein family tree is composed of lophotrochozoan and vertebrate CRALBP anchor sequences, included to suggest putative function. The relationship of these clades has high support and suggests putative function of these lophotrochozoan transcripts may be similar to that of vertebrate CRALBP which would mean the presence of a second CRAL-TRIO shuttle protein which functions in the photoreceptor

cells. Aspects of the visual cycle occurs in the retinal pigment epithelium in vertebrates (Lamb and Pugh 2004) and within the retinal pigment cells of insects (T. Wang and Montell 2005). The rhodopsin-retinochrome cycle occurs within the photoreceptor cells, as described in squid (Hara and Hara 1991; Terakita, Hara, and Hara 1989), but based on our findings of homologous vertebrate and insect dehydrogenases, the visual cycle of molluscs may occur in elsewhere in the molluscan retina, such as the pigment cells similar to vertebrate and insects requiring retinal isomers to be shuttled outside the photoreceptor cells.

In *Drosophila*, PINTA has been shown to shuttle 11-*cis* and all-*trans* retinol within the retinal pigment cells for retinoid synthesis (T. Wang, Jiao, and Montell 2007). We used PINTA as a query protein for our transcriptome searches, and our results hint at the presence of a PINTA-like protein within Lophotrochozoa. The presence of a homologous PINTA-like protein in molluscs offers another line of evidence that may suggest visual cycle activity outside of the photoreceptors of molluscs. RALBP delivers retinal products to and from retinochrome and rhodopsin, but it is still unclear if or how the retinal isomers leave or enter the photoreceptor cells in other visual cycle processes. Interestingly, genomic and transcriptome searches of PINTA in lepidopteran species were unable identify a PINTA ortholog (Macias-Muñoz, Rangel Olguin, and Briscoe 2019; Macias-Muñoz, McCulloch, and Briscoe 2017). Instead, Lepidopteran CTD31, also a CRAL-TRIO shuttle protein, was proposed to serve a similar function to PINTA, delivering retinal isomers between an unidentified retinochrome-like opsin in pigment cells and visual opsins of the photoreceptors. A CTD31 query protein was not used in this study; however, if PINTA is a *Drosophila* specific protein, it could explain

the absence of any other arthropod transcripts within the PINTA-like clade. Functional assays and localization studies of PINTA or possible CTD31-like transcripts could help us understand the purpose of these hypothesized shuttle proteins and their role in the molluscan visual cycle as well as the presence of visual cycle machinery outside the photoreceptor cells of molluscs.

Reconstruction of the carotenoid oxygenase protein family did not reveal the presence of a homologous RPE65 protein in protostomes but revealed a possible duplication event of BCOI unique to protostomes. The topology of carotenoid oxygenase tree suggests a duplication event of BCOI following the protostome-deuterostome split with this duplication leading to the neo-functionalization of one product in vertebrates giving rise to RPE65, the isomerohydrolase present in the vertebrate visual cycle. Another duplication of BCOI may have occurred prior to the Lophotrochozoa-Ecdysozoa lineage in which Ecdysozoa lost one of duplication products while Lophotrochozoa seems to have kept both. Transcripts within these lophotrochozoan clades were identified as BCOI-like and retinoid isomerohydrolase-like when blasted against the NCBI database. This topology could reflect that ninaB serves in both an isomerohydrolase and metabolic capacity in molluscs in a homologous manner to its activity in *Drosophila*. NinaB possesses both vitamin A cleavage and retinoid isomerase properties (Oberhauser et al. 2008) and is conserved across distantly related arthropods (C. Chai et al. 2019); however, these studies in insects showed that it is expressed in the neural tissues of flies and functions in the metabolism of retinoids outside the retina. Based on this evidence, it is highly unlikely to serve in the pivotal step of isomerization of all-*trans* to 11-*cis* retinol in molluscs either.

The dehydrogenases described in vertebrate and insect visual cycles serve two roles: to reduce all-*trans* retinal to all-*trans* retinol and oxidize 11-*cis* retinol to 11-*cis* retinal. It is difficult to assume function of such closely related proteins from our analysis when functional assays are required, but our results suggest the presence of multiple dehydrogenases within molluscs. The protein phylogeny of the short-chain dehydrogenase/reductase family show the presence of possible RDH2, RDH3, RDH5, RDH13, RDH14, and/or RDH16 homologs in molluscs. Retinol dehydrogenases within vertebrates are known to have overlapping functions, likely resulting from large gene duplications in which the dehydrogenases can compensate for deficiencies caused by mutation (Maeda et al. 2007) or in which dehydrogenases serve the same function but are localized to different tissues or subcellular structures (Belyaeva et al. 2008). For example, RDH2, RDH3, RDH5, RDH7, RDH14 and RDH16 oxidize retinal isomers, although the isomer preference varies by protein although some have showed enzymatic activity for multiple isomers (Jang et al. 2001; X. Chai et al. 1995; Su et al. 1998). From our results, the relationship of the clades as described may suggest the putative function of transcripts in the reduction of retinol to retinal in molluscs but mixed node support throughout these clades (BS: 0 to 100) do not confidently support this claim.

The short-chain dehydrogenase/reductase protein phylogeny also suggests the presence of a reducing dehydrogenase homologous to RDH13. In the vertebrate retinoid visual cycle, RDH8 and RDH12 function in photoreceptor cells to reduce all-*trans* retinal released from the photopigment to all-*trans* retinol (Maeda et al. 2007) and PDH serves the same function in reduction in insects (X. Wang et al. 2010). RDH13 is

part of the RDH11-like retinol dehydrogenases. Though the physiological function of RDH13 is yet unclear, it was shown to reduce retinal within the photoreceptor cells of humans (Belyaeva et al. 2008). As described, vertebrate RDH13 resolved in a clade with protostome sequences. This relationship may imply the presence of an RDH13-like retinol dehydrogenase in molluscs, but the low node support also does not allow us to confirm this relationship without functional testing. Vertebrate RDH8 sequences resolved as an isolated clade suggesting no presence of RDH8 in protostomes. Predicted RDH12-like sequences from the bivalve species *Mizuhopecten yessoensis* and *Crassostrea ariakensis* were included as an anchor sequence within this study in addition to vertebrate sequences; however, both resolved outside the clade containing the confirmed vertebrate sequences. The bivalve RDH12 predicted sequences used resolved in clades of lophotrochozoans, sister to a clade of *Drosophila* sequences including PDH and RDHB, but these sequences are likely labeled based on annotation software used to analyze large transcriptomes and genomes. Without functional studies or a verified sequence resolving within the same clade as the bivalve RDH12-like sequences to suggest putative function, it is difficult to assign an appropriate name to these sequences. Another important note is that within the short-chain dehydrogenase/reductase phylogeny, there are many clades composed solely of protostome sequences. The tree is not resolved enough to hypothesize in depth evolutionary relationships, but the presence of isolated protostome clades may suggest the presence of retinol dehydrogenases or a close ancestor of these proteins unique to these animals.

A major caveat to the interpretations of our results is our reliance on transcriptomes, in large part because the absence of a gene of interest does not necessarily mean that the gene is absent in a given species. Transcriptomes represent biological snapshots of an organism's expressed genes at with specific spatial and temporal dimensions. Even if the transcriptome is complete by quantitative standards, expression levels of genes vary by life stage, activity, and tissue. To combat some of these issues, we were particular about the organisms selected as well as our focus on transcriptomes from light sensitive tissue. However, due to the scope of the study and some of the organisms used, many transcriptomes were taken from the whole animals without control over what stage in life cycle or environmental conditions samples were gathered. Thus, we have tried to make general conclusions regarding the proteins identified and discussed and emphasize points for future research. From the findings we present here, we propose a novel light independent retinoid visual cycle present in molluscs, but more focused transcriptomic and functional studies will be required to verify these hypotheses.

Proposed retinoid visual cycle in molluscs

Here, we propose a novel molecular pathway of a light independent retinoid visual cycle in molluscs based on the results described from this study and published literature (Figure 3-3). Molluscs possess both monostable and bistable photopigments, similar to vertebrates and insects, respectively (Kojima et al. 1997). A light independent visual cycle may function for the turnover of retinal caused by the release from degraded bistable photopigments or the processing of metabolic products as is the case in insects (X. Wang et al. 2010, 2012; Montell 2012) or in the preparation of monostable

photopigments for phototransduction as in vertebrates (Lamb and Pugh 2004). In our proposed molluscan visual cycle model, a molluscan RDH13-like enzyme reduces the all-*trans* retinal to an all-*trans* retinol. The activity of a molluscan RDH13-like dehydrogenase may occur in the photoreceptor cells similar to humans, but the localization of this activity in molluscs is unknown. Our results lend evidence to further visual cycle activity outside the photoreceptor cells, likely adjacent cells such as pigment cells. This activity outside the photoreceptor cells is similar to what is described in vertebrates and insects, thus the resulting all-*trans* retinol requires a shuttle protein for transport from the photoreceptor cells to the pigment cells. In squid, RALBP has been shown to shuttle retinoid isomers within the photoreceptor cells (Hara and Hara 1991) but no evidence suggests activity outside photoreceptor cells. Inversely, in *Drosophila*, while PINTA preferentially binds all-*trans* retinol, it has been localized to pigment cells but shown little evidence of function in the photoreceptor cells (T. Wang and Montell 2005). We were unable to identify a homologous sequence to IRBP, the intercellular shuttle protein in vertebrates, but given findings of potentially homologous dehydrogenases and PINTA in molluscs, we propose that continuation of the visual cycle in molluscs extends outside the photoreceptor cells.

Once in the adjacent cells of the photoreceptors, such as the pigment cells, the all-*trans* retinol is bound by a PINTA-like shuttle protein to an isomerase or isomerohydrolase which has yet to be identified. In vertebrates, there is an extra step involving a large protein complex composed of LRAT, CRBP, RPE65, and RDH5, wherein CRBP, a lipocalin shuttle protein, delivers the retinol to LRAT allowing the conversion of the all-*trans* retinol to a retinyl. The retinyl is used for storage of the

retinoid in a non-cytotoxic, stable state (Travis et al. 2007; MacDonald and Ong 1988; Ong, MacDonald, and Gubitosi 1988; Saari, Bredberg, and Farrell 1993). Our analysis showed no evidence of LRAT or CRBP in protostomes and supports previous hypotheses that LRAT and CRBP mechanisms are a novel pathway unique to vertebrates (Albalat 2012). We were also unable to identify an RPE65 homolog in protostomes. However, both the mollusc pathway being described here and insects require an isomerase to isomerize the all-*trans* retinol to an 11-*cis* retinol which has yet to be identified in either (Montell 2012). Following isomerization from all-*trans* to 11-*cis* retinol, the retinol is reduced by one of the many reducing dehydrogenase candidates identified in our studied. Vertebrates use RDH5 in this step within the retinal pigment epithelium of the retina (Strauss 2005) though RDH2, RDH3, RDH7 and RDH14 have been shown to oxidize isomers of retinol to retinal (Belyaeva and Kedishvili 2002; Laudet, Schubert, and Albalat 2011; Su et al. 1998), and RDHB oxidizes retinol to retinal in *Drosophila* (X. Wang et al. 2012). While our results did not show the presence of an RDHB homolog, one or a combination of RDH2/16, RDH3/7, RDH5, and RDH14 homologs are suggested to oxidize 11-*cis* retinol to an 11-*cis* retinal within the pigment cells. The newly formed 11-*cis* retinal chromophore is then shuttled back to the photoreceptor cells by an unknown shuttle protein. Once within the photoreceptors is then delivered back to the rhodopsin, resetting the pigment and allowing a signaling cascade to begin anew. In this model, the 11-*cis* retinal from this dark regeneration cycle may be shuttled to the opsin by RALBP once in the photoreceptor cells; however, our study does indicate the presence of RALBP in molluscs outside of squid for the first time.

We were unable to identify a retinyl storage pathway in molluscs similar to that present in vertebrates. To date, neither insects nor molluscs have a described mechanism for storage of retinoid isomers. Retinochrome was suggested to serve as a storage molecule of retinal in squid given the equal expression level and inherent ability of retinochrome to revert all-*trans* retinal to 11-*cis* retinal in long periods darkness (Ozaki et al. 1983). We suspect that retinochrome does act as the storage molecule of retinal in the molluscan visual cycle. Following isomerization of the 11-*cis* retinal to all-*trans* retinal, the rhodopsin interacts with RALBP releasing the all-*trans* retinal. The RALBP delivers the retinal to retinochrome where, in the presence of light, retinochrome absorbs a photon of light and isomerizes the all-*trans* retinal to an 11-*cis* retinal. When in high demand for functional retinal, RALBP can shuttle the 11-*cis* retinal back to rhodopsin to reset the photopigment. Otherwise, the 11-*cis* retinal is retained by the retinochrome and can be reconverted back to all-*trans* retinal given sufficient time in the dark.

Conclusion

In this study, we assembled and queried transcriptomes from available data across molluscs in order to begin inventorying the presence or absence of homologous vertebrate and insect visual cycle proteins. From our findings, we propose a new model for the retinoid visual cycle in molluscs. There appears to be a degree of homology between molluscan transcripts with the vertebrate and arthropod retinoid visual cycles, and while the results presented here serve as a starting point, future studies are required to fully understand the functions of the transcripts identified. There are many questions left unanswered by the new proposed model. There is no work showing the

activity of pigment cells in molluscs on a molecular level, making our hypothesis as to their action based purely on our findings here and the similarities to that of the work in arthropods. Assuming our prediction regarding molluscan pigment cells is accurate, without identification of an IRBP-like shuttle protein, it is unclear how the all-*trans* retinal exits or how the 11-*cis* retinal enters the photoreceptor cells.

Due to the scope of this study, these findings are reliant on the quality of the transcriptomes assembled and searched. We used as many available ocular or photosensitive tissues as possible, though several transcriptomes were sequenced from whole animals or likely un-fasted specimens leading to possible contaminations and dilution in the quality of transcripts. A more targeted approach of transcriptomic sequencing of molluscan ocular tissues may help focus the available transcripts and clarify some of the phylogenetic relationships presented in the gene trees here. This study is intended to act as a strong background and stepping-stone towards understanding the retinoid visual cycle in these understudied organisms. Functional assays of the transcripts identified here may enlighten as to some of the processes proposed in this study, while localization studies will be necessary to complete the picture as to whether a dark regeneration cycle exists within molluscs and the details of its functions.

References

Albalat, Ricard. 2009. “The Retinoic Acid Machinery in Invertebrates: Ancestral Elements and Vertebrate Innovations.”

———. 2012. “Evolution of the Genetic Machinery of the Visual Cycle : A Novelty of the Vertebrate Eye ?” 29 (5): 1461–69. https://doi.org/10.1093/molbev/msr313.

Altschul, Stephen F., Warren Gish, Webb Miller, Eugene W. Myers, and David J. Lipman. 1990. “Basic Local Alignment Search Tool.” *Journal of Molecular Biology* 215 (3): 403–10. https://doi.org/10.1016/S0022-2836(05)80360-2.

Arshavsky, Vadim Y. 2010. “Vision: The Retinoid Cycle in Drosophila.” *Current Biology* 20 (3): R96–98. https://doi.org/10.1016/j.cub.2009.12.039.

Belyaeva, Olga V., and Natalia Y. Kedishvili. 2002. “Human Pancreas Protein 2 (PAN2) Has a Retinal Reductase Activity and Is Ubiquitously Expressed in Human Tissues.” *FEBS Letters* 531 (3): 489–93. https://doi.org/10.1016/S0014-5793(02)03588-3.

Belyaeva, Olga V., Olga V. Korkina, Anton V. Stetsenko, and Natalia Y. Kedishvili. 2008. “Human Retinol Dehydrogenase 13 (RDH13) Is a Mitochondrial Short-Chain Dehydrogenase/Reductase with a Retinaldehyde Reductase Activity.” *FEBS Journal* 275 (1): 138–47. https://doi.org/10.1111/j.1742-4658.2007.06184.x.

Borner, Janus, Peter Rehm, Ralph O. Schill, Ingo Ebersberger, and Thorsten Burmester. 2014. “A Transcriptome Approach to Ecdysozoan Phylogeny.” *Molecular Phylogenetics and Evolution* 80 (1): 79–87. https://doi.org/10.1016/j.ympev.2014.08.001.

Capella-Gutiérrez, Salvador, José M. Silla-Martínez, and Toni Gabaldón. 2009. “TrimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses.” *Bioinformatics* 25 (15): 1972–73. https://doi.org/10.1093/bioinformatics/btp348.

Chai, Chunli, Xin Xu, Weizhong Sun, Fang Zhang, Chuan Ye, Guangshu Ding, Jiantao Li, et al. 2019. “Characterization of the Novel Role of NinaB Orthologs from Bombyx Mori and Tribolium Castaneum.” *Insect Biochemistry and Molecular Biology* 109 (December 2018): 106–15. https://doi.org/10.1016/j.ibmb.2019.03.004.

Chai, Xiyun, Yan Zhai, Gabriela Popescu, and Joseph L. Napoli. 1995. “Cloning of a CDNA for a Second Retinol Dehydrogenase Type II. Expression of Its MRNA Relative to Type I.” *Journal of Biological Chemistry* 270 (47): 28408–12. https://doi.org/10.1074/jbc.270.47.28408.

Chen, Kevin, Dannie Durand, and Martin Farach-Colton. 2000. “NOTUNG: A Program for Dating Gene Duplications and Optimizing Gene Family Trees.” *Journal of Computational Biology* 7 (3–4): 429–47.

Chiari, Ylenia, Vincent Cahais, Nicolas Galtier, and Frédéric Delsuc. 2012. “Phylogenomic Analyses Support the Position of Turtles as the Sister Group of Birds and Crocodiles (Archosauria).” *BMC Biology* 10 (July). https://doi.org/10.1186/1741-7007-10-65.

Darby, Charlotte A., Maureen Stolzer, Patrick J. Ropp, Daniel Barker, and Dannie Durand. 2016. “Xenolog Classification.” *Bioinformatics* 33 (December 2016): btw686. https://doi.org/10.1093/bioinformatics/btw686.

Durand, Dannie, Bjarni V. Halldórsson, and Benjamin Vernot. 2006. “A Hybrid Micro-Macroevolutionary Approach to Gene Tree Reconstruction.” *Journal of Computational Biology* 13 (2): 320–35. https://doi.org/10.1089/cmb.2006.13.320.

Feuda, Roberto, Sinead C. Hamilton, James O. McInerney, and Davide Pisani. 2012. “Metazoan Opsin Evolution Reveals a Simple Route to Animal Vision.” *Proceedings of the National Academy of Sciences* 109 (46): 18868–72. https://doi.org/10.1073/pnas.1305990110.

Findlay, John B.C., and Darryl J.C. Pappin. 1986. “The Opsin Family of Proteins.” *The Biochemical Journal* 238 (3): 625–42. https://doi.org/10.1042/bj2380625.

Fu, Limin, Beifang Niu, Zhengwei Zhu, Sitao Wu, and Weizhong Li. 2012. “CD-HIT: Accelerated for Clustering the next-Generation Sequencing Data.” *Bioinformatics* 28 (23): 3150–52. https://doi.org/10.1093/bioinformatics/bts565.

Geniza, Matthew, and Pankaj Jaiswal. 2017. “Tools for Building de Novo Transcriptome Assembly.” *Current Plant Biology* 11–12 (December): 41–45. https://doi.org/10.1016/j.cpb.2017.12.004.

Gollapalli, Deviprasad R., and Robert R. Rando. 2003. “All-Trans-Retinyl Esters Are the Substrates for Isomerization in the Vertebrate Visual Cycle.” *Biochemistry* 42 (19): 5809–18. https://doi.org/10.1021/bi0341004.

Gonzalez-Fernandez, Federico. 2002. “Evolution of the Visual Cycle: The Role of Retinoid-Binding Proteins.” *Journal of Endocrinology* 175: 75–88. https://doi.org/10.1677/joe.0.1750075.

González, Vanessa L., Sénia C.S. Andrade, Rüdiger Bieler, Timothy M. Collins, Casey W. Dunn, Paula M. Mikkelsen, John D. Taylor, and Gonzalo Giribet. 2015. “A Phylogenetic Backbone for Bivalvia: An RNA-Seq Approach.” *Proceedings of the Royal Society B: Biological Sciences* 282 (1801): 1–9. https://doi.org/10.1098/rspb.2014.2332.

Goodheart, Jessica A., Adam L. Bazinet, Ángel Valdés, Allen G. Collins, and Michael P. Cummings. 2017. “Prey Preference Follows Phylogeny: Evolutionary Dietary Patterns within the Marine Gastropod Group Cladobranchia (Gastropoda: Heterobranchia: Nudibranchia).” *BMC Evolutionary Biology* 17 (1): 1–14. https://doi.org/10.1186/s12862-017-1066-0.

Grabherr, Manfred G, Brian J Haas, Moran Yassour, Joshua Z Levin, Dawn A Thompson, Ido Amit, Xian Adiconis, et al. 2011. “Full-length Transcriptome Assembly from RNA-Seq Data without a Reference Genome.” *Nature Biotechnology* 29: 644–52. https://doi.org/10.1038/nbt.1883.

Haas, Brian J., Alexie Papanicolaou, Moran Yassour, Manfred Grabherr, Philip D. Blood, Joshua Bowden, Matthew Brian Couger, et al. 2013. “De Novo Transcript Sequence Reconstruction from RNA-Seq Using the Trinity Platform for Reference Generation and Analysis.” *Nature Protocols* 8 (8): 1494–1512. https://doi.org/10.1038/nprot.2013.084.

Haeseleer, Françoise, Geeng Fu Jang, Yoshikazu Imanishi, Carola A.G.G. Driessen, Masazumi Matsumura, Peter S. Nelson, and Krzysztof Palczewski. 2002. “Dual-Substrate Specificity Short Chain Retinol Dehydrogenases from the Vertebrate Retina.” *Journal of Biological Chemistry* 277 (47): 45537–46. https://doi.org/10.1074/jbc.M208882200.

Hamdorf, Kurt, and Joachim Schwemer. 1975. “Photoregeneration and the Adaptation Process in Insect Photoreceptors.” *Photoreceptor Optics*, no. 3: 263–89. https://doi.org/10.1007/978-3-642-80934-7_16.

Hara, Tomiyuki, and Reiko Hara. 1972. “Cephalopod Retinochrome.” *Photochemistry of Vision* 7 (1): 720–46. https://doi.org/10.1016/S0076-6879(82)81031-8.

———. 1976. “Distribution of Rhodopsin and Retinochrome in the Squid Retina.” *Journal of General Physiology* 67 (6): 791–805. https://doi.org/10.1085/jgp.67.6.791.

———. 1991. “Retinal-Binding Protein: Function in a Chromphore Exchange System in the Squid Visual Cell.”

Hillman, Peter, Shaul Hochstein, and Baruch Minke. 1983. “Transduction in Invertebrate Photoreceptors: Role of Pigment Bistability.” *Physiological Reviews* 63 (2): 668–772. https://doi.org/10.1152/physrev.1983.63.2.668.

Jang, Geeng Fu, J. Preston Van Hooser, Vladimir Kuksa, Joshua K. McBee, Yu Guang He, Jacques J.M. Janssen, Carola A.G.G. Driessen, and Krzysztof Palczewski. 2001. “Characterization of a Dehydrogenase Activity Responsible for Oxidation of 11-Cis-Retinol in the Retinal Pigment Epithelium of Mice with a Disrupted RDH5 Gene: A Model for the Human Hereditary Disease Fundus Albipunctatus.” *Journal of Biological Chemistry* 276 (35): 32456–65. https://doi.org/10.1074/jbc.M104949200.

Jin, Minghao, Songhua Li, Walid N. Moghrabi, Hui Sun, and Gabriel H. Travis. 2005. “Rpe65 Is the Retinoid Isomerase in Bovine Retinal Pigment Epithelium.” *Cell* 122 (3): 449–59. https://doi.org/10.1016/j.cell.2005.06.042.

Katoh, Kazutaka, and Daron M. Standley. 2013. “MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability.” *Molecular Biology and Evolution* 30 (4): 772–80. https://doi.org/10.1093/molbev/mst010.

Kearse, Matthew, Richard Moir, Amy Wilson, Steven Stones-Havas, Matthew Cheung, Shane Sturrock, Simon Buxton, et al. 2012. “Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data.” *Bioinformatics* 28 (12): 1647–49. https://doi.org/10.1093/bioinformatics/bts199.

Kocot, Kevin M., Torsten H. Struck, Julia Merkel, Damien S. Waits, Christiane Todt, Pamela M. Brannock, David A. Weese, et al. 2017. “Phylogenomics of Lophotrochozoa with Consideration of Systematic Error.” *Systematic Biology* 66 (2): 256–82. https://doi.org/10.1093/sysbio/syw079.

Kocot, Kevin M., Christiane Todt, Nina T. Mikkelsen, and Kenneth M. Halanych. 2019. “Phylogenomics of Aplacophora (Mollusca, Aculifera) and a Solenogaster without a Foot.” *Proceedings of the Royal Society B: Biological Sciences* 286 (1902). https://doi.org/10.1098/rspb.2019.0115.

Kocot, Kevin M, Johanna T Cannon, Christiane Todt, Mathew R Citarella, Andrea B Kohn, Achim Meyer, Scott R Santos, et al. 2011. “Phylogenomics Reveals Deep Molluscan Relationships.” *Nature* 477 (7365): 452–56. https://doi.org/10.1038/nature10382.

