



Theses and Dissertations

2005-06-23

Crustacean phylogenetic systematics and opsin evolution

Megan L. Porter

Brigham Young University - Provo

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>

 Part of the [Microbiology Commons](#)

BYU ScholarsArchive Citation

Porter, Megan L., "Crustacean phylogenetic systematics and opsin evolution" (2005). *Theses and Dissertations*. 557.

<https://scholarsarchive.byu.edu/etd/557>

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

CRUSTACEAN PHYLOGENETIC SYSTEMATICS
AND OPSIN EVOLUTION

by

Megan L. Porter

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Microbiology and Molecular Biology

Brigham Young University

August 2005

Copyright © 2005 Megan L. Porter

All Rights Reserved

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a dissertation submitted by

Megan L. Porter

This dissertation has been read by each member of the following graduate committee and
by majority vote has been found to be satisfactory.

Date

Keith A. Crandall, Chair

Date

David A. McClellan

Date

Dennis K. Shiozawa

Date

Jack W. Sites

Date

Michael F. Whiting

Date

Barry M. Willardson

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the dissertation of Megan L. Porter in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Keith A. Crandall
Chair, Graduate Committee

Accepted for the Department

Alan R. Harker
Department Chair

Accepted for the College

R. Kent Crookston
Dean, College of Biology and Agriculture

ABSTRACT

CRUSTACEANS AND OPSIN EVOLUTION

Megan L. Porter

Department of Microbiology and Molecular Biology

Doctor of Philosophy

Composed of a chromophore bound to an integral membrane protein (opsin), visual pigments are phenotypically characterized by the wavelength of maximal absorption (λ_{\max}). The underlying molecular mechanism controlling λ_{\max} is the interaction between the opsin amino acid sequence and the chromophore. While a plethora of studies have looked at structure/function relationships in vertebrate opsins, fewer studies have investigated similar issues in invertebrates. Furthermore, those few studies undertaken in invertebrate systems suggest different mechanisms of spectral tuning and photoactivation compared to vertebrate systems. This dissertation research is focused on expanding our knowledge of opsin evolution in invertebrate systems, particularly from non-insect taxa.

First, issues related to opsin evolution and the maintenance of supposedly ‘non-functional’ genes were explored in a review of regressive and reverse evolution.

Second, in order to place studies of crustacean opsin evolution in context, phylogenetic studies of two crustacean groups (Mysidae and Decapoda) were completed. Studies of Mysidae utilized 16S mtDNA, and 18S and 28S rDNA to reconstruct phylogenetic relationships and assess newly developed Bayesian methods of assessing pattern heterogeneity. Using this suite of genetic markers, there are incongruencies between current taxonomy and inferred phylogenetic relationships. Studies of Decapoda assessed phylogenetic relationships and estimated divergence times using 16S mtDNA, H3 nDNA, and 18S and 28S rDNA sequence data in conjunction with a set of eight fossil calibrations. Reconstructed phylogenies show support for two well supported nodes corresponding to the Pleocyemata and the informal ‘Reptantia’ and place the emergence of the Decapod lineage in the early Devonian (407 MYA).

Finally, opsin sequences and spectral sensitivity data from species within the Mysidae and Decapoda were combined with previously characterized invertebrate sequences to investigate opsin evolution. Standard d_N/d_S methods did not detect any evidence of selection. Methods investigating selection on amino acid properties, however, identified four properties (coil tendencies, compressibility, power to be at the middle of the alpha helix, and refractive index) to be under positive destabilizing selection. These properties occurred mostly at sites in transmembrane helices and included residues previously identified to affect spectral tuning as well as identifying novel sites.

ACKNOWLEDGEMENTS

I owe many people more thanks than I can express on these pages; to those who are too numerous to name who have given me support, encouragement, friendship, camaraderie, lively debate, and help with field work, analyses, or urgently needed specimens – thank you. Without you, the long days of my graduate career would not have been so productive, meaningful, and enjoyable. I am eternally indebted to my research advisor, Keith, for giving me the independence to find my own way (including making my own mistakes), but always being nearby to help me when I stumbled. Thank you for always believing in me. My appreciation goes to my graduate committee for their sage counsel and unyielding encouragement. I am also grateful for the guidance and support of Dr. Tom Cronin, who took me under his wing and gave an unknown graduate student his time and friendship.

Many thanks to those organizations who saw fit to fund parts of this research, including the National Science Foundation, the Cave Conservancy Foundation, the Cave Research Foundation, the American Association of University Women, the Brigham Young University Doctoral Fellowship, and the Department of Integrative Biology, the Department of Microbiology & Molecular Biology, and the College of Biology and Agriculture, Brigham Young University.

To my parents, I send all of my love and admiration for always believing in my quest to be an eternal student. Also many thanks are due to Dr. Horton H. Hobbs III, for sharing with me his love for science, setting me on this course of biological adventure, and for continuing to be my mentor and friend. Finally, special thanks are extended to Marcos Pérez-Losada and Katharina Dittmar, both of whom have traveled the world with me during my pursuit of this degree. If not for their devoted friendship and support through all of the best (and worst) times, I would not have survived to finish this research. You both have my unending adoration.

To study history one must know in advance that one is attempting something fundamentally impossible, yet necessary and highly important. To study history means submitting to chaos and nevertheless retaining faith in order and meaning.

It is a very serious task, young man, and possibly a tragic one.

- Father Jacobus (from Hesse's Magister Ludi)

TABLE OF CONTENTS

Abstract	v
Acknowledgements	vii
Table of Contents	