Kojima, Daisuke, Akihisa Terakita, Toru Ishikawa, Yasuo Tsukahara, Akio Maeda, and Yoshinori Shichida. 1997. “A Novel G(o)-Mediated Phototransduction Cascade in Scallop Visual Cells.” *Journal of Biological Chemistry* 272 (37): 22979–82. https://doi.org/10.1074/jbc.272.37.22979.

Kusakabe, Takehiro G., Noriko Takimoto, Minghao Jin, and Motoyuki Tsuda. 2009. “Evolution and the Origin of the Visual Retinoid Cycle in Vertebrates.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1531): 2897–2910. https://doi.org/10.1098/rstb.2009.0043.

Lamb, Trevor D., and Edward N. Pugh. 2004. “Dark Adaptation and the Retinoid Cycle of Vision.” *Progress in Retinal and Eye Research* 23 (3): 307–80. https://doi.org/10.1016/j.preteyeres.2004.03.001.

Langmead, Ben, and Steven L. Salzberg. 2012. “Fast Gapped-Read Alignment with Bowtie 2.” *Nature Methods* 9 (4): 357–59. https://doi.org/10.1038/nmeth.1923.

Laudet, Vincent, Michael Schubert, and Ricard Albalat. 2011. “Evolution of Retinoid and Steroid Signaling : Vertebrate” 3: 985–1005. https://doi.org/10.1093/gbe/evr084.

Le, Si Quang, and Olivier Gascuel. 2008. “An Improved General Amino Acid Replacement Matrix.” *Molecular Biology and Evolution* 25 (7): 1307–20. https://doi.org/10.1093/molbev/msn067.

Li, Bo, and Colin N. Dewey. 2014. “RSEM: Accurate Transcript Quantification from RNA-Seq Data with or without a Reference Genome.” *Bioinformatics: The Impact of Accurate Quantification on Proteomic and Genetic Analysis and Research*, 41–74. https://doi.org/10.1201/b16589.

Li, Weizhong, and Adam Godzik. 2006. “Cd-Hit: A Fast Program for Clustering and Comparing Large Sets of Protein or Nucleotide Sequences.” *Bioinformatics* 22 (13): 1658–59. https://doi.org/10.1093/bioinformatics/btl158.

Lindgren, Annie R., Molly S. Pankey, Frederick G. Hochberg, and Todd H. Oakley. 2012. “A Multi-Gene Phylogeny of Cephalopoda Supports Convergent Morphological Evolution in Association with Multiple Habitat Shifts in the Marine Environment.” *BMC Evolutionary Biology* 12 (1). https://doi.org/10.1186/1471-2148-12-129.

MacDonald, Paul N., and David E. Ong. 1988. “A Lecithin: Retinol Acyltransferase Activity in Human and Rat Liver.” *Biochemical and Biophysical Research Communications* 156 (1): 157163.

Macias-Muñoz, Aide, Kyle J. McCulloch, and Adriana D. Briscoe. 2017. “Copy Number Variation and Expression Analysis Reveals a Nonorthologous Pinta Gene Family Member Involved in Butterfly Vision.” *Genome Biology and Evolution* 9 (12): 3398–3412. https://doi.org/10.1093/gbe/evx230.

Macias-Muñoz, Aide, Aline G. Rangel Olguin, and Adriana D. Briscoe. 2019. “Evolution of Phototransduction Genes in Lepidoptera.” *Genome Biology and Evolution* 11 (8): 2107–24. https://doi.org/10.1093/gbe/evz150.

Maeda, Akiko, Tadao Maeda, Wenyu Sun, Houbin Zhang, Wolfgang Baehr, and Krzysztof Palczewski. 2007. “Redundant and Unique Roles of Retinol Dehydrogenases in the Mouse Retina.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (49): 19565–70. https://doi.org/10.1073/pnas.0707477104.

Mata, Nathan L., Walid N. Moghrabi, Jung S. Lee, Tam V. Bui, Roxana A. Radu, Joseph Horwitz, and Gabriel H. Travis. 2004. “Rpe65 Is a Retinyl Ester Binding Protein That Presents Insoluble Substrate to the Isomerase in Retinal Pigment Epithelial Cells.” *Journal of Biological Chemistry* 279 (1): 635–43. https://doi.org/10.1074/jbc.M310042200.

Miller, Mark A., Wayne Pfeiffer, and Terri Schwartz. 2010. “Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees.” *2010 Gateway Computing Environments Workshop, GCE 2010*, 1–8. https://doi.org/10.1109/GCE.2010.5676129.

Montell, Craig. 2012. “Drosophila Visual Transduction.” *Trends in Neurosciences* 35 (6): 356–63. https://doi.org/10.1016/j.tins.2012.03.004.

Nguyen, Lam Tung, Heiko A. Schmidt, Arndt Von Haeseler, and Bui Quang Minh. 2015. “IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies.” *Molecular Biology and Evolution* 32 (1): 268–74. https://doi.org/10.1093/molbev/msu300.

Oberhauser, Vitus, Olaf Voolstra, Annette Bangert, Johannes von Lintig, and Klaus Vogt. 2008. “NinaB Combines Carotenoid Oxygenase and Retinoid Isomerase Activity in a Single Polypeptide,” 1–6. papers2://publication/uuid/08F13F7C-C545-4E3D-9C8E-89031DB3AC48.

Ong, David E., Paul N. MacDonald, and Angela M. Gubitosi. 1988. “Esterification of Retinol in Rat Liver. Possible Participation by Cellular Retinol-Binding Protein and Cellular Retinol-Binding Protein II.” *Journal of Biological Chemistry* 263 (12): 5789–96.

Ozaki, Koichi, Reiko Hara, Tomiyuki Hara, and Toshiaki Kakitani. 1983. “Squid Retinochrome. Configurational Changes of the Retinal Chromophore.” *Biophysical Journal* 44 (1): 127–37. https://doi.org/10.1016/S0006-3495(83)84285-4.

Pepe, Isidoro M., and Carlo Cugnoli. 1992. “Retinal Photoisomerase: Role in Invertebrate Visual Cells.” *Journal of Photochemistry and Photobiology. B, Biology* 13 (1): 5–17. http://www.ncbi.nlm.nih.gov/pubmed/1403367.

Ramirez, M. Desmond, Autum N. Pairett, M. Sabrina Pankey, Jeanne M. Serb, Daniel I. Speiser, Andrew J. Swafford, and Todd H. Oakley. 2016. “The Last Common Ancestor of Most Bilaterian Animals Possessed at Least Nine Opsins.” *Genome Biology and Evolution* 8 (12): 3640–52. https://doi.org/10.1093/gbe/evw248.

Saari, John C., D. Lucille Bredberg, and Donald F. Farrell. 1993. “Retinol Esterification in Bovine Retinal Pigment Epithelium: Reversibility of Lecithin:Retinol Acyltransferase.” *Biochemical Journal* 291 (3): 697–700. https://doi.org/10.1042/bj2910697.

Salvini-Plawen, L. von, and Ernst Mayr. 1977. “On the Evolution of Photoreceptors and Eyes.” *Journal of Evolutionary Biology* Vol. 10: 207–63. https://doi.org/10.1007/978-1-4615-6953-4_4.

Seppey, Mathieu, Mosè Manni, and Evgeny M. Zdobnov. 2019. “BUSCO: Assessing Genome Assembly and Annotation Completeness.” *Gene Prediction*, 227–45.

Smith, W. Clay, and Timothy H. Goldsmith. 1991. “The Role of Retinal Photoisomerase in the Visual Cycle of the Honeybee.” *The Journal of General Physiology* 97 (1): 143–65. https://doi.org/10.1085/jgp.97.1.143.

Soubrier, Julien, Mike Steel, Michael S.Y. Lee, Clio Der Sarkissian, Stéphane Guindon, Simon Y.W. Ho, and Alan Cooper. 2012. “The Influence of Rate Heterogeneity among Sites on the Time Dependence of Molecular Rates.” *Molecular Biology and Evolution* 29 (11): 3345–58. https://doi.org/10.1093/molbev/mss140.

Speiser, Daniel I., M. Sabrina Pankey, Alexander K. Zaharoff, Barbara A. Battelle, Heather D. Bracken-Grissom, Jesse W. Breinholt, Seth M. Bybee, et al. 2014. “Using Phylogenetically-Informed Annotation (PIA) to Search for Light-Interacting Genes in Transcriptomes from Non-Model Organisms.” *BMC Bioinformatics* 15 (1): 350. https://doi.org/10.1186/s12859-014-0350-x.

Sperling, Linda, and Ruth Hubbard. 1975. “Squid Retinochrome.” *The Journal of General Physiology* 65 (2): 235–51. https://doi.org/10.1085/jgp.65.2.235.

Stolzer, Maureen, Han Lai, Minli Xu, Deepa Sathaye, Benjamin Vernot, and Dannie Durand. 2012. “Inferring Duplications, Losses, Transfers and Incomplete Lineage Sorting with Nonbinary Species Trees.” *Bioinformatics* 28 (18): 409–15. https://doi.org/10.1093/bioinformatics/bts386.

Strauss, Olaf. 2005. “The Retinal Pigment Epithelium in Visual Function.” *Physiological Reviews* 85: 845–81. https://doi.org/10.1152/physrev.00021.2004.

Su, Jian, Xiyun Chai, Beverly Kahn, and Joseph L. Napoli. 1998. “CDNA Cloning, Tissue Distribution, and Substrate Characteristics of a Cis-Retinol/3α-Hydroxysterol Short-Chain Dehydrogenase Isozyme.” *Journal of Biological Chemistry* 273 (28): 17910–16. https://doi.org/10.1074/jbc.273.28.17910.

Terakita, Akihisa, Reiko Hara, and Tomiyuki Hara. 1989. “Retinal-Binding Protein as a Shuttle for Retinal in the Rhodopsin-Retinochrome System of the Squid Visual Cells.” *Vision Research* 29 (6): 639–52. https://doi.org/10.1016/0042-6989(89)90026-6.

Travis, Gabriel H., Marcin Golczak, Alexander R. Moise, and Krzysztof Palczewski. 2007. “Diseases Caused by Defects in the Visual Cycle: Retinoids as Potential Therapeutic Agents.” *Annual Review of Pharmacology and Toxicology* 47 (1): 469–512. https://doi.org/10.1146/annurev.pharmtox.47.120505.105225.

Vernot, Benjamin, Maureen Stolzer, Aiton Goldman, and Dannie Durand. 2008. “Reconciliation with Non-Binary Species Trees.” *Journal of Computational Biology* 15 (8): 981–1006. https://doi.org/10.1089/cmb.2008.0092.

Wald, George. 1935. “Carotenoids and the Visual Cycle.”

Wang, Tao, Yuchen Jiao, and Craig Montell. 2007. “Dissection of the Pathway Required for Generation of Vitamin A and for Drosophila Phototransduction.” *Journal of Cell Biology* 177 (2): 305–16. https://doi.org/10.1083/jcb.200610081.

Wang, Tao, and Craig Montell. 2005. “Rhodopsin Formation in Drosophila Is Dependent on the PINTA Retinoid-Binding Protein.” *Journal of Neuroscience* 25 (21): 5187–94. https://doi.org/10.1523/JNEUROSCI.0995-05.2005.

Wang, Xiaoyue, Tao Wang, Yuchen Jiao, Johannes von Lintig, and Craig Montell. 2010. “Requirement for an Enzymatic Visual Cycle in Drosophila.” *Current Biology* 20 (2): 93–102. https://doi.org/10.1016/j.cub.2009.12.022.

Wang, Xiaoyue, Tao Wang, Jinfei D. Ni, Johannes von Lintig, and Craig Montell. 2012. “The Drosophila Visual Cycle and De Novo Chromophore Synthesis Depends on RdhB.” *Journal of Neuroscience* 32 (10): 3485–91. https://doi.org/10.1523/JNEUROSCI.5350-11.2012.

Yang, Ziheng. 1995. “A Space-Time Process Model for the Evolution of DNA Sequences.” *Genetics* 139 (2): 993–1005.

Zapata, Felipe, Nerida G. Wilson, Mark Howison, Sónia C.S. Andrade, Katharina M. Jorger, Michael Schrodl, Freya E. Goetz, et al. 2014. “Phylogenomic Analyses of Deep Gastropod Relationships Reject Orthogastropoda.” *Proceedings of the Royal Society B: Biological Sciences*.

Zmasek, Christian M., and Sean R. Eddy. 2001. “ATV: Display and Manipulation of Annotated Phylogenetic Trees.” *Bioinformatics* 17 (4): 383–84. https://doi.org/10.1093/bioinformatics/17.4.383.

Figures and Tables

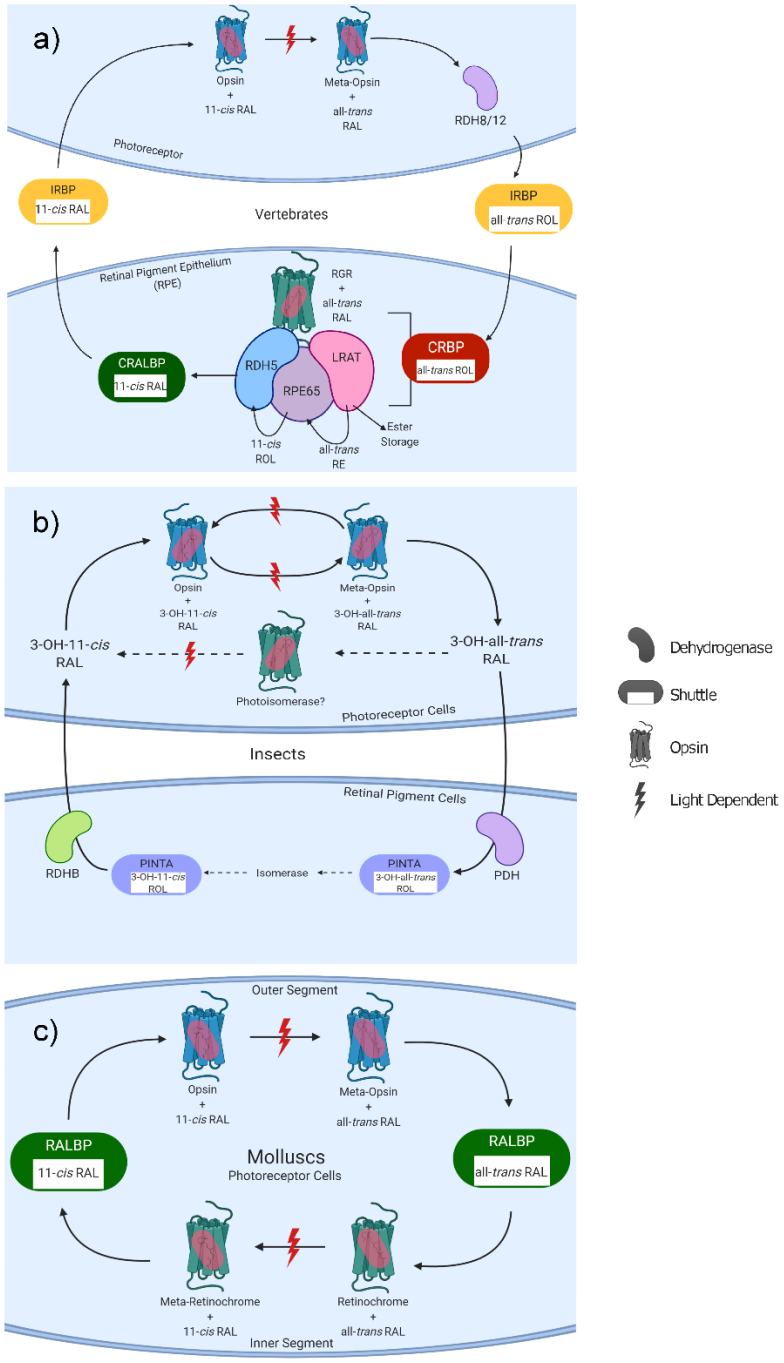


Figure 3-1 - Current understanding of the retinoid visual cycles identified in a) vertebrates, b) insects, and c) molluscs. Primary machinery is shown as uniform shapes consistent with function described in the legend. See text for more detail regarding step-wise pathway. Figures adopted in part from (a: Radu et al, 2008 and Albalat, 2012, b: Arshavsky, 2010, c: Hara and Hara, 1991).

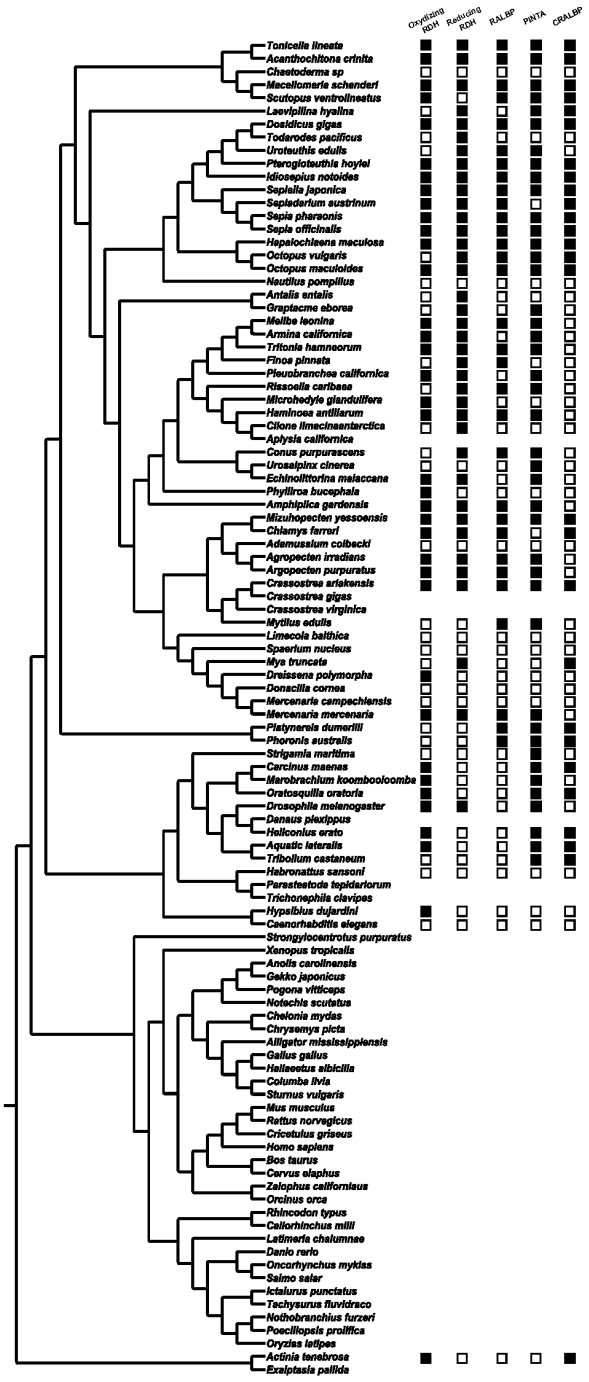


Figure 3-2 - Phylogenetic species tree showing absence and presence of key visual cycle proteins. Species include all protostome species with assembled and searched transcriptomes from this study and all vertebrate species used as anchors for respective gene trees. Absence and presence is represented by an empty or filled black box respectively. Species used for anchor species without transcriptome representation lack boxes.

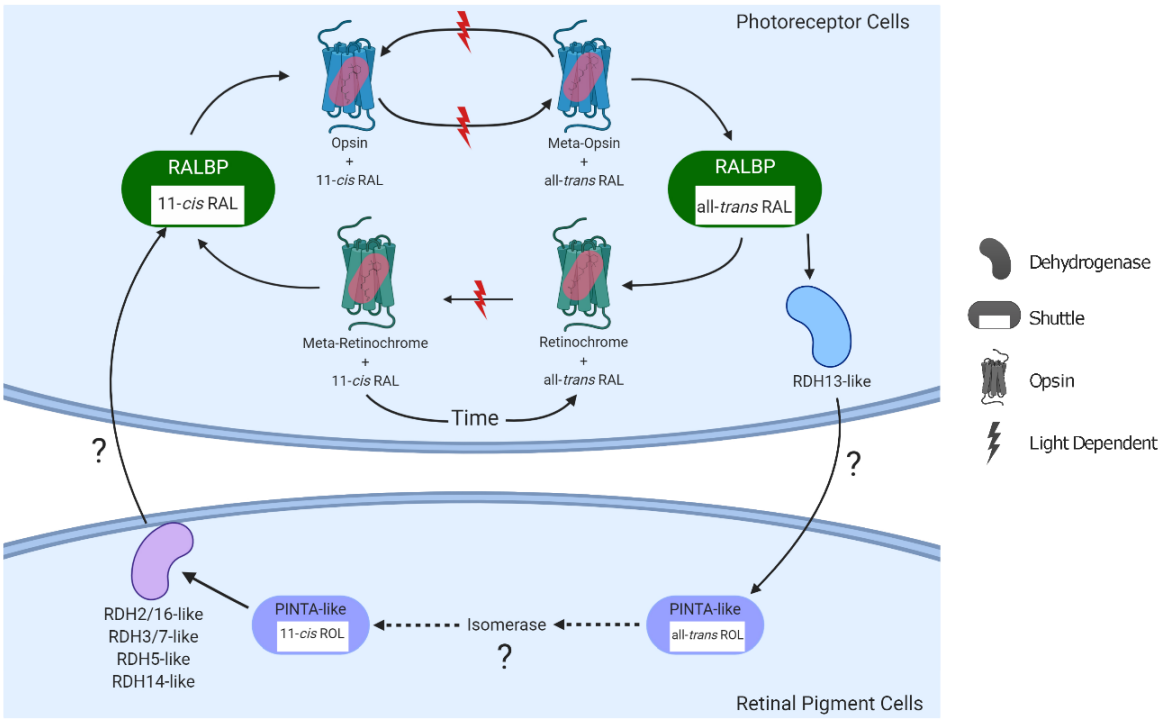


Figure 3-3 - Proposed model of molluscan light-independent visual cycle. Following bleaching of opsin or introduction of all-*trans* retinal, the retinal binds to RALBP shuttle protein and is delivered to a RDH13-like dehydrogenase reducing the all-*trans* retinal into an all-*trans* retinol. The retinol is then transported from the photoreceptor cells to an adjacent support cell, such as the pigment cells, by an unknown intercellular shuttle protein. Once within the pigment cells, the all-*trans* retinol is transported through the pigment cells by a PINTA-like shuttle protein and introduced to a yet unidentified isomerase isomerizing the all-*trans* retinol to an 11-*cis* retinol. Following isomerization, the 11-*cis* retinol is oxidized by an RDH2/16-like, RDH3/7-like, RDH5-like, or RDH14-like dehydrogenase to an 11-*cis* retinal. The 11-*cis* retinal can then be shuttled back to the photoreceptor cells and delivered to the rhodopsin to reset the photopigment. In addition, RALBP may deliver all-*trans* retinal from the rhodopsin to retinochrome. Retinochrome can photoisomerize all-*trans* retinal to an 11-*cis* retinal upon absorption of a photon when retinal is high demand, or act as a storage molecule, retaining 11-*cis* retinal and reverting it back to all-*trans* retinal given sufficient time.

Table 3-1 - Name of query proteins and the animal from which it was isolated.
Proteins are organized by gene family and functional motifs are described, if applicable.

Protein (Abbreviation) [Gene]	Species	Protein family/general function
Retinol dehydrogenase 5 (RDH5) [Rdh5]	*Mus musculus*	Short chain dehydrogenase - NAD-/NADP-dependent oxidoreductases
Retinol dehydrogenase 8 (RDH8) [Rdh8]	*Mus musculus*	Short chain dehydrogenase - NAD-/NADP-dependent oxidoreductases
Retinol dehydrogenase 11 (RDH11) [RDH11]	*Homo sapiens*	Short chain dehydrogenase - NAD-/NADP-dependent oxidoreductases
Retinol dehydrogenase 12 (RDH12) [Rdh12]	*Mus musculus*	Short chain dehydrogenase - NAD-/NADP-dependent oxidoreductases
Photoreceptor Dehydrogenase (PDH) [Pdh]	*Drosophila melanogaster*	Short chain dehydrogenase - NAD-/NADP-dependent oxidoreductases
Lecithin: retinol acyltransferase (LRAT) [Lrat]	*Mus musculus*	H-rev107 - transfer of acyl groups
Retinal pigment epithelium 65 (RPE65) [Rpe65]	*Mus musculus*	Carotenoid oxygenase - carotenoid cleavage
Beta, beta-carotene 15, 15'-dioxygenase (BCOI) [BCO1]	*Homo sapiens*	Carotenoid oxygenase - carotenoid cleavage
Beta, beta-carotene 9', 10'-oxygenase (BCOII) [BCO2]	*Homo sapiens*	Carotenoid oxygenase - carotenoid cleavage
Neither inactivation nor after potential B (ninaB) [ninab]	*Drosophila melanogaster*	Carotenoid oxygenase - carotenoid cleavage
Cellular retinoic acid-binding protein (CRABP) [CRABP2]	*Homo sapiens*	Fatty-acid binding protein family (FABP)
Interphotoreceptor retinoid-binding protein (IRBP) [Rbp3]	*Mus musculus*	Peptidase_S41
Cytoplasmic retinol binding protein (CRBP) [Rbp1]	*Mus musculus*	Lipocalin - fatty-acid binding
Cellular retinaldehyde-binding protein (CRALBP) [Rlbp1]	*Mus musculus*	CRAL-TRIO - small lipophilic molecule binding
Prolonged depolarization afterpotential is not apparent (PINTA) [pinta]	*Drosophila melanogaster*	CRAL-TRIO - small lipophilic molecule binding
Retinal-binding protein (RALBP) [N/A]	*Todarodes pacificus*	CRAL-TRIO - small lipophilic molecule binding

CHAPTER 4. MUTATION OF AMINO ACIDS LINING THE BINDING POCKET OF SCALLOP RETINOCHROME ARE RESPONSIBLE FOR FINE SPECTRAL TUNING OF PHOTOPIGMENTS.

G. Dalton Smedley[1] and Jeanne M. Serb[1]

[1]Department of Ecology, Evolution, and Organismal Biology; Iowa State University; Ames, Iowa, U.S.A. 50011

Abstract

The relationship between genotype and phenotype is nontrivial due to often complex molecular pathways that make it difficult to unambiguously relate phenotypes to specific genotypes. Photopigments, however, present an opportunity to directly relate the amino acid sequence to the phenotype in the form of the absorbance peak or λ_{max}. We conducted a series of mutagenesis experiments, mutating gene sequence of retinochrome from the common bay scallop, *Argopecten irradians,* to match another closely related scallop retinochrome from the king scallop, *Pecten maximus,* at three amino acid sites of interest. The mutated *A. irradians* was then expressed and spectrally analyzed allowing us to compare λ_{max} of the wild-type scallop retinochromes and the mutant retinochromes in order to understand the effect (red or blue shift) on the absorbance by the photopigment. Our results show that the mutation of amino acids lining the binding pocket of opsins may be responsible for fine spectral tuning or small changes in the λ_{max} of these light sensitive proteins. Introduction of an amino acid with a hydroxyl group in close proximity to the bound retinal showed a consistent blue shift ranging from 7nm to 14nm highlighting the effect of changes in polarity to the λ_{max}. Mutations of amino acids deeper within helices of the opsin suggest the effect of

conformational changes on the binding pocket caused by the addition or removal of bulky amino acids. These alterations in the shape of the binding pocket may be responsible for fine spectral tuning in which amino acids compensate or overpower the effects of one another. This study is a step towards being able to map the relationship between genotype and phenotype of a photopigment and to development of more thorough prediction models of opsin function.

Introduction

One of the important questions in evolutionary biology is how changes to the genotype alter the phenotype of an organism. This relationship is far from trivial, as genes are part of complex pathways making it difficult to unambiguously relate phenotypes to specific genotypes, while at the same time, relating the phenotypic changes to the organism's ecological and physiological environments. The phototransduction system offers a rare opportunity to examine this relationship because it has been shown that mutation in a single amino acid can drastically alter the protein's phenotype allowing an organism to change its perception (Yokoyama 2000b). Phototransduction is the molecular conversion of light into a change in the electrical potential of a cell (Shichida and Matsuyama 2009). The functional unit of phototransduction is the photopigment, an apoprotein, opsin, covalently bound to retinal isomer. Opsins are a family of G-protein coupled receptors (GPCRs), a large group of seven transmembrane proteins. These proteins bind a retinal chromophore via a Schiff base linkage at a highly conserved lysine residue within the binding pocket of the opsin (Terakita 2005).

The bound retinal chromophore is the light absorbing molecule of the photopigment. When unbound to an opsin, retinal absorbs within the UV range of the light spectra; however, when the chromophore is covalently bound to the opsin, the absorbance is shifted into the visual spectrum (Nathans 1990; Hara, Hara, and Takeuchi 1967). When a photon of light is absorbed, the retinal is photoisomerized from 11-*cis* retinal to all-*trans* retinal isomer or vice versa, depending on the class of opsin (Shichida and Matsuyama 2009; Ramirez et al. 2016; Chang et al. 1995; Porter et al. 2012). Photopigments absorb only a portion of the light spectrum, and amaximum absorbance of the photopigment (λ_{max}) represents the phenotype.

The question of how the genotype of the opsin protein determines the absorption peak of the photopigment is longstanding. In vertebrates, the difference between λ_{maxes} of middle- and long-wavelength-sensitive pigments has been attributed to three-sites, and the difference λ_{max} of red and green pigments has been attributed to five sites of interest. These individual sites are responsible for small nm shifts, the cumulation of these mutations leads to a functional shift in physiological function of the protein (Yokoyama 2000b). Many studies have highlighted the interactions of the chromophore with amino acids of the binding pocket via theoretical calculations (Beppu and Kakitani 1994; Beppu 1997), mutagenesis studies (Nathans 1990; Oprian et al. 1991; Yokoyama 2000b, 2000a; Van Hazel et al. 2013) and quantum mechanics/molecular mechanics (QM/MM) modeling (Melaccio, Ferré, and Olivucci 2012; Ferré and Olivucci 2003; González-Luque et al. 2000) in vertebrates to try and better understand the effect of the structure of the protein and chromophore on the function of the protein. What these and other studies show is that photopigment function is not regulated by the amino acid

sequence of opsin *per se*, but by complex patterns of van der Waals, hydrogen bonding and polar interactions that the amino acids generate affecting the environment with which the chromophore is bound to the opsin protein in the binding pocket. These interactions are all noncovalent, and it is suggested that the polar interactions between the chromophore and the amino acid residues within the binding pocket of the opsin is responsible for the spectral tuning of the absorbance peak (Beppu and Kakitani 1994; Irving, Byers, and Leermakers 1970).

Spectral tuning is the modulation of the absorbance peaks or λ_{max} of a photopigment by alteration of the amino acid side chains which affect the chromophore binding pocket (Lin et al. 1998). Most work investigating the spectral tuning of opsins has focused on the visual opsins due to the importance of these visual pathways for the fitness of an organism and the evolution of complex traits. To truly understand the relationship between genotype and phenotype of opsins requires extensive mutagenesis studies comparing the effect of amino acids at sites of interest within the protein and their effect on the absorbance peak. Functional assays of visual opsins have begun dissecting what characteristics of the protein or photic environment are responsible for spectral tuning (Merbs and Nathans 1993; Hauser, van Hazel, and Chang 2014; Hunt et al. 2009; Hope et al. 1997), but the vast majority of this work has been done in vertebrate systems (e.g., monostable G_t-protein coupled opsins), with few exceptions in non-vertebrate species (e.g. bistable G_q-protein coupled opsins) (Wakakuwa et al. 2010; Briscoe 2002). Such work has enabled our understanding of spectral tuning in primate color vision (Irving, Byers, and Leermakers 1970; Carvalho et al. 2012) and shifts into UV-sensitivity (Shi, Radlwimmer, and Yokoyama 2001;

Yokoyama et al. 2006). Despite our understanding of vertebrate visual systems, we cannot assume that the amino acids changes identified in vertebrates also explain invertebrate systems. The crystal structure of invertebrate opsins show important differences compared to vertebrate opsin, such as greater organization in the cytoplasmic region correlating to bistability (Varma et al. 2019; Shimamura et al. 2008). This longstanding focus on vertebrate opsins has left us with a severe lack of understanding for how spectral tuning functions in invertebrates.

In order to identify amino acids involved in the spectral tuning of photopigments, we first need to be able to compare the amino acid sequence and protein function of two closely related species directly. Two examples of retinochrome were identified from eye transcriptomes from two closely related scallop species in our lab (Serb et al, in prep). Retinochrome (RTC) is a mollusc specific non-visual opsin and one of the few examples of a non-vertebrate, non-visual opsins which has been long studied. It was first identified in cephalopods (Brown and Brown 1958; Hara and Hara 1972; Hara, Hara, and Takeuchi 1967) and later in gastropods (Ozaki et al. 1986) with the most recent identification in scallops (Ramirez et al. 2016). It is part of the “photoisomerase” opsin group, recently renamed to RGR/retinochrome/peropsin as part of the tetraopsin group (Ramirez et al. 2016). These opsins are unique in their preference to bind all-*trans* retinal as a chromophore as opposed to the 11-*cis* retinal commonly used by most opsins. After its identification in the mid to late 1960’s, there was expansive work done to fully characterize the function of retinochrome in varying pH environments and given different retinal derivatives (Sperling and Hubbard 1975) and its function in molluscs (Hara and Hara 1968, 1973; Azuma, Azuma, and Kito 1974). Another interesting

characteristic of retinochrome is the location of the counterion. In other described opsins, there is a conserved glutamic acid necessary to stabilize the binding of the chromophore via a Schiff base linkage found within the third transmembrane helix. However, a series of mutagenesis experiments in retinochrome showed that this counterion in squid retinochrome resides in the IV-V loop, a site conserved across molluscan retinochromes (Terakita, Yamashita, and Shichida 2000). Retinochrome is a single copy gene, according to genomic analyses in scallops (Serb et al., *in prep*), with a narrow range of experimental peak absorbance (λ_{max}) observed in tested species (Hara, Hara, and Takeuchi 1967). Retinochrome is now poised to be used as a powerful model system to study the characteristics of spectral tuning of invertebrate opsin proteins and broaden our understanding of the relationship between genotype and phenotype.

Here, we hypothesize that amino acid residues forming the binding pocket of retinochrome are important in the spectral tuning of the protein. We propose that alteration in the shape or electrostatic environment of the binding pocket result in changes to the noncovalent interactions between the bound retinal and the retinochrome will thereby change the wavelength at which the bound chromophore absorbs light. To investigate this hypothesis, we cloned and expressed retinochrome *in vitro* from two closely related scallop species: the common bay scallop, *Argopecten irradians* and the king scallop, *Pecten maximus.* Using the spectra and alignment of amino acid sequences of the scallop retinochromes, we identified amino acid residues that may be responsible for λ_{max} shifts between these two retinochrome samples. We then created mutants of *A. irradians* retinochrome (Airr-RTC) mutating non-conserved

sites of interest to match that of *P. maximus* retinochrome (Pmax-RTC) in order to test the effect of alteration of specific amino acids on the absorbance of Airr-RTC using photospectroscopy. Our findings suggest that the mutation of sites within the binding pocket are responsible for altering the λ_{max} of retinochrome, but that there may be a hierarchy in the effect of the amino acid residues on the shift in absorbance. This study also highlights the potential of retinochrome as a model to more fully understand the interactions of amino acids within a photosensitive protein and their effects on the absorbance of the bound chromophore.

Methods

Retinochrome Cloning and Insertion in Vector

Previously assembled transcriptomes (Serb et al., *in prep*) were used to identify retinochrome sequences in *Argopecten irradians* and *Pecten maximus*. Based on those transcripts, UTR-specific primers were designed to amplify the complete coding regions from cDNA (Table 4-1). Scallop RNA was extracted from eye tissue using the RiboPure RNA extraction kit (Ambion) and converted to cDNA libraries. PCR was carried out with a reaction mixture equaling 50uL, containing, 5uL of 10x buffer, 1.5uL of 25mM MgCl, 4uL of 2.5mM dNTPs, 0.2uL Platinum Taq, 1uL of 10uM of forward and reverse primers (Table 4-1), and 1uL of 3uM template cDNA. The thermocycler protocol used was as follows, with variation in primer annealing temperatures: 95°C for 2min; 35 cycles of 95°C for 30s, primer temperature for 40s, and 72°C for 2min; and 72°C for 10 min. PCR products were size-screened using a 1% agarose gel electrophoresis, bands of expected size (923bp) were gel extracted (Qiagen Qiaquick Gel Extraction kit) and cloned using chemically competent *E. coli* cells following the manufacturer's protocol

(TOPO TA Cloning Kit with pCR2.1-TOPO). The identity of positive colonies from blue-white screening was confirmed by Sanger DNA sequencing using an ABI 3730 Capillary Electrophoresis Genetic Analyzer at the Iowa State University DNA Facility. The genes were then inserted into the expression vector p1D4-hrGFP II (Morrow and Chang 2010) to generate our working plasmids for retinochrome from *A. irradians* (Airr-RTC) and *P. maximus* (Pmax-RTC). These expression plasmids served as the templates for site-directed mutagenesis.

Modeling and Site Identification

In order to identify amino acids lining the binding pocket which may be responsible for altering the λ_{max} of retinochrome, amino acid sequences and predicted 3D models were compared. Amino acid sequences of *Argopecten irradians* retinochrome and *Pecten maximus* retinochrome were aligned using MAFFT v7.221 (Katoh and Standley 2013). To identify sites hypothesized to cause changes in λ_{max} of retinochrome, amino acid sequences of *A. irradians* and *P. maximus* retinochrome were submitted to GPCR-I-TASSER (Zhang et al. 2015) to create 3D models of each protein. The resulting models were then submitted to COACH (Yang, Roy, and Zhang 2013b, 2013a), a meta-server used to predict the active interaction sites within protein-ligand interactions. COACH outputs a list of amino acid sites it has predicted to interact with the ligand when bound based on proximity of the amino acid to the bound ligand model. This list of predicted sites was compared to the alignment of Airr-RTC and Pmax-RTC, specifically to identify predicted interaction sites that are also not conserved sites between the two species, revealing one predicted interaction site which differed between species.

The second approach to identifying amino acids responsible for altering the retinochrome binding pocket environment was based on the role of possible polar interactions between amino acids and variation in the shape or electrostatic environment of the binding pocket plays a role in spectral tuning of the λ_{max} of opsins. 3D models from COACH were loaded into UCSF Chimera v1.4 (Pettersen et al. 2004), a visualization software for molecular analyses and model comparison. Using Chimera, predicted interaction sites were differentially highlighted based on whether the amino acids were conserved between *A. irradian* and *P. maximus* amino acid sequences. Non-conserved amino acids outside of or far from the binding pocket were disregarded, as they are less likely to affect the polar or shape of the binding pocket. The distances of the predicted interaction sites to the active side chains of non-conserved amino acids were then individually measured. Distances less than 3.5 angstroms (maximal hydrogen bond length) were searched for, revealing two sites as targets for site-directed mutagenesis.

Site-directed Mutagenesis

In order to create the seven mutants (possible combinations of the three sites of interest explained previously), the Airr-RTC expression plasmid served as the template for the preceding mutagenesis experiments. DNA Polymerase PfuTurbo (Agilent, Santa Clara, CA) was used for all cloning experiments following the manufacturer's instructions with the same thermocycler profile with varying annealing temperatures dependent on the specifications of each primer set (Table 4-1). Following PCR protocol, reaction was subjected to a 2.5-h digestion with DpnI (New England Biolabs, Ipswich, MA) at 37°C. Five microliters of the reaction was then used to transform TOP10

chemically competent *E. coli* cells (Thermo Fisher Scientific, Waltham, MA). Positive colonies containing the correct mutant sequence were confirmed by Sanger sequencing. Overlapping primers were developed to create mutant Airr-RTC (Table 4-1). Single mutant primers were used in the creation of all mutants except for one double mutant plasmid. Due to the proximity of the selected sites, a set of primers including two mutation sites was used to guarantee mutation of both sites without removal of either. Mutant products were then used as templates for subsequent mutagenesis. Plasmids were amplified by incubating positively identified colonies in 1L liquid LB culture with 50 ug/mL kanamycin. Plasmids were purified using QIAGEN (Hilden, Germany) HiSpeed Plasmid Maxi Kit according to the manufacturer's instructions with the product sequenced to confirm mutant identity.

Cell Culture, Expression, and Pull-down

To express wild-type and mutant retinochrome proteins *in vitro*, 15 plates (Corning Falcon Standard Tissue Culture Dishes, 10cm, ref. 353003; Tewksbury, MA) of confluent HEK293T cells (ATCC, Manassas, VA) were transfected with 8 mg DNA and 20 mL 293fectin Transfecting Reagent (Thermo Fisher Scientific) per plate, according to the manufacturer's instructions. Plates were incubated for 24 hours before the minimum essential medium (MEM) was exchanged with new MEM containing 5 µmol all-*trans* retinal. Due to the addition of light sensitive all-*trans* retinal at this step, all subsequent culturing and experimentation was conducted in a dark-lab environment under dim red light. The plates were incubated for another 24 hours before cells were harvested by scraping the plates twice with 5 mL bufferA (3 mmol MgCl2, 140 mmol NaCl, 50 mmol HEPES pH 6.6, aprotinin [10 mg/mL], leupeptin [10 mg/mL]). All

subsequent centrifugation and incubation steps were at 4°C or on ice. Cells were collected by pellet following centrifugation (10 min at 1620 relative centrifugal force [RCF]) and resuspended in 10 mL buffer A. Cells were washed two times in total following the same protocol.

After a second wash, cells were resuspended in 2 mL per plate of buffer A with 5 µmol all-*trans* retinal to regenerate the photopigment. The cell suspension was nutated for 1 hour at 4°C. The regenerated cells were pelleted by centrifugation for collection at 38360 RCF for 20 minutes and resuspended in solubilization buffer (buffer A plus 1% n-dodecyl b-D-maltoside and glycerol [20% w/v]) using 1 mL solubilization buffer per plate. The solubilized cells were nutated for 1 hour at 4C. After the hour nutation, the mixture was centrifuged for 20 min at 42,740 RCF. The supernatant was then added to a 100 mL slurry resin (1:1 v/v resin/resin buffer) composed of 1D4 antibody (University of British Columbia, Canada) conjugated to sepharose beads and nutated for 30 minutes. The resin was washed three times with 5 mL washing buffer (buffer A with 1% n-dodecyl b-D-maltoside and glycerol [20% w/v] without aprotinin and leupeptin), and the protein was eluted with 2 mL elution buffer (washing buffer with 40 mmol Rho1D4 peptide [TETSQVAPA]), adapted from Oprian et. al. (1987). To concentrate the protein sample, eluate was concentrated to ~300 uL using 4 Amicon Ultra 0.5 mL 10 kDa centrifugal filters (Millipore, Billerica, MA).

Photospectroscopy

Ultraviolet-visible absorption spectra (250–750 nm) of purified proteins was measured at 15°C using a Hitachi U-3900 spectrophotometer (Chiyoda, Tokyo, Japan).

Data analysis was performed on the mean value of five spectral measurements with the software UV Solutions v4.2 (Hitachi). To test the proteins for photoreactivity, “dark” absorbance was measured first for each protein, e.g., the naive protein that has been incubated and regenerated with the all-*trans* retinal chromophore. Retinochrome proteins were tested independently with all-trans retinal because (1) retinochrome preferentially binds all-*trans* retinal preferentially (Hara and Hara 1967) and (2) retinochrome forms a stable pigment only in the presence of all-*trans* retinal (Faggionato and Serb 2017). The maximum absorbance of the all-*trans* retinal when the apoprotein is not present is 380 nm. Thus, any light-dependent isomerization converting free all-*trans* retinal to 11-cis retinal will be undetectable in the experimental system. Therefore, the most plausible explanation for any observed change in spectral absorbance is due to a conformational change of the retinal covalently bonded to the apoprotein.

For the “light” spectra, extracted proteins were bleached at different wavelengths according to λ_{max} identified from the “dark” spectra and then the absorbance was measured. Extracted proteins were first exposed to light at ~474 nm using two blue LEDs (MR16-B24-15-DI; superbrightleds.com) simultaneously irradiating both transparent sides of the cuvette (Hellma Analytics 104002B-10-40; Müllheim, Germany) for 3 minutes and the absorbance was recorded. A final exposure to white light was performed, followed by an absorbance measurement. Mean spectra was plotted using R scripts (R Core Team, 2017) with interpolation of the data points being performed with the R function smooth.spline. The differential spectra were calculated from two adjacent light treatments (i.e., the absorbance spectrum before exposure to light [dark spectra]

was subtracted from spectra after the exposure to blue light [blue spectra]). This approach minimized the likelihood of observed differences in spectrum being a result of unrelated factors such as handling of the cuvette between light exposures and absorbance measurements or degradation of the protein sample other than light treatments themselves. These methods were carried out for the wild-type proteins of *A. irradians* and *P. maximus* and then the seven mutant *A. irradians* proteins.

Results

Wild-type Retinochrome Sequences and Spectra

Cloned retinochrome sequence of *Argopecten irradians* and *Pecten maximus* retinochrome sequences were 308 and 307 amino acids, respectively. The two proteins were 92% similar, with 25 different amino acid residues. Many of these differences occur close to the C and N terminals of the protein within the transmembrane helices 1 and 7. There was high conservation within the binding pocket (99%), and a majority of these site differences result in the same biochemical classification of amino acid. Residues surrounding sites important for the function of retinochrome such as the conserved lysine for chromophore binding and counterion are conserved between scallop samples, as are amino acids around these sites. The lysine residue which forms the Schiff base is at site 274 in both retinochromes, and the counterion is a glutamate at site 160.

To determine the λ_{max} of *Pecten maximus* retinochrome (Pmax-RTC) and *Argopecten irradians* retinochrome (Airr-RTC), both proteins were expressed in HEK293T cells and incubated with all-*trans* retinal. Difference spectra inset into the

dark/light spectra also reveal a sharp peak signifying the formation of a photopigment. Following protein pulldown and concentration of samples, the spectra of retinochrome proteins were tested before exposure to light (Figure 4-1). The absorbance peak (λ_{max}) of Pmax-RTC was at 502nm while the absorbance peak of Airr-RTC was around 510nm. Analysis of the spectra after exposure of the protein to blue light showed no absorbance. These results suggest the property of retinochrome as a monostable opsin.

Site Identification and Mutants

GPCR-i-TASSER and COACH were used to predict ligand interaction sites. Squid rhodopsin (2Z73, (Murakami and Kouyama 2008)) was selected by the programs as the template for creating the 3D models and retinal was suggested as the ligand for predicting interaction sites. For both Airr-RTC and Pmax-RTC, 18 sites were identified as potential ligand interaction sites. All the predicted interaction sites were conserved with two exceptions: site 170 of Pmax-RTC (tryptophan), which was not predicted to be an interaction site from the Airr-RTC model, and site 188 of Airr-RTC (tyrosine) (Figure 4-2a, b), which was not predicted to be an interaction site in Pmax-RTC. Site 170 is conserved between retinochrome samples; however, site 188 is a tryptophan in Pmax-RTC rather than a tyrosine as is the case with Airr-RTC. Both amino acids are classified as hydrophobic with similar pKa scores between the present amino and indole ring of the residues and the hydroxyl group of the tyrosine (around pKa =10). However, the position of the tyrosine's hydroxyl group allows more direct change in polarity of the binding pocket (Shimono et al. 2001) and therefore may affect the absorbance of the photon by the bound retinal (Figure 4-2c). Site 188 was selected as the first mutant site due to these principles.

To identify other sites that may be responsible for altering the shape of or electrostatic environment within the binding pocket, the non-conserved sites were compared to the 3D models. Within UCSF Chimera, the 3D models were manipulated allowing measurements to be recorded of the distance between active groups of the amino acid residues in relation to one another as well as the retinal model. Many non-covalent polar interactions range from 0.5 angstroms to 3.5 angstroms, thus this range was used as a criterion for identifying sites of interest through bond measurements. Site 184 (Figure 4-2a, b) is a conserved methionine in both Pmax-RTC and Airr-RTC predicted by COACH to function as a ligand interaction site. A close site which is not conserved is site 181. Site 181 is a methionine in Airr-RTC but a leucine in Pmax-RTC. The primary difference between these nonpolar side chains is the presence of the thioether of the methionine. This thioether allows for oxidation of the residue and acts as a strong hydrogen bond acceptor. Bond measurements show that the sulfur of Met181 is only 2.850 angstroms from the side chain hydrogens of Met184 (Figure 4-2c). Site 181 was selected as a second site of interest due its potential in changing the polarity of an amino acid lining the binding pocket, site 184.

Using the same logic, non-conserved site 193 was also identified as a site of interest. Site 193 is a valine in Airr-RTC but an alanine in Pmax-RTC, and although it was not identified as a ligand interaction site, site 189, a leucine, was predicted by COACH as an interaction site for the ligand (Figure 4-2a, b). The isopropyl group of the valine is much bulkier than that of the methyl group found in alanine. Given the proximity of the hydrogens found on the carbon groups of the leucine to that of the

valine (between 2.735 angstroms and 3.202 angstroms) it is suggested that the location of predicted interaction site 193 is altered in the space of the protein (Figure 4-2c).

The seven mutant proteins of Airr-RTC were successfully created, comprising all the combinations of mutations at the respective sites. Wild-type Airr-RTC was used as the backbone template for the single mutant proteins. Mutants are labeled with a delta indicating presence of a mutation followed the wild-type and resulting amino acid, respectively, flanking the site number: Airr-RTC-ΔM181L, Airr-RTC-ΔY188W, and Airr-RTC-ΔV193A. Double mutant proteins used the single mutant retinochromes as a template using the same primers used to create the single mutants; however, due to the proximity of sites 188 and 193, new overhang primers were designed to include the mutations at both sites (Table 4-1). This ensured the production of a double mutant without the chance of reverse mutation. The resulting mutants are: Airr-RTC-ΔM181L-ΔY188W, Airr-RTC-ΔM181L-ΔV193A, Airr-RTC-ΔY188W-ΔV193A. Finally, the production of the triple mutant was carried out using Airr-RTC-ΔY188W-ΔV193A as the template and the primer set for site 181: Airr-RTC- ΔM181L-ΔY188W-ΔV193A.

Mutant Spectra

Spectra were analyzed for each mutant protein before and after exposure to blue light (Figure 4-3). The inset differential spectra of each mutant retinochrome showed sharp peaks showing the formation of photopigments. From the single mutants, site 181 (Met to Leu) had little change from the wild-type with a λ_{max} of 512nm, while site 188 (Tyr to Trp) and site193 (Val to Ala) resulted in a blue shift with λ_{maxes} of 505nm and 498nm, respectively. The blue shift from sites 188 and 193 suggest their role as spectral

tuning sites. Both double mutants containing mutations in site 188 (Airr-RTC-ΔM181L-ΔY188W and Airr-RTC-ΔY188W-ΔV193A) also shifted blue, with λ_{maxes} of 498nm and 502nm, respectively. The double mutations at sites 181 and 193 had little to no change in the absorbance peak from the wild-type (λ_{max} = 512nm). Considering the λ_{max} of the single mutants at sites 181 and 193, the combination of the two shows the effects at site 181 either compensate for or outweigh the effects at site 193. The triple mutant (Airr-RTC-ΔM181L-ΔY188W-ΔV193A) shows a λ_{max} was around 500nm, a slight blue shift from the wild-type Pmax-RTC. All spectra showed the expected bleaching known to be a trait of retinochrome resulting in no absorbance following the exposure of light.

Discussion

Predicted Ligand Interaction Sites Alter λ_{max}

In this study, we investigate how alterations in genotype affect the phenotype of scallop retinochrome. We address the hypothesis that amino acids lining the binding pocket may in part be responsible for fine spectral tuning of the λ_{max} of retinochrome. Our results show that mutations forming the binding pocket of the Airr-RTC to match that of Pmax-RTC do cause a distinct blue shift in the λ_{max}. These findings highlight the spectral tuning effect of amino acids away from the Schiff-base linkage site and the counterion and suggest that these secondary sites may be responsible for smaller adjustments in λ_{max} or fine spectral tuning.

Using the spectra and sequences of wild-type retinochrome isolated from closely related scallop species, we identified amino acids which may alter the shape or electrostatic environment of the opsin binding pocket. The bound retinal chromophore of

an opsin acts as the light absorbing aspect of the photopigment, but the shape (Hirayama et al. 1994) and electrostatic environment (Shimono et al. 2001) of the opsin binding pocket is hypothesized to be responsible to determining the λ_{max} of the photopigment. Previous work on spectral tuning has focused on non-conserved amino acids surrounding the sites of the Schiff base or sites surrounding the counterion (Shimono et al. 2001). However, in the case of the scallop retinochrome, the sequences are highly conserved around both the Schiff base binding site and the counterion, thus the difference in λ_{max} must be attributed to non-conserved sites elsewhere in the protein.

Site 188 is a tyrosine in *A. irradians* identified by COACH as a site of interest. The effect of presence or absence of amino acids with hydroxyl groups on the λ_{max} of photopigments is well documented in both vertebrate and invertebrate models (Shimono et al. 2001; Merbs and Nathans 1993; Nakayama and Khorana 1991). The results of these studies suggest that the change in λ_{max} is due to the potential dipole moments created by the oxygen of present hydroxyl groups in proximity of the chromophore. The 3D model of scallop retinochrome shows the hydroxyl group of the tyrosine actually inserted into the aromatic ring of the retinal. This impossible orientation is likely a result of the software using 11-*cis* retinal compared to the preferred all-*trans* retinal used by retinochrome. The straight poly-carbon chain of the all-*trans* retinal rests in the binding pocket of the opsin differently than the 11-*cis* retinal. It would be expected that the 3D model would show a greater distance between the chromophore and the hydroxyl group of the tyrosine, but the potential for dipole moments with the hydroxyl group would still affect the electrostatic environment of the binding pocket. This hydroxyl group is absent in the tryptophan at site 188 of *Pecten maximus*. Spectral analyses of all Airr-RTC

mutants possessing a mutation at site Y188W show a blue shift in λ_{max} ranging from 7nm to 14nm. This consistent blue shift supports previous findings but further supports the role of amino acids possessing hydroxyl groups in the spectral tuning and that its effect also functions in relation to the chromophore as well as the Schiff base binding site and counterion.

The second and third mutant sites, M181L and V193A, investigate the effects of amino acids in proximity to predicted interaction sites rather than those in direct proximity to the retinal chromophore. Our results point to a hierarchy in the effect of amino acid residues. Retinochrome single mutant possessing a mutation at M181L and double mutant possessing mutation in M181L and V193A show no change in λ_{max} (512nm); however, the single mutant V193A shows a large blue shift (498nm). The lack of shift in the double mutant between these sites show the effect posed by M181L may compensate for the alteration caused by the mutation at V193A. The mutation of hydrophobic residues in bacterial rhodopsins near the Schiff base, the counterion, and when near the chromophore have been shown to cause drastic λ_{max} shifts (up to 80nm) (Greenhalgh et al. 1993). The side chains of hydrophobic residues tend to be bulky due to the saturation of hydrogens and mutation of hydrophobic residues may alter the geometry of the binding pocket by reducing the size of residues deep to the binding pocket allowing residues lining the binding pocket to fill the space.

The effect of the single mutant V193A is curious and more drastic than expected; however, this may suggest a stronger effect on the shape of the binding pocket on the λ_{max} of the photopigment than anticipated. Our spectra do not show such extreme shifts in λ_{max} as previously reported, but this could be due to their distance to the Schiff-base,

counterion, or chromophore directly. The effect the mutations cause in shape of the binding pocket may be lessened due to their secondary effect on those amino acids which line the binding pocket and thereby are in closer proximity to the chromophore. To further investigate the effect of shape changes of the binding pocket, future work should include mutating sites lining the binding pocket with bulkier amino acids in closer proximity to the chromophore and binding pocket.

Scallop Retinochrome as a Model Photopigment

In this study, we investigate how alterations in genotype affect the phenotype of scallop retinochrome. We focus on the potential of amino acids lining the binding pocket and their potential in fine spectral tuning of the λ_{max} of retinochrome. Here, we also present the first instance of cloning, expression, and spectral analysis of retinochrome from multiple scallops species. Our results show highly conserved amino acid sequences between retinochrome of two closely related scallop species, *Pecten maximus* and *Argopecten irradians*, but different λ_{max} values. These λ_{max} values are consistent with reported variation in absorbance peaks of other retinochromes from cephalopods (Terakita, Yamashita, and Shichida 2000). Expression and spectral analysis showed absorbance peaks within the known range of previous studies of retinochrome (490nm to 522nm) (Hara and Hara 1972; Hara, Hara, and Takeuchi 1967). The retinochromes of cephalopods were also compared to the identified scallop sequences. The cephalopod retinochrome (squid, octopus, and cuttlefish) shared a similar length, but the lack of conservation (~25% conserved) presented too many sites possibly responsible for spectral tuning between scallops and cephalopods to infer their effect. However, the high level of conservation within scallop retinochrome helps narrow

the number of amino acids possibly responsible for the difference of 8nm between protein samples. This range is not large, and likely does not affect the function of the protein as a putative visual cycle protein, but the small range allows the use of scallop retinochrome to fine focus the effects of spectral tuning in photopigments.

The workflow we describe shows the utility of scallop retinochrome as a model for understanding the relationship between genotype and phenotype of photopigments. With the rapid expansion of molecular data available due to advances in sequencing technology, the identification of novel opsin proteins has increased in tandem. However, the process of understanding their function is limited by the complexity of expression and analysis experiments. Rather than express and test these opsins individually as they have classically described, some researchers have begun building QM/MM models to explain how the 3D structure and amino acid sequences of opsin proteins define the λ_{max} (Melaccio, Ferré, and Olivucci 2012; Hirayama et al. 1994; Ferre and Olivucci 2004; Coto et al. 2006). Both *Argopecten irradians* and *Pecten maximus* species are economically important and found readily in aquaculture for their use as a food source making it easy to acquire samples, and work with these species do not present the same hurdles often required when working with vertebrate species. Our lab has presented our expression system previously (Faggionato and Serb 2017) with use in the study of visual opsins in *A. irradians*. However, unlike those visual opsins, retinochrome is more readily expressed within the HEK293T cells requiring half the number of plates and thereby materials for greater absorbance and more profound peaks. We suggest future studies use scallop retinochrome as a robust, easily accessible alternative to vertebrate opsins in the study of how amino acid sequence informs the absorbance of

photopigments. The use of retinochrome as a model will enable continued mutagenesis experiments to fine tune the effect of amino acid changes within opsins at sites of interest on the λ_{max}. This growing dataset can then be used to expand current QM/MM models to more accurately infer function of newly identified opsins based on sequence without the need for extensive molecular testing.

Conclusion

The work presented here shows the first instance of expression and spectral analysis of retinochrome from scallop species. These results show the absorbance peaks of scallop retinochrome to fall within the blue-green range of the visible spectrum similar to previously defined cephalopod retinochrome. We propose methods to identify amino acid sites potentially responsible for spectral tuning of photopigments. Previous work shows the effect of altering amino acids near sites of import, such as the Schiff base binding site and counterion, but from these mutagenesis experiments, we show that sites elsewhere in the protein may be responsible for smaller differences or fine tuning in the λ_{max} of opsin proteins. Our findings also highlight the effectiveness of scallop retinochrome as a model system for investigating the relationship between genotype and phenotype of photopigments requiring fewer resources and easier access to samples.

This study is limited in part by the accuracy of the 3D models which were used as background for the identification of the sites of interest. Retinochrome does not currently have a crystallographic model with or without a bound retinal. GPCRs are historically difficult to create x-ray models of due to their necessity to maintain portions of the cell membrane to retain their structure, but further work would benefit from a more well-

defined 3D structure. Future work should continue to explore the effect of different classes of amino acids around the chromophore binding site and counterion as well as sites that may affect the electrostatic environment or geometry of the binding pocket. The unique characteristics of retinochrome as a monostable opsin and its preferential binding of all-*trans* rather than 11-*cis* retinal are also worth further investigation. Comparison of closely related opsins, such as scallop retinochrome and visual opsins, would allow use of a similar workflow to identified how the genotype determines these characteristics. Studies to further map the effects amino acid relationships within the opsin and their effect on spectral tuning will allow the development of more thorough, complex mathematical models to infer phenotype from genotype of newly identified opsin proteins.

References

Azuma, Masami, Katsu Azuma, and Yuji Kito. 1974. “Circular Dichroism of Squid Retinochrome.” *Journal of Chemical Information and Modeling* 53: 1689–99. https://doi.org/10.1017/CBO9781107415324.004.

Beppu, Yoshitaka. 1997. “Theoretical Study of Color Control MEchanism in Retinal Proteins: Orientational Effects of Aromatic Amino Acid Residues upon Opsin Shift.”

Beppu, Yoshitaka, and Toshiaki Kakitani. 1994. “Theoretical Study of Color Control Mechanism in Retinal Proteins” 59 (6): 660–69.

Briscoe, Adriana D. 2002. “Letter to the Editor Homology Modeling Suggests a Functional Role for Parallel Amino Acid Substitutions.” *Molecular Biology and Evolution* 19 (6): 983–86. https://doi.org/10.1006/gyno.1999.5675.

Brown, Paul K., and Patricia S. Brown. 1958. “Visual Pigments of the Octopus and Cuttlefish.” *Nature* 182 (4645): 1288–90. https://doi.org/10.1038/1821288a0.

Carvalho, Livia S., Wayne L. Davies, Phyllis R. Robinson, and David M. Hunt. 2012. “Spectral Tuning and Evolution of Primate Short-Wavelength-Sensitive Visual Pigments.” *Proceedings of the Royal Society B: Biological Sciences* 279 (1727): 387–93. https://doi.org/10.1098/rspb.2011.0782.

Chang, Belinda S. W., Keith A. Crandall, John P. Carulli, and Daniel L. Hartl. 1995. “Opsin Phylogeny and Evolution: A Model for Blue Shifts in Wavelength Regulation.” *Molecular Phylogenetics and Evolution*. https://doi.org/10.1006/mpev.1995.1004.

Coto, Pedro B., Angela Strambi, Nicolas Ferre, and Massimo Olivucci. 2006. “The Color of Rhodopsins at the Ab Initio Multiconfigurational Perturbation Theory Resolution.” *Proceedings of the National Academy of Sciences of the United States of America* 103 (46): 17154–59. https://doi.org/10.1073/pnas.0604048103.

Faggionato, Davide, and Jeanne M. Serb. 2017. “Strategy to Identify and Test Putative Light-Sensitive Non-Opsin G-Protein-Coupled Receptors: A Case Study.” *Biological Bulletin* 233 (1): 70–82. https://doi.org/10.1086/694842.

Ferre, Nicolas, and Massimo Olivucci. 2004. “Energy Storage of Rhodopsin Resolved at the Multiconfigurational Perturbation Theory Level” 101 (52): 17908–13.

Ferré, Nicolas, and Massimo Olivucci. 2003. “Probing the Rhodopsin Cavity with Reduced Retinal Models at the CASPT2/CASSCF/AMBER Level of Theory.” *Journal of the American Chemical Society* 125 (23): 6868–69. https://doi.org/10.1021/ja035087d.

González-Luque, Remedios, Marco Garavelli, Fernando Bernardi, Manuela Merchán, Michael A. Robb, and Massimo Olivucci. 2000. “Computational Evidence in Favor of a Two-State, Two-Mode Model of the Retinal Chromophore Photoisomerization.” *Proceedings of the National Academy of Sciences of the United States of America* 97 (17): 9379–84. https://doi.org/10.1073/pnas.97.17.9379.

Greenhalgh, D. A., D. L. Farrens, S. Subramaniam, and H. G. Khorana. 1993. “Hydrophobic Amino Acids in the Retinal-Binding Pocket of Bacteriorhodopsin.” *Journal of Biological Chemistry* 268 (27): 20305–11.

Hara, Tomiyuki, and Reiko Hara. 1967. “Rhodopsin and Retinochrome in the Squid Retina.” *Nature* 214.

———. 1968. “Regeneration of Squid Retinochrome.” *Nature* 219: 450–54. https://doi.org/10.1038/219450a0.

———. 1972. “Cephalopod Retinochrome.” *Photochemistry of Vision* 7 (1): 720–46. https://doi.org/10.1016/S0076-6879(82)81031-8.

———. 1973. “Isomerization of Retinal Catalysed by Retinochrome in the Light.” *Nature* 242 (115): 39–43. https://doi.org/10.1038/10.1038/newbio242039a0.

Hara, Tomiyuki, Reiko Hara, and Jitsuzo Takeuchi. 1967. “Rhodopsin and Retinochrome in the Octopus Retina.” *Nature* 214: 572–73. https://doi.org/10.1038/214572a0.

Hauser, Frances E., Ilke van Hazel, and Belinda S W Chang. 2014. “Spectral Tuning in Vertebrate Short Wavelength-Sensitive 1 (SWS1) Visual Pigments: Can Wavelength Sensitivity Be Inferred from Sequence Data?” *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 3228: 529–39. https://doi.org/10.1002/jez.b.22576.

Hazel, Ilke Van, Amir Sabouhanian, Lainy Day, John A. Endler, and Belinda Sw Chang. 2013. “Functional Characterization of Spectral Tuning Mechanisms in the Great Bowerbird Short-Wavelength Sensitive Visual Pigment (SWS1), and the Origins of UV/Violet Vision in Passerines and Parrots.” *BMC Evolutionary Biology* 13 (1). https://doi.org/10.1186/1471-2148-13-250.

Hirayama, Junichi, Yasushi Imamoto, Yoshinori Shichida, Tŏru Yoshizawa, Alfred E. Asato, Robert S.H. Liu, and Naoki Kamo. 1994. “SHAPE OF THE CHROMOPHOIW BINDING SITE IN Pharaonis PHOBORHODOPSIN FROM A STUDY USING RETINAL ANALOGS.” *Photochemistry and Photobiology* 60 (4): 388–93. https://doi.org/10.1111/j.1751-1097.1994.tb05121.x.

Hope, Andrew J., Julian C. Partridge, Kanwaljit S. Dulai, and David M. Hunt. 1997. “Mechanisms of Wavelength Tuning in the Rod Opsins of Deep-Sea Fishes.” *Proceedings of the Royal Society B: Biological Sciences* 264 (1379): 155–63. https://doi.org/10.1098/rspb.1997.0023.

Hunt, David M., Livia S. Carvalho, Jill A. Cowing, and Wayne L. Davies. 2009. “Evolution and Spectral Tuning of Visual Pigments in Birds and Mammals.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1531): 2941–55. https://doi.org/10.1098/rstb.2009.0044.

Irving, Charles S., Gary W. Byers, and Peter A. Leermakers. 1970. “Spectroscopic Model for the Visual Pigments. Influence of Microenvironmental Polarizability.” *Biochemistry* 9 (4): 858–64. https://doi.org/10.1021/bi00806a020.

Katoh, Kazutaka, and Daron M. Standley. 2013. “MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability.” *Molecular Biology and Evolution* 30 (4): 772–80. https://doi.org/10.1093/molbev/mst010.

Lin, Steven W., Gerd G. Kochendoerfer, Kate S. Carroll, Dorothy Wang, Richard A. Mathies, and Thomas P. Sakmar. 1998. “Mechanisms of Spectral Tuning in Blue Cone Visual Pigments: Visible and Raman Spectroscopy of Blue-Shifted Rhodopsin Mutants.” *Journal of Biological Chemistry* 273 (38): 24583–91. https://doi.org/10.1074/jbc.273.38.24583.

Melaccio, Federico, Nicolas Ferré, and Massimo Olivucci. 2012. “Quantum Chemical Modeling of Rhodopsin Mutants Displaying Switchable Colors.” *Physical Chemistry Chemical Physics* 14 (36): 12485–95. https://doi.org/10.1039/c2cp40940b.

Merbs, Shannath L., and Jeremy Nathans. 1993. “Role of Hydroxyl-Bearing Amino Acids in Differentially Tuning the Absorption Spectra of the Human Red and Green Cone Pigments.” *Photochemistry and Photobiology* 58 (5): 706–10. https://doi.org/10.1111/j.1751-1097.1993.tb04956.x.

Morrow, James M., and Belinda S.W. Chang. 2010. “The P1D4-HrGFP II Expression Vector: A Tool for Expressing and Purifying Visual Pigments and Other G Protein-Coupled Receptors.” *Plasmid* 64 (3): 162–69. https://doi.org/10.1016/j.plasmid.2010.07.002.

Murakami, Midori, and Tsutomu Kouyama. 2008. “Crystal Structure of Squid Rhodopsin.” *Nature* 453 (7193): 363–67. https://doi.org/10.1038/nature06925.

Nakayama, Tomoko A., and H. Gobind Khorana. 1991. “Mapping of the Amino Acids in Membrane-Embedded Helices That Interact with the Retinal Chromophore in Bovine Rhodopsin.” *Journal of Biological Chemistry* 266 (7): 4269–75.

Nathans, Jeremy. 1990. “Determinants of Visual Pigment Absorbance: Identification of the Retinylidene Schiff’s Base Counterion in Bovine Rhodopsin.” *Biochemistry* 29 (41): 9746–52. https://doi.org/10.1021/bi00493a034.

Oprian, Daniel D., Robert S. Molday, Randal J. Kaufman, and H. Gobind Khorana. 1987. “Expression of a Synthetic Bovine Rhodopsin Gene in Monkey Kidney Cells.” *Proceedings of the National Academy of Sciences of the United States of America* 84 (24): 8874–78. https://doi.org/10.1073/pnas.84.24.8874.

Oprian, Daniel D., Sandra L. Pelletier, Ana B. Asenjo, and Ning Lee. 1991. “Design, Chemical Synthesis, and Expression of Genes for the Three Human Color Vision Pigments.” *Biochemistry* 30 (48): 11367–72. https://doi.org/10.1021/bi00112a002.

Ozaki, Koichi, Akihisa Terakita, Reiko Hara, and Tomiyuki Hara. 1986. “Rhodopsin and Retinochrome in the Retina of a Marine Gastropod, Conomulex Luhuanus.” *Vision Research* 26 (5): 691–705. https://doi.org/10.1016/0042-6989(86)90083-0.

Pettersen, Eric F., Thomas D. Goddard, Conrad C. Huang, Gregory S. Couch, Daniel M. Greenblatt, Elaine C. Meng, and Thomas E. Ferrin. 2004. “UCSF Chimera - A Visualization System for Exploratory Research and Analysis.” *Journal of Computational Chemistry* 25 (13): 1605–12. https://doi.org/10.1002/jcc.20084.

Porter, Megan L., Joseph R. Blasic, Michael J. Bok, Evan G. Cameron, Thomas Pringle, Thomas W. Cronin, and Phyllis R. Robinson. 2012. “Shedding New Light on Opsin Evolution.” *Proc Biol Sci* 279 (1726): 3–14. https://doi.org/10.1098/rspb.2011.1819.

Ramirez, M. Desmond, Autum N. Pairett, M. Sabrina Pankey, Jeanne M. Serb, Daniel I. Speiser, Andrew J. Swafford, and Todd H. Oakley. 2016. “The Last Common Ancestor of Most Bilaterian Animals Possessed at Least Nine Opsins.” *Genome Biology and Evolution* 8 (12): 3640–52. https://doi.org/10.1093/gbe/evw248.

Serb, Jeanne M., Davide Faggionato, G. Dalton Smedley, Arun Seetharam, Andrew J. Severin, and Autum N. Pairett. n.d. “Expression and Spectral Analysis of Twelve Opsins Reveals Astonishing Photochemical Diversity in the Common Bay Scallop Argopecten Irradians (Mollusca: Bivalvia).”

Shi, Yongsheng, F. Bernhard Radlwimmer, and Shozo Yokoyama. 2001. “Molecular Genetics and the Evolution of Ultraviolet Vision in Vertebrates.” *Proceedings of the National Academy of Sciences of the United States of America* 98 (20): 11731–36. https://doi.org/10.1073/pnas.201257398.

Shichida, Yoshinori, and Take Matsuyama. 2009. “Evolution of Opsins and Phototransduction.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 364 (1531): 2881–95. https://doi.org/10.1098/rstb.2009.0051.

Shimamura, Tatsuro, Kenji Hiraki, Naoko Takahashi, Tetsuya Hori, Hideo Ago, Katsuyoshi Masuda, Koji Takio, Masaji Ishiguro, and Masashi Miyano. 2008. “Crystal Structure of Squid Rhodopsin with Intracellularly Extended Cytoplasmic Region.” *Journal of Biological Chemistry* 283 (26): 17753–56. https://doi.org/10.1074/jbc.C800040200.

Shimono, Kazumi, Yukako Ikeura, Yuki Sudo, Masayuki Iwamoto, and Naoki Kamo. 2001. “Environment around the Chromophore in Pharaonis Phoborhodopsin: Mutation Analysis of the Retinal Binding Site.” *Biochimica et Biophysica Acta - Biomembranes* 1515 (2): 92–100. https://doi.org/10.1016/S0005-2736(01)00394-7.

Sperling, Linda, and Ruth Hubbard. 1975. “Squid Retinochrome.” *The Journal of General Physiology* 65 (2): 235–51. https://doi.org/10.1085/jgp.65.2.235.

Terakita, Akihisa. 2005. “The Opsins.” *Genome Biology* 6 (3): 213. https://doi.org/10.1186/gb-2005-6-3-213.

Terakita, Akihisa, Takahiro Yamashita, and Yoshinori Shichida. 2000. “Highly Conserved Glutamic Acid in the Extracellular IV-V Loop in Rhodopsins Acts as the Counterion in Retinochrome, a Member of the Rhodopsin Family.” *Proceedings of the National Academy of Sciences of the United States of America* 97 (26): 14263–67. https://doi.org/10.1073/pnas.260349597.

Varma, Niranjan, Eshita Mutt, Jonas Mühle, Valérie Panneels, Akihisa Terakita, Xavier Deupi, Przemyslaw Nogly, Gebhard F.X. Schertler, and Elena Lesca. 2019. “Crystal Structure of Jumping Spider Rhodopsin-1 as a Light Sensitive GPCR.” *Proceedings of the National Academy of Sciences of the United States of America* 116 (29): 14547–56. https://doi.org/10.1073/pnas.1902192116.

Wakakuwa, Motohiro, Akihisa Terakita, Mitsumasa Koyanagi, Doekele G. Stavenga, Yoshinori Shichida, and Kentaro Arikawa. 2010. “Evolution and Mechanism of Spectral Tuning of Blue-Absorbing Visual Pigments in Butterflies.” *PLoS ONE* 5 (11): 1–8. https://doi.org/10.1371/journal.pone.0015015.

Yang, Jianyi, Ambrish Roy, and Yang Zhang. 2013a. “BioLiP: A Semi-Manually Curated Database for Biologically Relevant Ligand-Protein Interactions.” *Nucleic Acids Research* 41 (D1): 1096–1103. https://doi.org/10.1093/nar/gks966.

———. 2013b. “Protein-Ligand Binding Site Recognition Using Complementary Binding-Specific Substructure Comparison and Sequence Profile Alignment.” *Bioinformatics* 29 (20): 2588–95. https://doi.org/10.1093/bioinformatics/btt447.

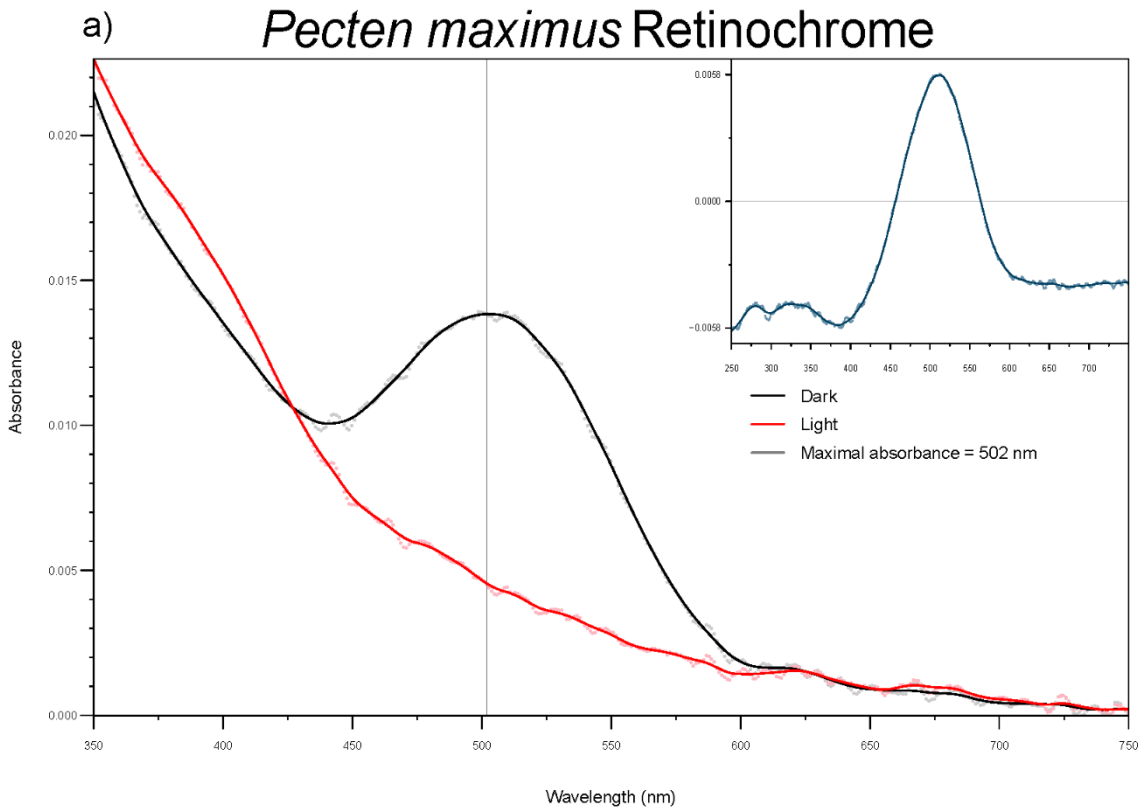
Yokoyama, Shozo. 2000a. *Molecular Evolution of Vertebrate Visual Pigments*. *Progress in Retinal and Eye Research*. Vol. 19. https://doi.org/10.1016/S1350-9462(00)00002-1.

———. 2000b. “Phylogenetic Analysis and Experimental Approaches to Study Color Vision in Vertebrates.” *Methods in Enzymology* 315 (1997): 312–25. https://doi.org/10.1016/s0076-6879(00)15851-3.

Yokoyama, Shozo, William T. Starmer, Yusuke Takahashi, and Takashi Tada. 2006. “Tertiary Structure and Spectral Tuning of UV and Violet Pigments in Vertebrates.” *Gene* 365 (1-2 SPEC. ISS.): 95–103. https://doi.org/10.1016/j.gene.2005.09.028.

Zhang, Jian, Jianyi Yang, Richard Jang, and Yang Zhang. 2015. “GPCR-I-TASSER: A Hybrid Approach to G Protein-Coupled Receptor Structure Modeling and the Application to the Human Genome.” *Structure* 23 (8): 1538–49. https://doi.org/10.1016/j.str.2015.06.007.

Figures and Tables



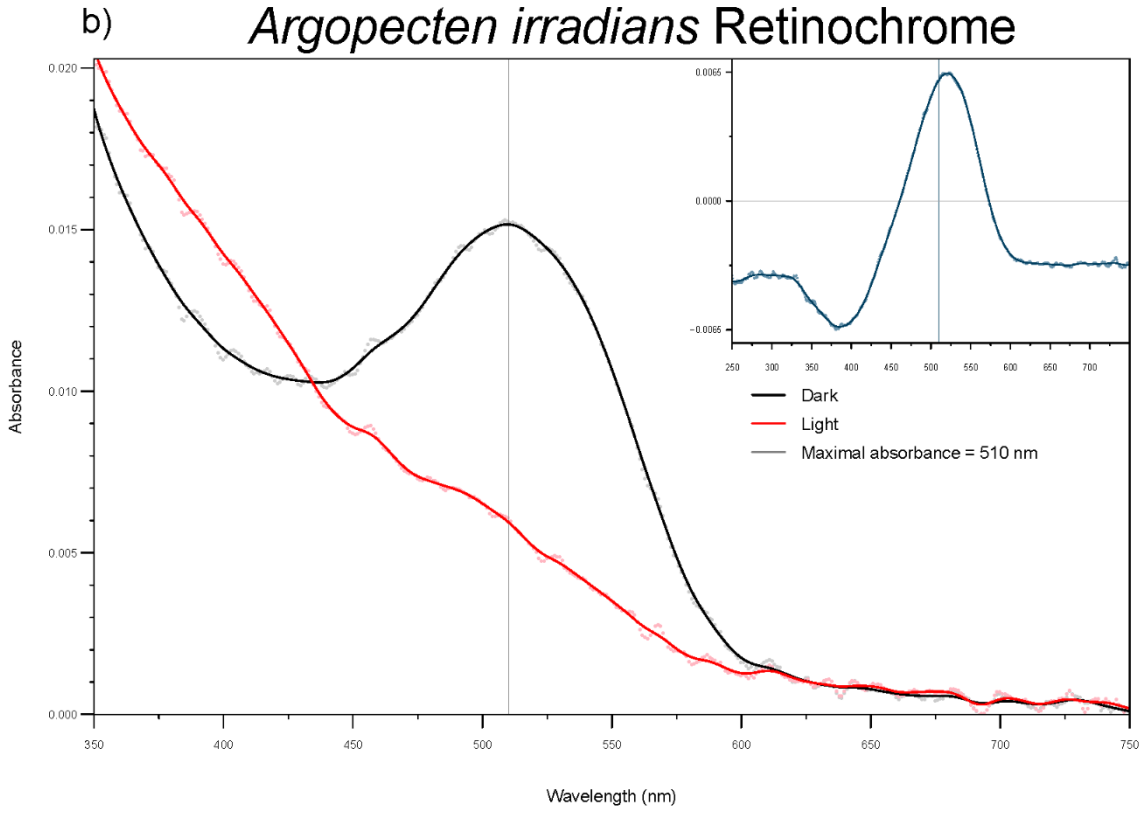


Figure 4-1 - Spectral analysis of Airr-RTC and Pmax-RTC. Black curve shows the absorbance spectrum of the protein prior to its exposure to light, while the red curve shows the absorbance spectrum of the protein following exposure to light for 3 minutes. The inset spectra are the difference in spectra of dark minus light indicating the presence of the photopigment. The grey vertical line of each spectrum highlights the lambda max.

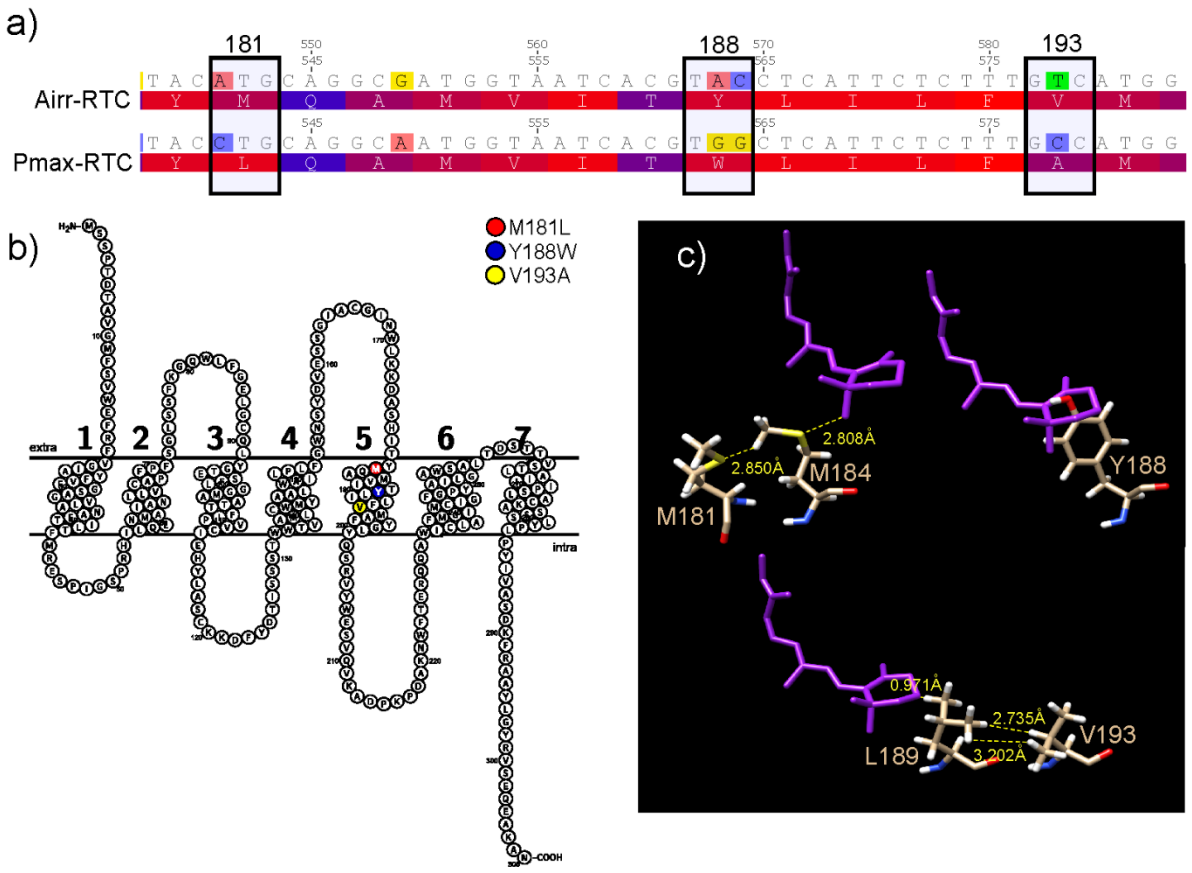


Figure 4-2 - Comparison of Airr-RTC and Pmax-RTC amino acid sequence and light absorption spectra. a) Alignment of amino acid sequences of Pmax-RTC and Airr-RTC with red boxes showing the nucleotide mutation site and resulting amino acid. Large numbers show the number of the amino acid, with smaller numbers showing the nucleotide numbers. Coloration of the amino acid boxes describe the amount of hydrophobicity of the side chains ranging from red (more hydrophobic) to blue (less hydrophobic). b) Snakeplot showing secondary structure of Airr-RTC. The sites of the mutagenesis protocol are highlighted in red (M181), blue (Y188), and yellow (V193). c) 3D modeling of bound retinal chromophore and sites of interest. The retinal chromophore is highlighted by a purple color with the amino acid chains colored by atom composition. Yellow dash bars and numbers show the distances between structures in angstroms. Amino acid residue identities are labeled in beige next to the residue.

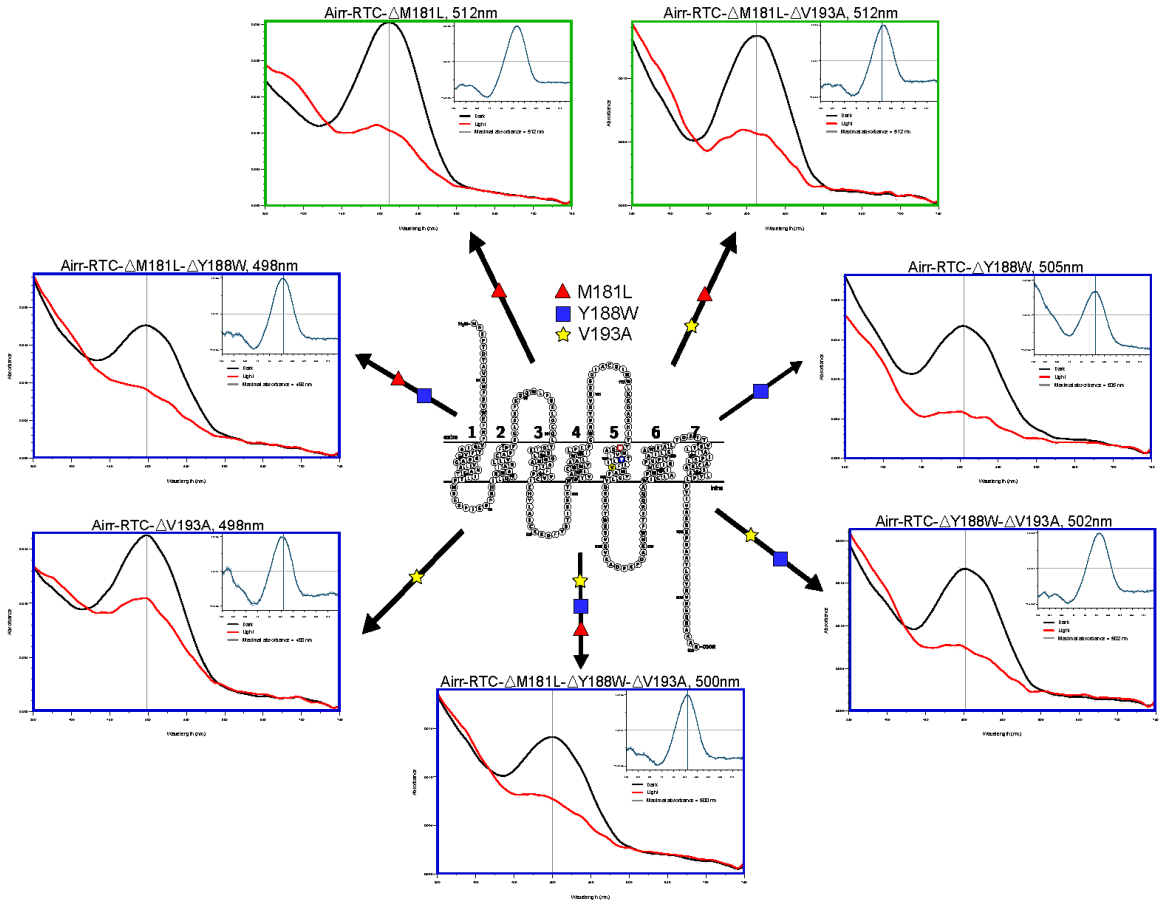


Figure 4-3 - Absorption spectra of mutant Argopecten irradians retinochrome. Black curves show the plot of dark (unexposed) spectra and red curves show absorption after 3-minute exposure to blue light. Vertical black lines highlight absorption peaks of mutants, the names of which reside above the spectra along with the value of the lambda max in nanometers. Colored boxes around the spectra show the shift (blue or green) compared to the wild-type Airr-RTC. Colored symbols (red triangle, blue square, and yellow star) correspond to the snakeplot and presence of mutations at sites 181, 188, and 193, respectively, within the protein of the resulting spectra.

Table 4-1 - Primer sequences and annealing temperatures used in cloning of wild-type retinochromes and the creation of the seven mutants. Primers are read 5 prime to 3 prime and the annealing temperatures are in Celsius.

Cloning Target	Sequence (5' to 3')	Annealing temp (°C)
Airr-RTC: Out-UTR_F Out-UTR_R In-UTR_F In-UTR_R BamHI site EcoR1 site	 TGCATGGCAGTGGCTCGGAA ACGTCACTCGTTTCCTGTCTCAACA CACATTTGATAGAATTGCTCTCG CCTGACTGAAAATAGATAAATCTCTG ATGCGGATCCCACCATGAGCTCCCCTACAGATACCG GCATGAATTCTTGGCCTTGGCTTCCTGTTC	 Step up 49 – 54°C Step up 49 – 54°C Step up 49 – 54°C Step up 49 – 54°C 55°C 55°C
Pmax-RTC: Out-UTR_R In-UTR_F In-UTR_R BamHI site EcoR1 site	 CCACGGACGCGGGGGTATTG GCACAGTGTTAGATAGAGCTCGAGGG TGCCTGGCGGAGGACCTTCA GCGGATCCCACCATGTCGTCACCTACTGATAC GCATGAATTCTTGGCCTTGGCTTCCTGC	 Step up 49 – 54°C Step up 49 – 54°C Step up 49 – 54°C 55°C 55°C
Airr-RTC ΔM181L: Forward Reverse	 GAGTCACATTACGTACCTGCAGGCGATGGTAATC TAATGGTAGCGGACGTCCATGCATTACACTGAGC	 60°C 60°C
Airr-RTC ΔW188Y: Forward Reverse	 AGGCGATGGTAATCACGTGGCTCATTCTCTTTGTCATGG GTACTGTTTCTCTTACTCGGTGCACTAATGGTAGCGGAC	 60°C 60°C
Airr-RTC ΔV193A: Forward Reverse	 GTACCTCATTCTCTTTGCCATGGCGTTTTACGGAC CAGGCATTTTGCGGTACCGTTTCTCTTACTCCATG	 60°C 60°C
Airr-RTC ΔW188Y+V193A: Forward Reverse	 CACGTGGCTCATTCTCTTTGCCATGGCGTTTTAC GTAAAACGCCATGGCAAAGAGAATGAGCCACGTG	 60°C 60°C

CHAPTER 5. CONCLUSION

In this dissertation, I examined the molecular machinery and the evolutionary origin of visual processes in the phylum Mollusca. In chapter II, I expanded on previous molecular datasets to investigate the phylogenetic relationships of families within Pectinoidea, a superfamily of bivalves. My analyses give consistent support for the non-monophyly of the Propeamussiidae, with a subset of species as the sister group to the Pectinidae, the Propeamussiidae type species as sister to the Spondylidae, and the majority of propeamussiid taxa sister to the Spondylidae + *Pr. dalli*. This topology describes previously unidentified relationships among the families of Pectinoidea. Ancestral state estimations using this phylogeny suggest a single origin of eyes in Pectinoidea with multiple instances of loss. These results demonstrate the necessity for a more comprehensive taxonomic sampling from the family Propeamussiidae in order to better understand the evolutionary relationships of pectinoidean families as well as revising the taxonomy of the Propeamussiidae.

In Chapter III, I took a bioinformatic approach to develop a model of the retinoid visual cycle in molluscs. Using known vertebrate and insect retinoid visual cycle proteins as query, I searched 50 publicly available molluscan transcriptomes to identify putative homologs. Then, I used phylogenetic methods to test for homology of the molluscan blast results with vertebrate and insect visual proteins. A major finding of this study is the absence of a homologous retinyl storage pathway, which may support the hypothesis that retinochrome functions as a storage molecule in molluscs rather than, or in addition to, its role as a photoisomerase. This role would allow retinochrome to act as a storage molecule for retinal isomers to alleviate the inherent toxicity to cells of retinal.

Based on my findings of putative homologous retinal dehydrogenases and retinal shuttle proteins, I propose a new retinoid visual cycle pathway similar in structure to that described in insects and vertebrates. Beginning within the photoreceptor cells, the all-*trans* retinal is released from the opsin and transported to the pigment cells by an unknown mechanism. Within the pigment cells, a dehydrogenase homologous to vertebrate RDH13 reduces the all-*trans* retinal to a retinol. The all-*trans* retinol is transported by a PINTA-like shuttle protein to an isomerase, the identity of which is unclear. The newly formed 11-*cis* retinol is then acted upon by a dehydrogenase homologous to several vertebrate dehydrogenases which oxidizes the retinol back to a retinal. The 11-*cis* retinal is shuttled back to the photoreceptor cells and back to the opsin, completing the cycle. This pathway also suggests the presence of visual cycle activity outside the photoreceptor cells of the molluscan eye, but the interphotoreceptor shuttle protein and isomerohydrolase remain unidentified.

In the fourth chapter, I used the spectral analyses and amino acid sequences of two closely related scallop retinochromes to investigate the relationship of genotype and phenotype via spectral tuning of an invertebrate opsin. First, I cloned, expressed, and spectrally analyzed retinochrome from *Argopecten irradians* (Airr-RTC) and *Pecten maximus* (Pmax-RTC). The λ_{max} I observed for these proteins fell within the 490nm to 520nm range of previously described mollusc retinochromes. Using 3D modeling and ligand binding prediction software, I identified three sites of interest between these scallop retinochromes and mutated Airr-RTC to match Pmax-RTC at these three sites to check the effect of mutations at these sites on spectral tuning. My spectral analyses showed that amino acids with hydroxyl groups lining the binding pocket can be used to

fine tune the λ_{max} of photopigments by altering the polar environment of the binding pocket. These results also suggest spectral tuning is altered by the shape of the binding pocket and that conformational changes to the binding pocket by amino acid mutation may be antagonistic or compensated by one another. Alterations in the shape of the binding pocket cause differences in the proximity of amino acid side chains on the retinal chromophore and thereby the λ_{max} of the opsin. However, the effects of changing the shape of the binding pocket still require much work to understand the relationship in shape and absorbance.

My dissertation work illustrates the various hierarchical levels of biological organization that are involved in complex traits, such as vision and the potential for molluscs as model systems for a variety of molecular and evolutionary studies. In order to better clarify the phylogenetic framework necessary to understand the evolution of complex traits within molluscs, future work will require the expansion of taxonomic datasets particularly from underrepresented groups, such as the Propeamussiidae of Bivalvia. While the proposed pathway for a molluscan retinoid visual cycle is a good start towards understanding the molecular pathways of vision in molluscs, there are still many questions about how and where retinal is recycled in molluscan eyes. The hypothesis I pose in this dissertation suggests the presence of visual cycle processes outside the photoreceptor cells of molluscs; however, to date, there is no molecular data specifically analyzing pigment cells of molluscs. The next step should be to focus studies on adjacent cells within the eyes of molluscs to flesh out our understanding of visual systems in the eyes of molluscs. As proposed in chapter IV, it is possible that processes of the visual cycle occur in pigment or adjacent cells similar to that described

in insects. However, expansion of the current visual cycle will require the identification and localization of molecular machinery within cells adjacent to the photoreceptors in molluscan eyes. Finally, mutagenesis experiments investigating the effect of different classes of amino acids at sites of importance will help us understand the relationship between genotype and phenotype of photopigments. There is still much to investigate as to how specific amino acids affect the absorbance or function of opsins at sites important to spectral tuning. Understanding the relationship between genotype and phenotype of opsins will allow the development of more thorough models to infer the function of untested opsins without extensive expression experiments.

APPENDIX A. SUPPLEMENTAL FIGURES AND TABLES FROM CHAPTER 2

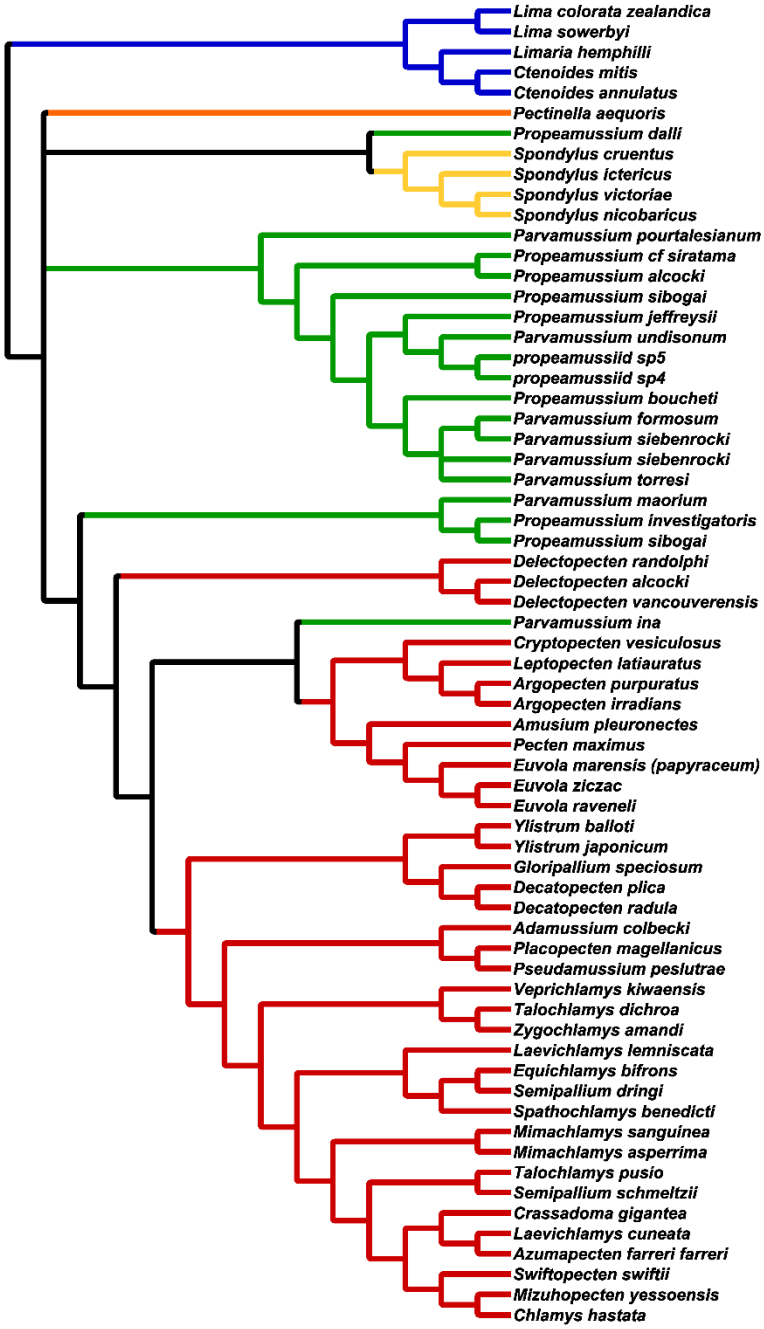


Figure 2-S1 - Bayesian inference majority-rule consensus phylogram based on combined 12S, 16S, 18S, 28S rRNAs and histone H3 sequences. Taxa are color coded by family.

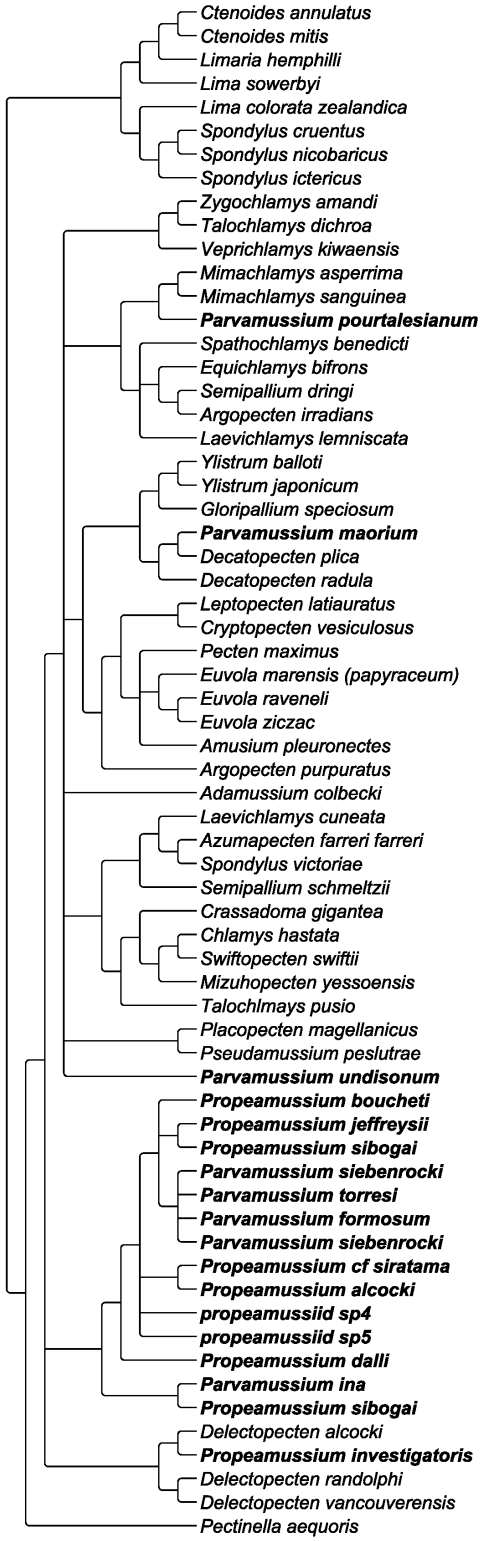


Figure 2-S2 - Maximum likelihood phylogram based on mitochondrial gene sequences (12S, 16S rRNAs). Propeamussiidae sensu lato are in bold.

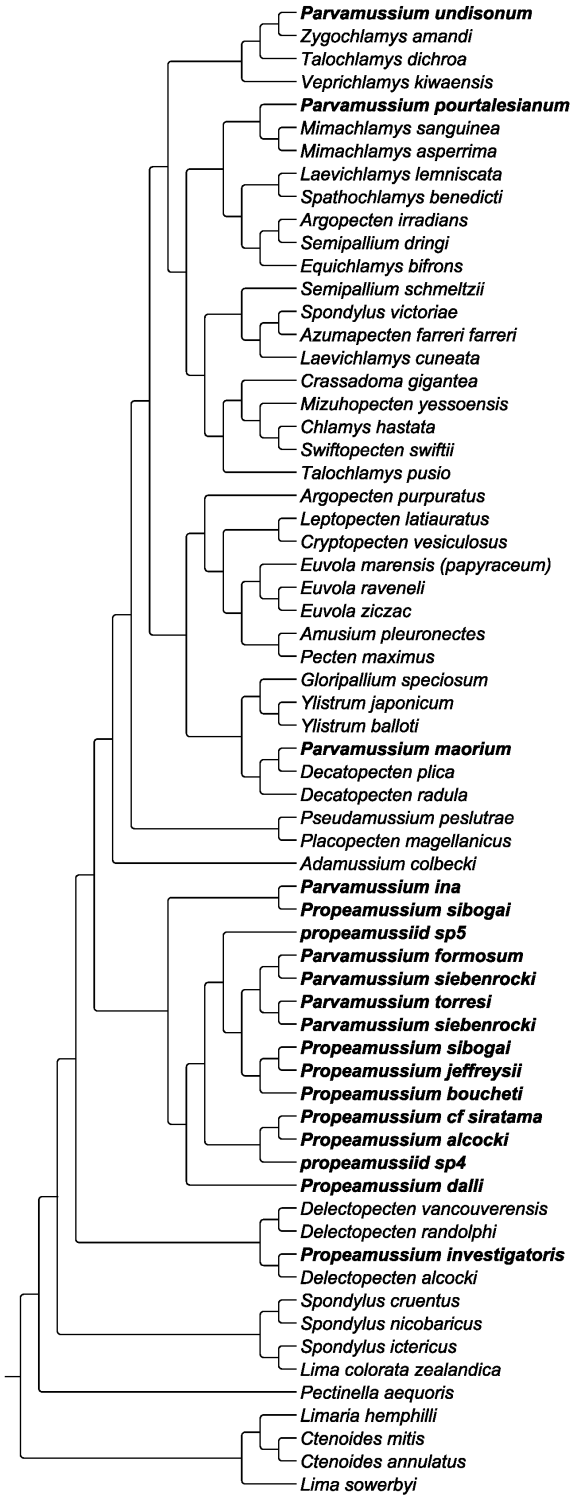


Figure 2-S3 - Bayesian inference majority-rule consensus phylogram based on mitochondrial gene sequences (12S, 16S rRNAs). Propeamussiidae sensu lato are in bold.

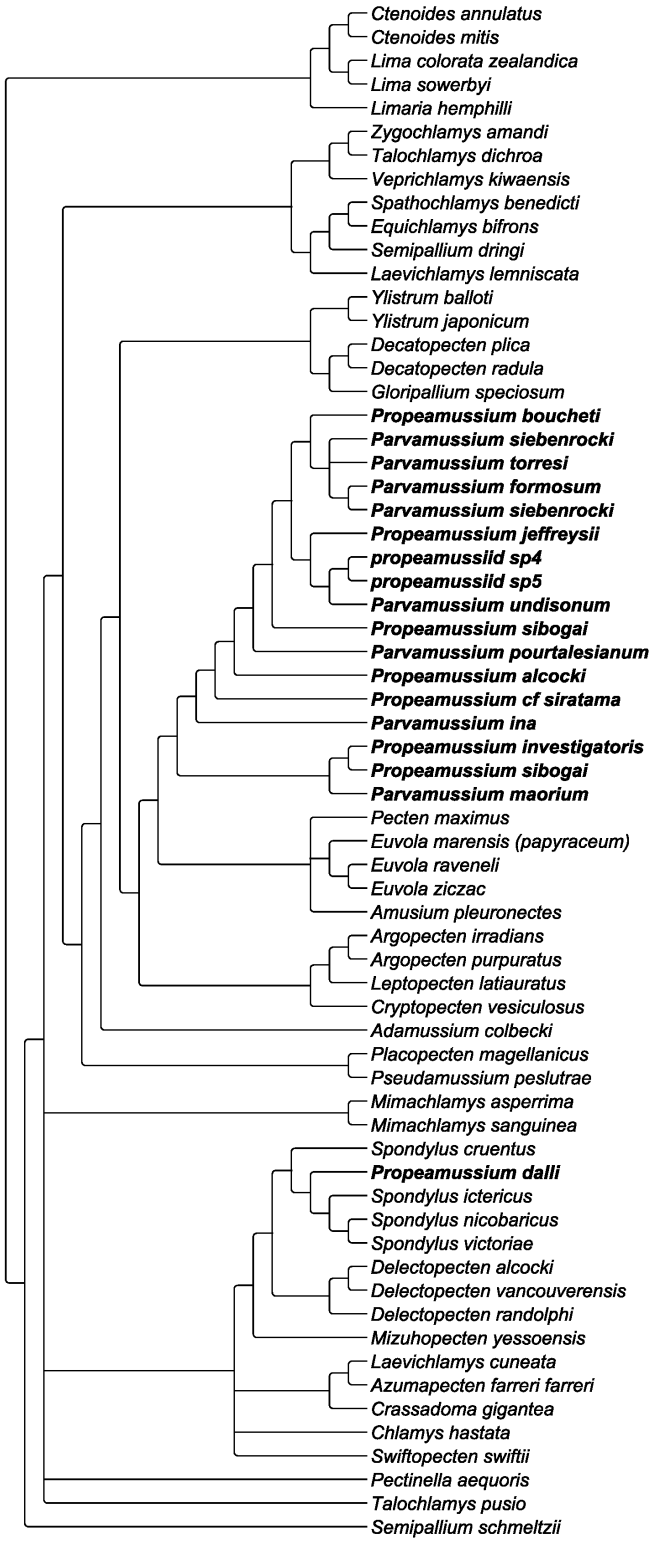


Figure 2-S4 - Maximum likelihood phylogram based on nuclear gene sequences (18S rRNA, 28S rRNA and histone H3). Propeamussiidae sensu lato are in bold.

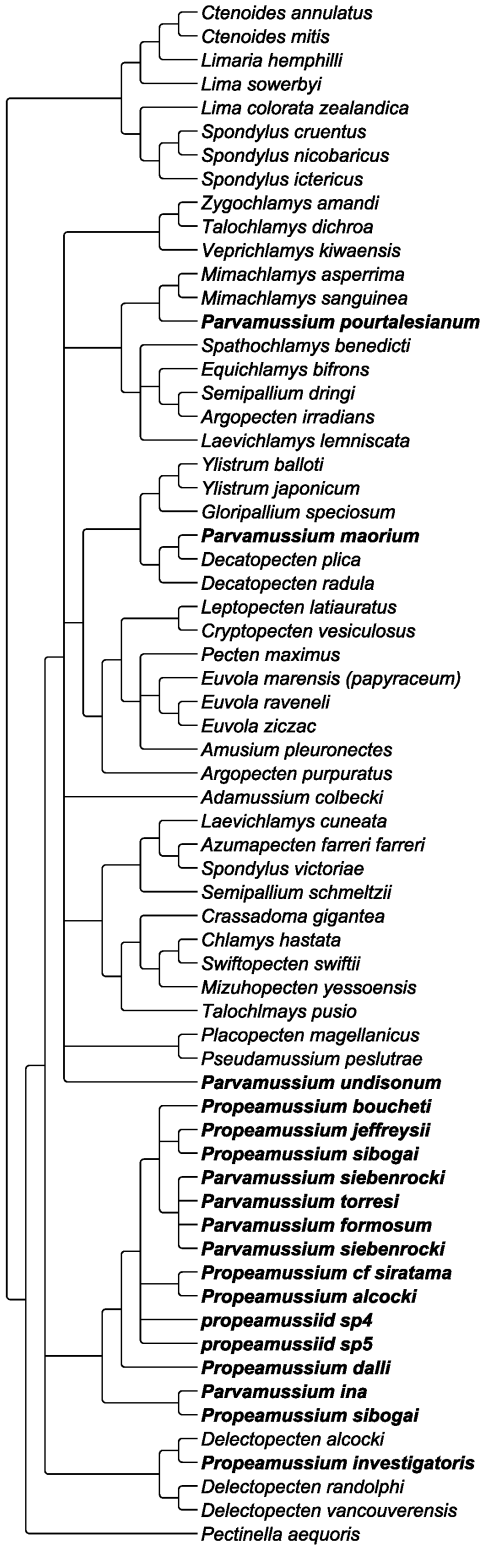


Figure 2-S5 - Bayesian inference majority-rule consensus phylogram based on nuclear gene sequences (18S rRNA, 28S rRNAs and histone H3). Propeamussiidae sensu lato are in bold.

Table 2-S1 Genbank accession numbers for 65 specimens included in the molecular phylogeny. Outgroup species indicated by asterisk (*). The phylogeny ID corresponds to the tip labels of Figure 2 and Figure S1. When available, morphological vouchers are listed by museum and collection accession number: AMNH = American Museum of Natural History; MNHN = Muséum National d'Histoire Naturelle, Paris, France; NIC = National Institute Water and Atmospheric Research, New Zealand Invertebrate Collection; QM = Queensland Museum, Australia; TM = Tepapa Museum, New Zealand; UF = Florida Museum of Natural History, Gainesville, Florida, United States; USC = University of the Sunshine Coast Pectinid Collection, Queensland, Australia; USNM = United States National Museum, Smithsonian Institution; N/A = no morphological specimen available. Names of MNHN and Pro-Natura International expeditions are in all capitals. DNA sequence data generated in this study are given as Genbank accession numbers in bold.

Genus	12S rRNA Genbank	16S rRNA Genbank	18S rRNA Genbank	28SrRNA Genbank	Histone H3 Genbank	Locality	Morphological Specimen
Pectinidae							
Adamussium colbecki (E.A. Smith, 1902)	EU379383	EU379437	**MH464058**	FJ263652	EU379491	Terra Nova Bay, Antarctica	USNM 1532171
Amusium pleuronectes (Linnaeus, 1758)	EU379415	EU379469	**MH464085**	HM630508	EU379523	Rayong Province, Thailand	USNM 1236642
Argopecten irradians (Lamarck, 1819)	EU379392	EU379446	**MH464072**	HM622700	EU379500	Gulf of Mexico, FL	USNM 1532203
Argopecten purpuratus (Lamarck, 1819)	EU379417	EU379471	**MH464088**	HM630495	EU379525	Tongo Bay, Chile	N/A
Azumapecten farreri (Jones & Preston, 1904)	HM622677	HM622678	**MH464066**	HM622680	HM622679	Aquaculture Facility in Qindao, china	USNM 1532174
Chlamys hastata (G.B. Sowerby II, 1842)	FJ263639	FJ263648	**MH464068**	FJ263658	FJ263667	San Juan Island, Washington, USA	USNM 1532198
Crassadoma gigantea (J.E. Gray, 1825)	FJ263635	FJ263644	**MH464067**	FJ263654	FJ263664	Santa Barbara, California, USA	N/A
Cryptopecten vesiculosus (Dunker, 1877)	HM630403	HM630404	**MH464108**	HM630406	HM630405	Miura City, Japan	USNM 1532242

Decatopecten plica (Linnaeus, 1758)	HM630435	HM630436	**MH464086**	HM630438	HM630437	Tateyama, Japan	USNM 1532233
Decatopecten radula (Linnaeus, 1758)	MH463999	HM630492		HM630494	HM630493	Sulawsi Island, Indonesia	UF 280376
Delectopecten alcocki (E.A. Smith, 1902)	**MH463998**	**MH464013**	**MH464062**	**MH464039**	**MH464025**	Bohol Island, Maribohoc Bay	MNHN IM-2007-33918
Delectopecten randolphi (Dall, 1897)	HM630488	HM630489	**MH464090**	HM630491	HM630490	Hitachi City, Japan	N/A
Delectopecten vancouverensis (Whiteaves, 1893)	HM630418	HM630420	**MH464107**	HM630417	HM630416	North Pacific Ocean; 32°36'N; 117°30.5'W	Scripps Inst Oceanography
Equichlamys bifrons (Lamarck, 1819)	HM561991	HM561992	**MH464055**	HM561994	HM561993	Tasmania	USNM 1532170
Euvola marensis (=papyraceum) (Weisbord, 1964)	HM630371	HM630372	**MH464084**	HM630374	HM630373	Gulf of Mexico, USA	TCWC 40985
Euvola raveneli (Dall, 1898)	EU379419	EU379473	**MH464091**	HM630487	EU379527	Gulf of Mexico off of St. Petersburg	UF 351301
Euvola ziczac (Linnaeus, 1758)	EU379430	EU379484	**MH464109**	HM630509	EU379538	Harrington Sound, Bermuda	Serb lab
Gloripallium speciosum (Reeve, 1853)	HM630465	HM630466	**MH464097**	HM630468	HM630467	Viti Levu Island, Fiji	UF 292110
Laevichlamys cuneata (Reeve, 1853)	HM622702	HM622703	**MH464061**	HM622705	HM622704	Tateyama City, Japan	USNM 1532172
Laevichlamys lemniscata (Reeve, 1853)	KP300588	KP300554	**MH464077**	KP300523	KP300491	ATIMO VATAE, Port Ehoala, Madagascar	MNHN IM-2009-21008
Leptopecten latiauratus (Conrad, 1837)	EU379393	EU379447	**MH464076**	HM622714	EU379501	Goleta Pier, California, USA	USNM 1532209
Mimachlamys asperrima **(**Lamarck. 1819)	HM540080	HM540081	**MH464052**	HM540083	HM540082	Hobart, Tasmania, Australia	USNM 1532169

Mimachlamys sanguinea (Linnaeus, 1758)	HM630479	HM630480	**MH464092**	HM630482	HM630481	Thailand	Serb lab
Mizuhopecten yessoensis (Jay, 1857)	HM630383	HM630384		HM630386	HM630385	Mutsu Bay, Japan	USNM 1532242
Pecten maximus (Linnaeus, 1758)	EU370400	EU379454	**MH464079**	HM630545	EU379508	Millport, Scotland	USNM 1532228
Placopecten magellanicus (Gmelin, 1791)	FJ263638	FJ263647	**MH464078**	JF263657	EU379647	Georges Bank, USA	USNM 1532222
Pseudamussium peslutrae (=septemradiatus) (Linnaeus, 1771)	EU379420	EU379474	**MH464094**	FJ263659	EU379528	Millport, Scotland	USNM 1532236
Semipallium dringi (Reeve, 1853)	EU379387	EU379441	**MH464065**	HM622672	EU379495	Ie Island, Okinawa, Japan	UF 352373
Semipallium schmeltzii (Dunker in Küster & Kobelt, 1888)	HM630483	HM630484	**MH464093**	HM630486	HM630485	Maruki hama, Bonotsu City, Japan	N/A
Spathochlamys benedicti (Verrill & Bush, 1897)	HM540103	HM540104	**MH464054**	HM540106	HM540105	W of Cedar Key, Levy Co, Florida	UF 369432
Swiftopecten swiftii (Bernardi, 1858)	KP300599	KP300565	**MH464105**	KP300532	KP300502	Japan	USNM 1532240
Talochlamys dichroa (Suter, 1909)	KP300577	KP300543	**MH464064**	KP300514	KP300480	Otago Peninsula, NZ	TM M297698
Talochlamys pusio (Linnaeus, 1758)	HM600764	HM600757	**MH464089**	HM600750	HM600737	Raxo, Ria de Pontevedra, Gallicia, Spain	USNM 1532235
Veprichlamys kiwaensis (Powell, 1933)	KP300586	KP300552	**MH464075**	KP300521	KP300489	Louisiville Ridge, Chatham Rise, NZ	NIC TAN0707/ 84
Ylistrum balloti (Bernardi, 1861)	HM4540088	HM540089	**MH464053**	HM540091	HM540090	Bunderberg, Queensland, Australia	USNM 1236641
Ylistrum japonicum (Gmelin, 1791)	HM622706	HM622707	**MH464073**	HM622709	HM622708	Oyano Island, Kumamoto, Japan	USNM 1236649

Zygochlamys amandi (Soot-Ryen, 1959)	HM485575	HM485576	**MH464050**	HM485578	HM485577	Puerto Montt, Chile	N/A
Propeamussiidae							
Parvamussium formosum (Melvill & Standen, 1907)	**MH464007**	**MH464020**	**MH464100**	**MH464044**	**MH464033**	MIRIKY - entre Nosy-be et Banc du Leven, 12°46'S; 48°11'E	MNHN IM-2007-38444
Parvamussium ina (Dautzenberg & Bavay, 1912)	**MH464002**	**MH464015**			**MH464028**	SALOMON2, 8°41' S; 157°24' E	MNHN IM-2007-33924
Parvamussium maorium Dell, 1956	KP300590	KP300556			KP300493	UTM - 42.7871700, -176.7222000, New Zealand	NIC 22965
Parvamussium pourtalesianum (Dall, 1886)	EU379411	EU379465	**MH464087**	HM600741	EU379519	Florida Straits, Florida, USA	UF 323764
Parvamussium puillandrei (Dijkstra & Maestrati, 2015)	**MH464009**	**MH464022**	**MH464102**	**MH464046**	**MH464035**	MIRIKY, Ouest du Cap d'Ambre, 12°06'S; 48°54'E	MNHN IM-2007-39021
Parvamussium siebenrocki (Sturany, 1901)	**MH464005**	**MH464018**	**MH464098**	**MH464042**	**MH464031**	MAINBAZA, Inhambane transect, Mozambique Channel, 23°31'S; 35°50'E	MNHN IM-2007-38282
Parvamussium siebenrocki (Sturany, 1901)	**MH464010**	**MH464023**	**MH464103**	**MH464047**	**MH464036**	MAINBAZA, Inhambane transect, Mozambique Channel, 23°31'S; 35°50'E	MNHN IM-2007-38307
Parvamussium torresi (E.A. Smith, 1885)	**MH464006**	**MH464019**	**MH464099**	**MH464043**	**MH464032**	MAINBAZA, Inhambane transect,	MNHN IM-2007-38296

						Mozambique Channel, 23°31'S; 35°50'E	
Parvamussium undisonum Dijkstra, 1995	**MH464012**		**MH464106**			SALOMON2, SE Isabel, 8°17' S;160°00' E	MNHN IM-2007-33931
Propeamussium alcocki (E.A. Smith, 1894)	KP300572	KP300537	**MH464083**		KP300474	AURORA 2007, 14°50'N; 123°12'E, Philippines	MNHN IM-2007-33735
Propeamussium boucheti Dijkstra & Maestrati, 2008	**MH464000**		**MH464056**		**MH464026**	EBISCO, Coral Sea, 21°06'S; 158°32'E	MNHN IM-2007-33932
Propeamussium cf siratama (Oyama in Kuroda, 1951)	**MH464001**	**MH464014**	**MH464057**		**MH464027**	SALOMON2, 6°55' S; 156°21' E	MNHN IM-2007-33901
Propeamussium dalli (E.A. Smith, 1885)	EU379416	EU379470	**MH464063**	HM600740	EU379524	Dry Tortugas, Florida, USA	UF 289879
Propeamussium investigatoris (E.A. Smith, 1906)	**MH464003**	**MH464016**	**MH464071**	**MH464040**	**MH464029**	SALOMON2, 8°17' S; 160°00' E	MNHN IM-2007-33930
Propeamussium jeffreysii (E.A. Smith, 1885)	**MH464004**	**MH464017**	**MH464074**	**MH464041**	**MH464030**	EBISCO, Coral Sea, 23°55'S; 161°44'E	MNHN IM-2007-33906
Propeamussium sibogai (Dautzenberg & Bavary, 1904)	**MH464011**	**MH464024**	**MH464104**	**MH464048**	**MH464037**	MIRIKY, entre Nosy-be et Banc du Leven, 12°40'S; 48°12'E	MNHN IM-2007-38389
Propeamussium sibogai (Dautzenberg & Bavary, 1904)	HM600762	HM600755	**MH464095**	HM600748	HM600735	NW of Nomamishiki, Kasasa-cho, Japan	USNM 1532237
Propeamussiidae spp4	**MH464008**	**MH464021**	**MH464101**	**MH464045**	**MH464034**	TARASOC, Huahine,	MNHN IM-2007-38558

						16°43'S; 150°38'W	
Spondylidae							
Spondylus cruentus Lischke, 1868	HM600761	HM600754	**MH464060**	HM600747	HM600734	Tateyama City, Japan	N/A
Spondylus ictericus Reeve, 1856	EU379423	EU379477	**MH464070**	HM600742	EU379531	Florida Keys, Florida, USA	UF 367487
Spondylus nicobaricus Schreibers, 1793	EU379424	EU379478	**MH464081**	HM600743	EU379532	West of New Briton, Papua New Guinea	UF 322550
Spondylus victoriae (=wrightianus) G.B. Sowerby II, 1860	KP300606	KP300571			KP300508	Stradbroke IS, QLD, Australia	USC SCALLOG23
Entoliidae							
Pectinella aequoris Dijkstra, 1991			**MH464082**	**MH464049**	**MH464038**	PANGLAO 2004	MNHN IM-2007-33872
Limidae							
*Ctenoides annulatus** (Lamarck, 1819)	EU379385	EU379439	**MH464051**	HM535655	EU379493	Bismark Archipelago, Papua New Guinea	UF 322180
*Ctenoides mitis** (Lamarck, 1807)	EU379386	EU379440	**MH464080**	HM600745	EU379494	Florida Keys, Florida, USA	UF 367478
*Lima colorata zealandica** G.B. Sowerby III, 1877	HM600760	HM600753	**MH464059**	HM600746	HM600733	North Cape, New Zealand	UF 332786
*Lima sowerbyi** Deshayes, 1863	HM600763	HM600756	**MH464096**	HM600749	HM600736	Masirah Island, Oman	UF 286387
*Limaria hemphilli** (Hertlein & A.M. Strong, 1946)	KP300584	KP300550	**MH464069**		KP300487		N/A

APPENDIX B. SUPPLEMENTAL FIGURES AND TABLES FROM CHAPTER 3

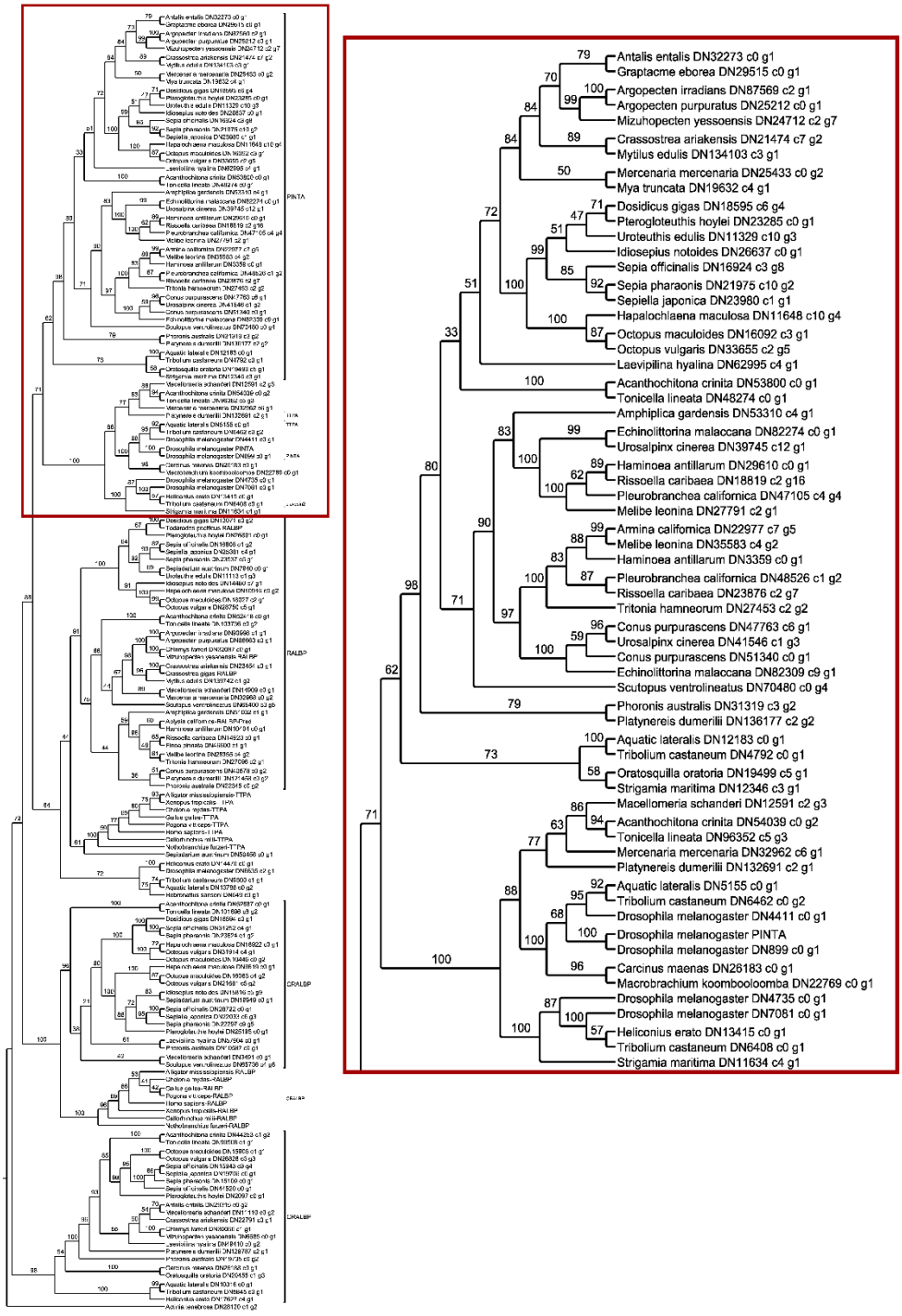


Figure 3-S1.1 – Maximum Likelihood phylogeny of CRAL-TRIO protein family. Bootstrap support values are labeled at nodes. Tree is magnified with tree region highlighted by respective colored boxes.

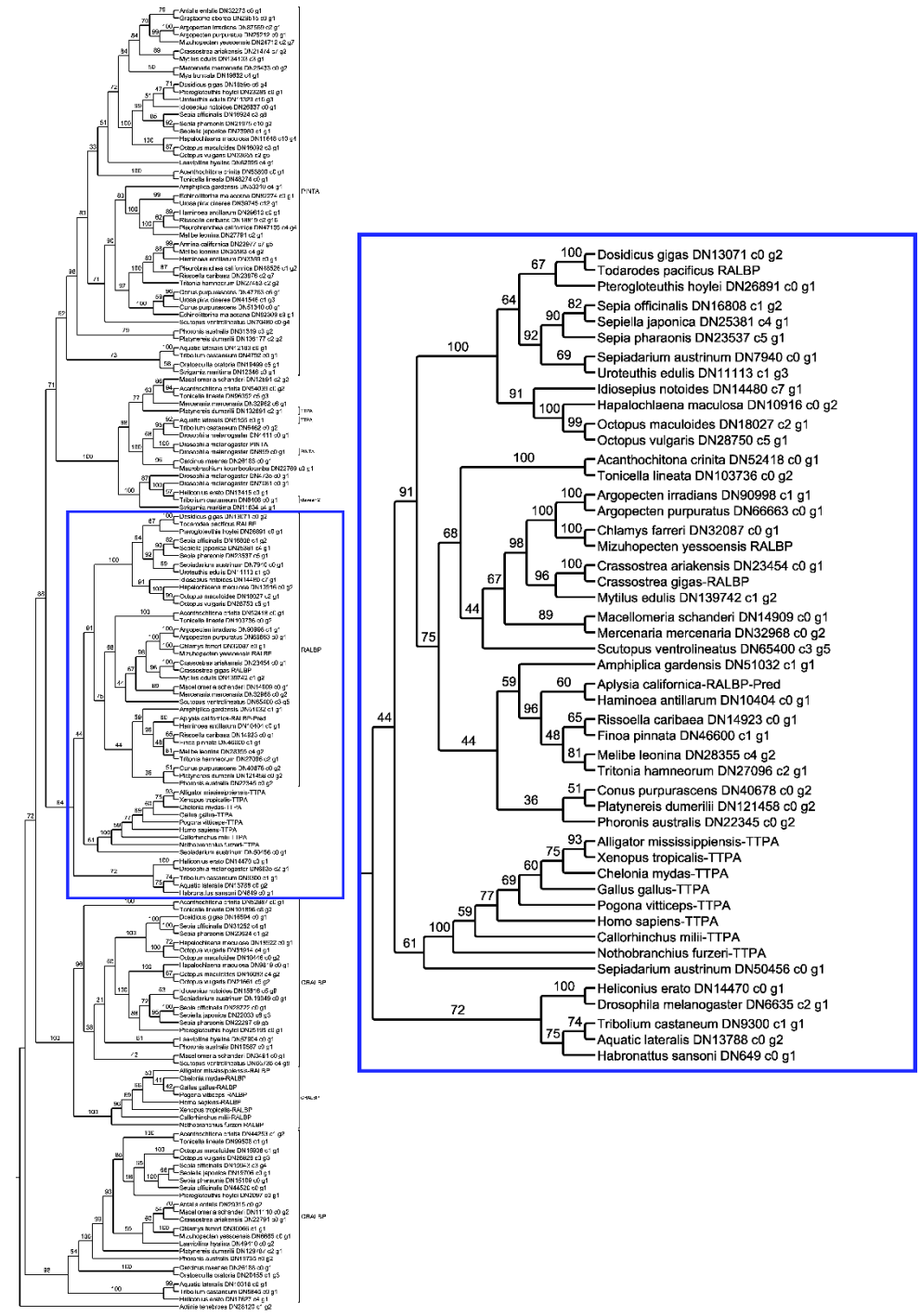


Figure 3-S1.2 – Maximum Likelihood phylogeny of CRAL-TRIO protein family. Bootstrap support values are labeled at nodes. Tree is magnified with tree region highlighted by respective colored boxes.

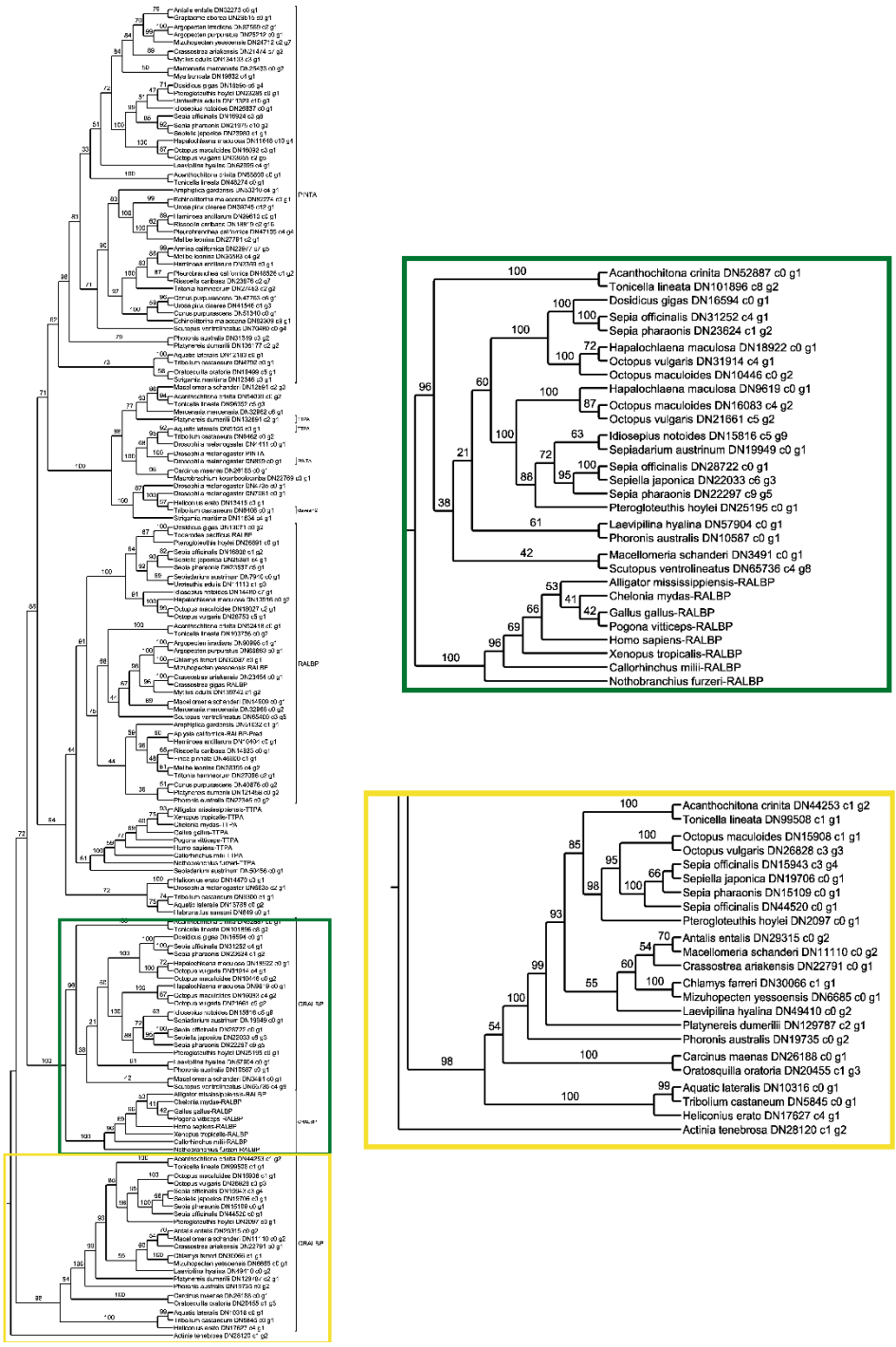


Figure 3-S1.3 – Maximum Likelihood phylogeny of CRAL-TRIO protein family. Bootstrap support values are labeled at nodes. Tree is magnified with tree region highlighted by respective colored boxes.

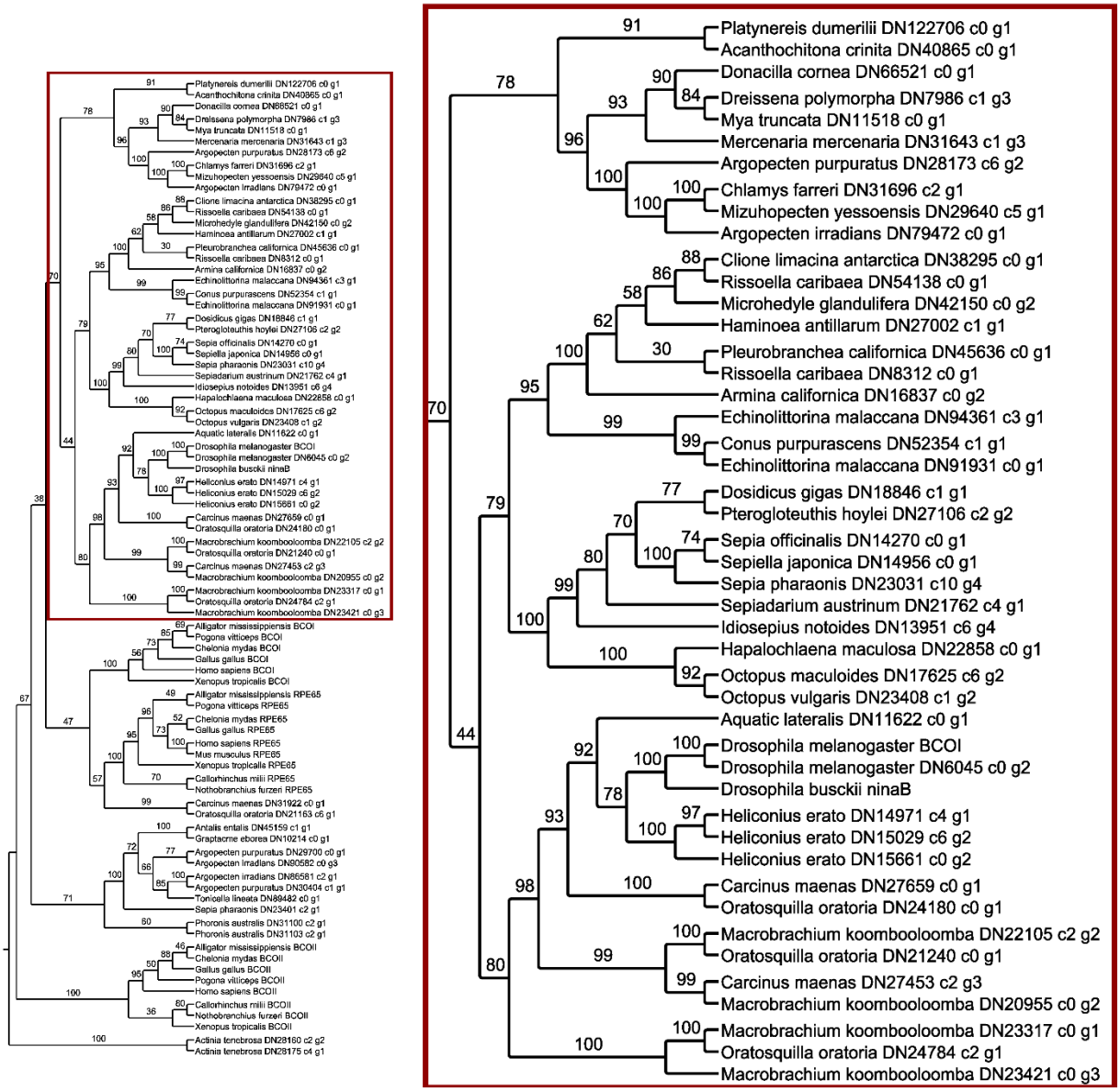


Figure 3-S2.1 – Maximum Likelihood phylogeny of carotenoid oxygenase protein family. Bootstrap support values are labeled at nodes. Tree is magnified with tree region highlighted by respective colored boxes.

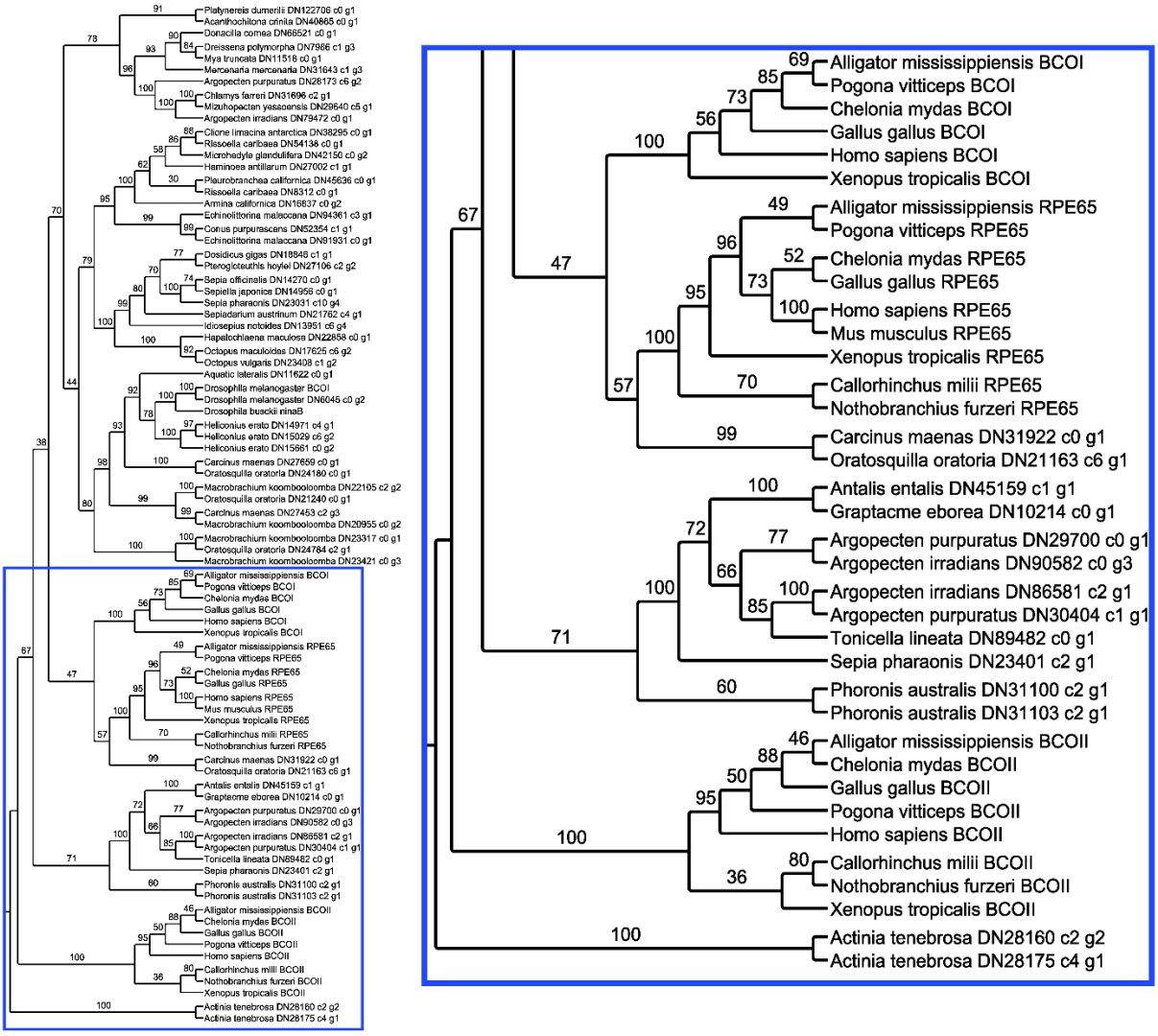


Figure 3-S2.2 – Maximum Likelihood phylogeny of carotenoid oxygenase protein family. Bootstrap support values are labeled at nodes. Tree is magnified with tree region highlighted by respective colored boxes.

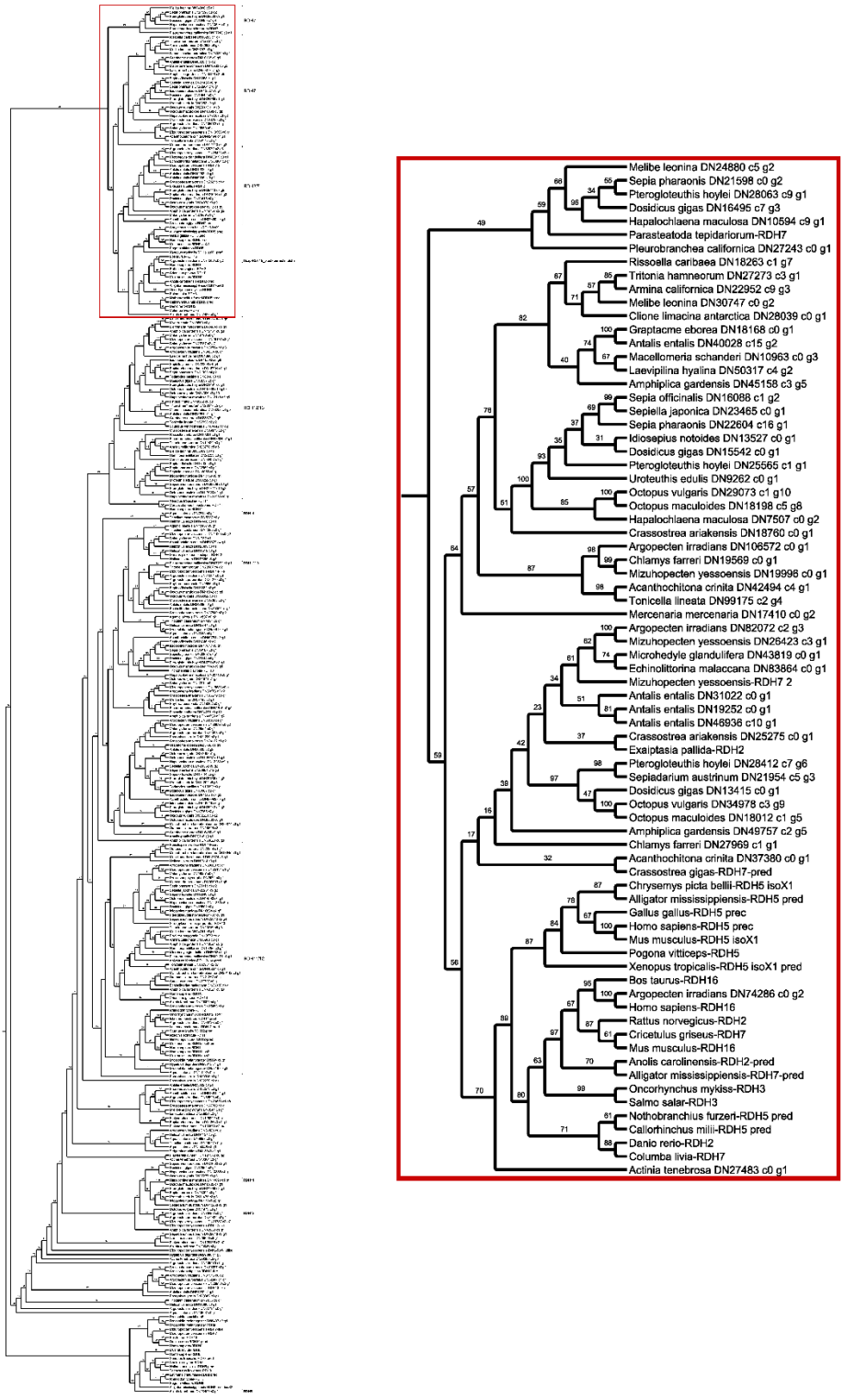


Figure 3-S3.1 – Maximum Likelihood phylogeny of short-chain dehydrogenase protein family. Bootstrap support values are labeled at nodes. Tree is magnified with tree region highlighted by respective colored boxes.

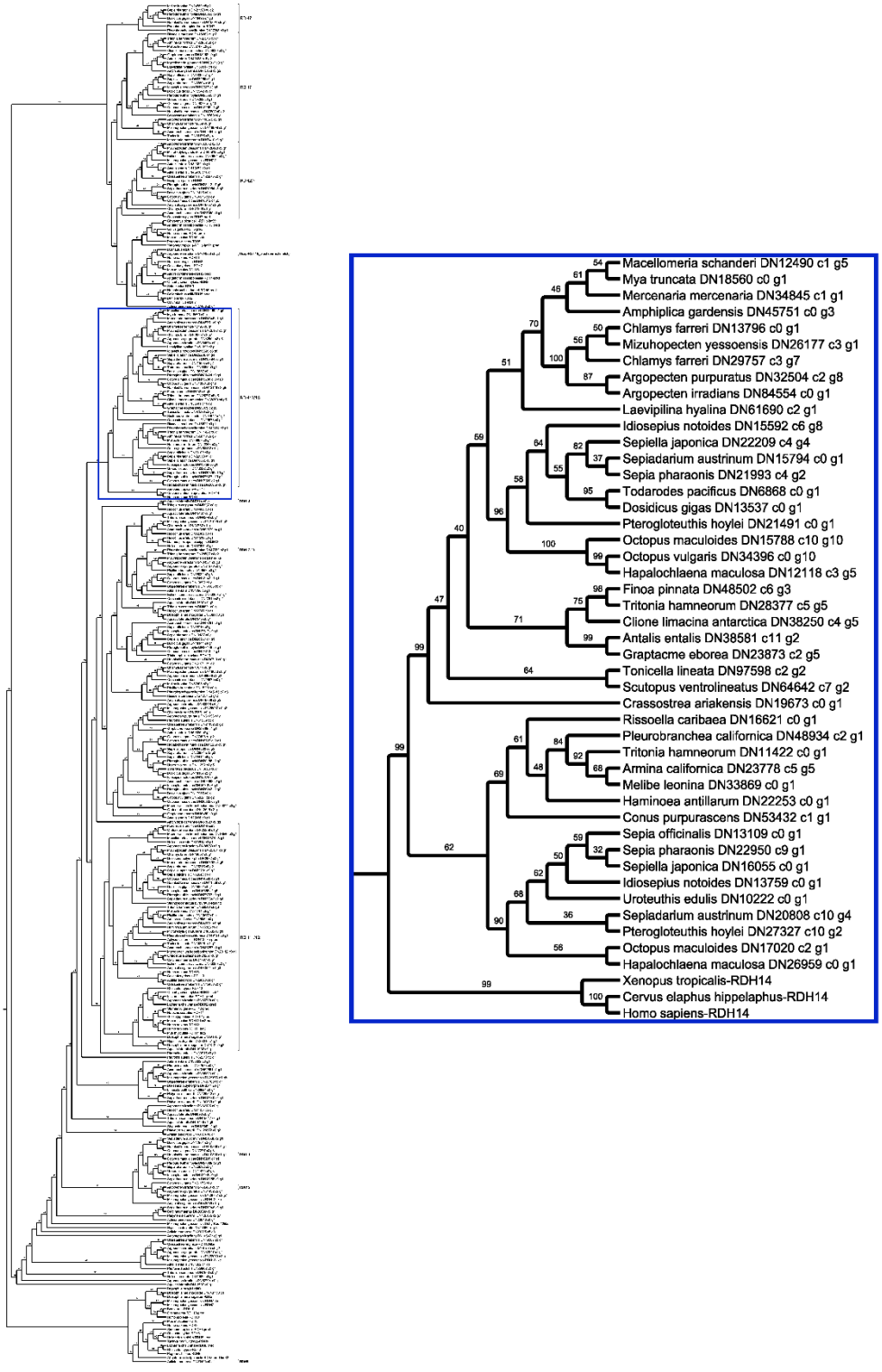


Figure 3-S3.2 – Maximum Likelihood phylogeny of short-chain dehydrogenase protein family. Bootstrap support values are labeled at nodes. Tree is magnified with tree region highlighted by respective colored boxes.

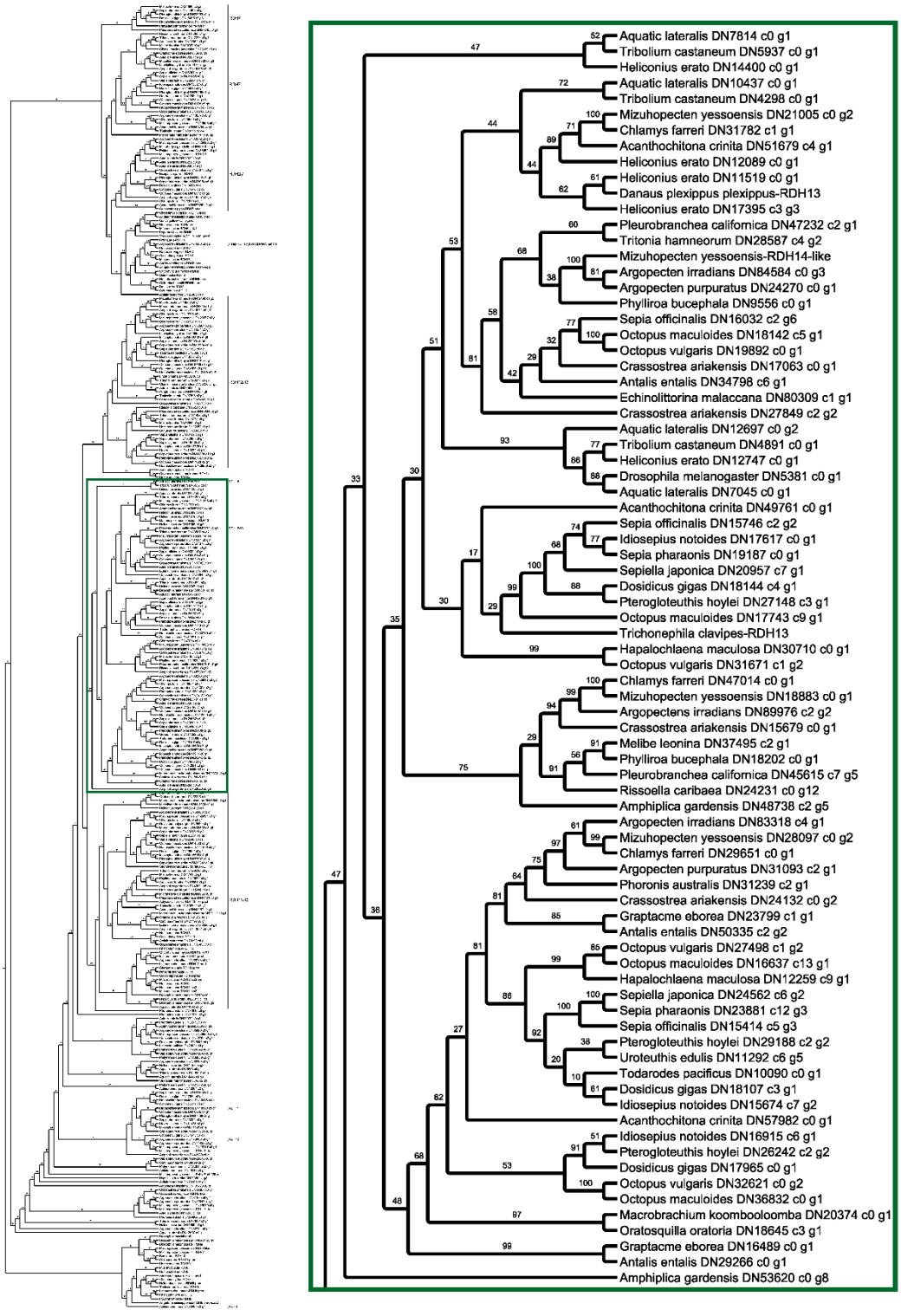


Figure 3-S3.3 – Maximum Likelihood phylogeny of short-chain dehydrogenase protein family. Bootstrap support values are labeled at nodes. Tree is magnified with tree region highlighted by respective colored boxes.

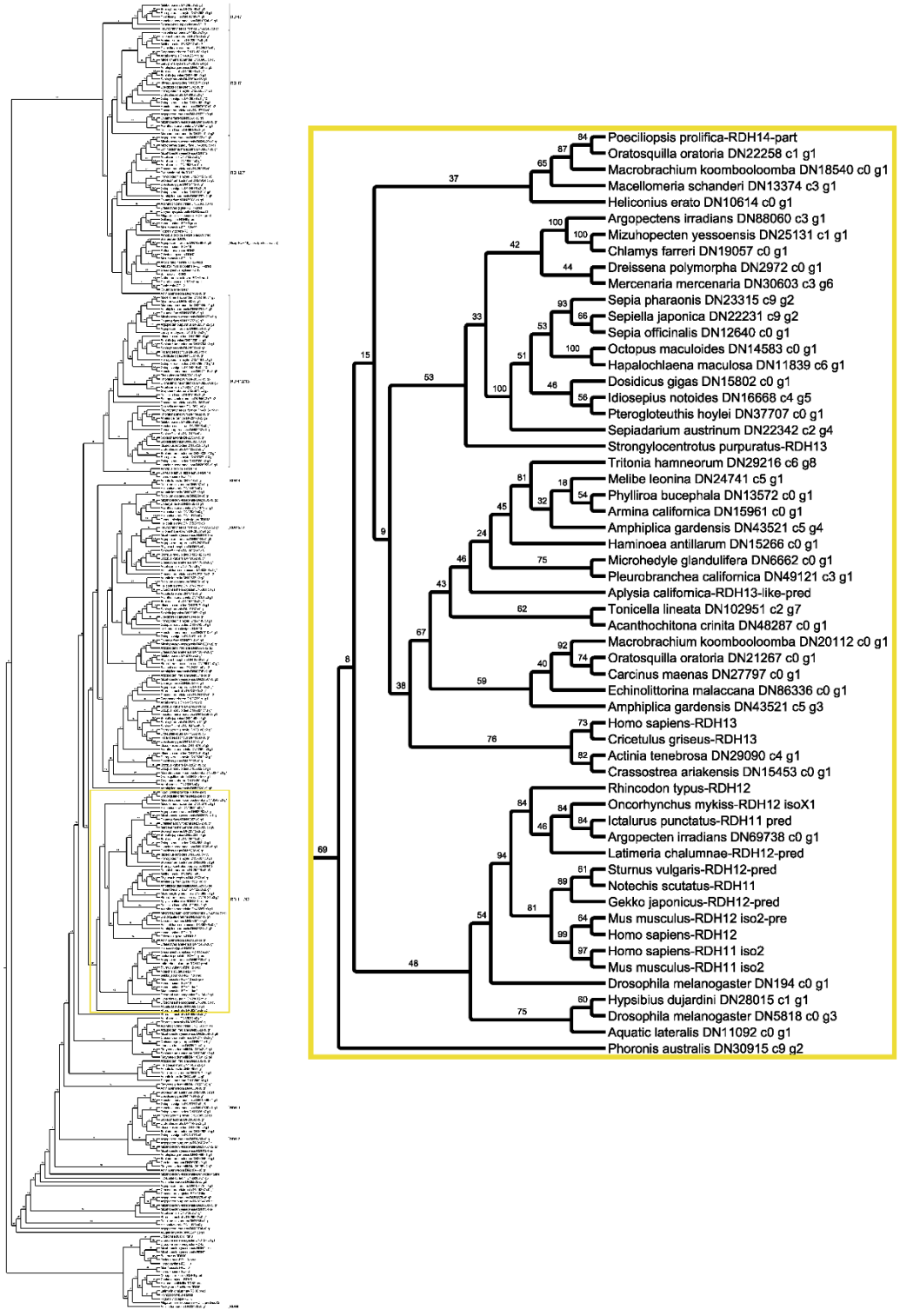


Figure 3-S3.4 – Maximum Likelihood phylogeny of short-chain dehydrogenase protein family. Bootstrap support values are labeled at nodes. Tree is magnified with tree region highlighted by respective colored boxes.

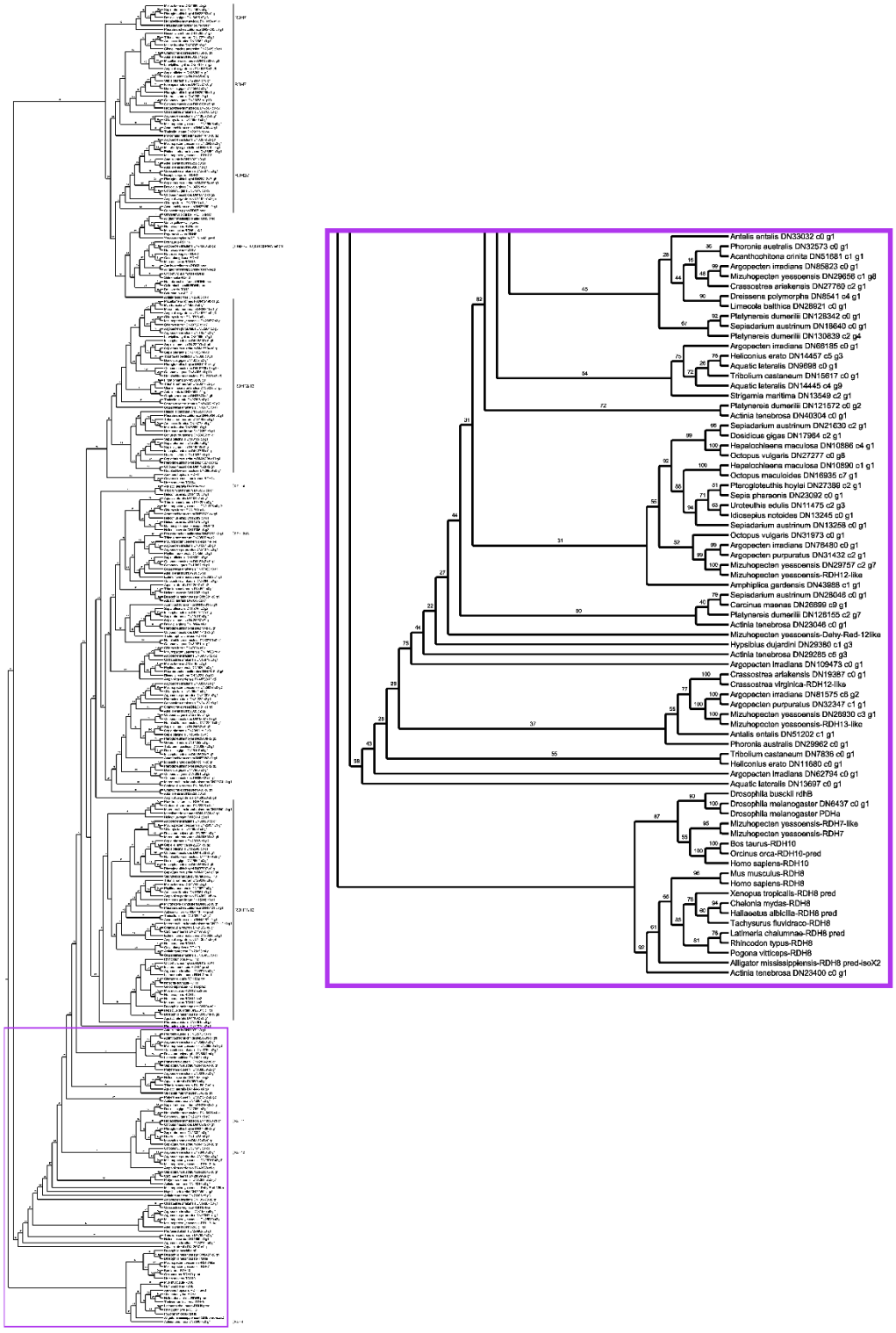


Figure 3-S3.5 – Maximum Likelihood phylogeny of short-chain dehydrogenase protein family. Bootstrap support values are labeled at nodes. Tree is magnified with tree region highlighted by respective colored boxes.

Table 3-S1 - Transcriptome identification information and assembly quality statistics.

Species	Class	Phylum	SRA#	Bowtie2	BUSCO	Ex90N50	N50
Macellomeria schanderi	Mollusca	Aplacophora	SRR2057023	87.95%	C:89.9(S:19.1)%	735/5605	Avg=917.5/1540
Scutopus ventrolineatus	Mollusca	Aplacophora	SR5110532	88.92%	C:73.7(S:38.1)%	572/66237	Avg=478/549
Adamussium colbecki	Mollusca	Bivalvia	SRR5349700	88.86%	C:17.4(S:13.8)%	395/18105	Avg=400.67/414
Argopecten irradians	Mollusca	Bivalvia	SRR5469239	93.42%	C:96.5(S:36.2)%	2091/26528	Avg=920.67/1757
Argopecten pupuratus	Mollusca	Bivalvia	SRR6849473	94.96%	C:97.1(S:15.0)%	1502/66469	Avg=1021.32/1706
Chlamys farreri	Mollusca	Bivalvia	SRR5130889	92.66%	C:86.1(S:58.5)%	2209/65222	Avg=1111.67/2100
Crassostrea ariakensis	Mollusca	Bivalvia	SRR6423968	94.00%	C:97.6(S:52.4)%	2117/45860	Avg=1079.11/1982
Donacilla cornea	Mollusca	Bivalvia	SRR1560084	90.41%	C:26.9(S:21.9)%	710/27608	Avg=482.15/569
Dreissena polymorpha	Mollusca	Bivalvia	SRR5000302	88.06%	C:38.1(S:35.4)%	695/18058	Avg=504.97/587
Limecola balthica	Mollusca	Bivalvia	SRR5758183	87.58%	C:34.0(S:23.5)%	852/13486	Avg=539.3/671
Mercenaria campechiensis	Mollusca	Bivalvia	SRR1560359	84.38%	C:18.5(S:12.5)%	411/22989	Avg=387.37/401
Mercenaria mercenaria	Mollusca	Bivalvia	SRR3095228	90.97%	C:92.3(S:41.8)%	1621/93583	Avg=845.69/1430
Mizuhopecten yessoensis	Mollusca	Bivalvia	SRR1185949	91.01%	C:92.8(S:63.6)%	1232/11967	Avg=995.68/1708
Mya truncata	Mollusca	Bivalvia	SRR5945865	90.37%	C:64.1(S:49.5)%	1029/36571	Avg=633.63/860
Mytilus edulis	Mollusca	Bivalvia	SRR6873090	74.88%	C:45.0(S:26.8)%	419/67921	Avg=391.29/414
Sphaerium nucleus	Mollusca	Bivalvia	SRR1561723	87.24%	C:32.7(S:17.0)%	448/139998	Avg=383.2/394
Dosidicus gigas	Mollusca	Cephalopoda	SRR5152122	94.55%	C:90.5(S:66.4)%	2248/6425	Avg=1027.24/1981
Hapalochlaena maculosa	Mollusca	Cephalopoda	SRR3105555	95.13%	C:91.9(S:69.3)%	1807/25275	Avg=782.98/1334
Idiosepius notoides	Mollusca	Cephalopoda	SRR2984343	92.66%	C:81.4(S:55.0)%	956/33260	Avg=556.15/698
Nautilus pompilius	Mollusca	Cephalopoda	SRR2857280	89.08%	C:85.2(S:26.5)%	1421/135641	Avg=745.15/1101
Octopus maculoides	Mollusca	Cephalopoda	SRR2047116	95.20%	C:96.1(S:68.8)%	1875/46734	Avg=821.19/1483
Octopus vulgaris	Mollusca	Cephalopoda	SRR2857274	86.31%	C:86.2(S:33.8)%	958/109466	Avg=543.28/662
Pterogioteuthis hoylei	Mollusca	Cephalopoda	SRR5527418	94.53%	C:94.2(S:64.6)%	1876/4083	Avg=987.88/1909

Sepia officinalis	Mollusca	Cephalopoda	SRR2856422	84.71%	C:84.3(S:36.1)%	713/136525	Avg=442.14/479
Sepia officinalis	Mollusca	Cephalopoda	SRR5204439	89.04%	C:91.2(S:62.7)%	1630/38743	Avg=767.93/1258
Sepia pharaonis	Mollusca	Cephalopoda	SRR3011300	93.72%	C:92.7(S:70.9)%	1708/55827	Avg=765.23/1346
Sepiadarium austrinum	Mollusca	Cephalopoda	SRR2973271	92.61%	C:84.7(S:60.7)%	1189/29879	Avg=595.2/797
Sepiella japonica	Mollusca	Cephalopoda	SRR2891216	93.01%	C:93.2(S:65.1)%	1388/68739	Avg=682.56/1061
Todarodes pacificus	Mollusca	Cephalopoda	SRR3472305	95.21%	C:68.8(S:60.1)%	1584/14918	Avg=692.88/1057
Uroteuthis edulis	Mollusca	Cephalopoda	SRR3498545	92.69%	C:87.5(S:66.9)%	1236/3372	Avg=661.87/989
Amphiplica gardensis	Mollusca	Gastropoda	SRR1505101	75.05%	C:56.6(S:32.2)%	394/95597	Avg=394.77/417
Armina californica	Mollusca	Gastropoda	SRR4124996	92.34%	C:96.3(S:58.9)%	1549/46810	Avg=817.1/1382
Chaetoderma sp.	Mollusca	Gastropoda	SRR1505105	46.07%	C:21.8(S:12.8)%	322/21519	Avg=322.84/320
Clione limacina antarctica	Mollusca	Gastropoda	SRR1505107	85.31%	C:76.7(S:42.8)%	1305/9007	Avg=767.12/1210
Conus purpurascens	Mollusca	Gastropoda	SRR6784970	92.24%	C:84.7(S:56.6)%	1582/58037	Avg=821.96/1392
Echinolittorina malaccana	Mollusca	Gastropoda	SRR1269556	83.86%	C:69.5(S:35.6)%	525/114790	Avg=478.15/536
Finoa pinnata	Mollusca	Gastropoda	SRR1505109	84.49%	C:72.4(S:52.0)%	873/20152	Avg=625.02/834
Haminoea antillarum	Mollusca	Gastropoda	SRR1505111	78.81%	C:73.3(S:57.0)%	931/7002	Avg=590.86/771
Melibe leonina	Mollusca	Gastropoda	SRR3738852	89.23%	C:95.0(S:48.9)%	1136/75244	Avg=695.75/1061
Microhedyle glandulifera	Mollusca	Gastropoda	SRR1505118	56.77%	C:31.4(S:21.3)%	356/140545	Avg=362.82/370
Phylliroa bucephala	Mollusca	Gastropoda	SRR5527414	95.59%	C:80.8(S:52.0)%	1396/5164	Avg=736.27/1117
Pleurobranchea californica	Mollusca	Gastropoda	SRR3928990	93.73%	C:95.5(S:59.5)%	2327/1817	Avg=843.83/1817
Rissoella caribaea	Mollusca	Gastropoda	SRR1505135	80.32%	C:84.4(S:41.9)%	825/95407	Avg=588.18/763
Tritonia hamneorum	Mollusca	Gastropoda	SRR4190242	94.21%	C:95.8(S:67.2)%	1354/79187	Avg=753.33/1172
Urosalpinx cinerea	Mollusca	Gastropoda	SRR1505141	86.79%	C:52.9(S:42.7)%	721/23931	Avg=530.31/656
Laevipilina hyalina	Mollusca	Monoplacophora	SRR1505115	78.59%	C:69.2(S:34.5)%	1582/58037	Avg=821.96/1392
Acanthochitona crinita	Mollusca	Polyplacophora	SRR5110524	90.20%	C:95.0(S:23.4)%	1460/71522	Avg=822.9/1301
Tonicella lineata	Mollusca	Polyplacophora	SRR6926331	91.73%	C:95.6(S:53.5)%	1369/152487	Avg=635.7/934
Antalis entalis	Mollusca	Scaphopoda	SRR5110529	79.62%	C:93.6(S:43.9)%	567/122966	Avg=487.24/553
Graptacme eborea	Mollusca	Scaphopoda	SRR2057020	79.48%	C:81.5(S:36.7)%	711/38266	Avg=585.71/763
Outgroups							

Platynereis dumerilii	Annelida	Polychaeta	SRR1742987	92.03%	C:90.4(S:22.7)%	1228/132693	Avg=714.96/1065
Habronattus sansoni	Arthropoda	Arachnida	SRR6381055	67.79%	C:11.0(S:10.3)%	543/3673	Avg=450.28/497
Strigamia maritima	Arthropoda	Chilopoda	SRR7280107	97.65%	C:98.4(S:16.1)%	2081/33600	Avg=1375.65/2768
Aquatic lateralis	Arthropoda	Insecta	DRR119267	97.59%	C:98.7(S:62.2)%	2758/18156	Avg=1406.41/2646
Drosophila melanogaster	Arthropoda	Insecta	SRR5003953	97.95%	C:95.8(S:75.2)%	2228/5269	Avg=1233.62/2093
Heliconius erato	Arthropoda	Insecta	SRR2076809	96.48%	C:96.5(S:48.7)%	2422/13580	Avg=1034/1965
Tribolium castaneum	Arthropoda	Insecta	SRR4446751	95.76%	C:97.2(S:47.4)%	1637/2643	Avg=1511.41/2604
Carcinus maenas	Arthropoda	Malacostraca	SRR1586326	88.66%	C:73.5(S:53.6)%	1241/46247	Avg=684.6/1024
Macrobrachium koombooloomba	Arthropoda	Malacostraca	SRR7402067	96.30%	C:84.9(S:63.6)%	423/4078	Avg=806.62/1709
Oratosquilla oratoria	Arthropoda	Malacostraca	SRR5804714	93.33%	C:96.2(S:47.9)%	2013/54863	Avg=876.28/1722
Actinia tenebrosa	Cnidaria	Anthozoa	SRR2437124	95.80%	C:79.7(S:60.3)%	1429/34393	Avg=747.88/1123
Caenorhabditis elegans	Nematoda	Chromadorea	SRR8717248	73.86%	C:22.8(S:22.1)%	831/4595	Avg=511.42/645
Phoronis australis	Phoronida	Not Assigned	SRR5811956	94.61%	C:96.6(S:59.1)%	2420/83019	Avg=1029/.79/2267
Hypsibius dujardini	Tardigrada	Eutardigrada	SRR1739983	95.69%	C:87.9(S:20.1)%	2249/13327	Avg=1130.7/2104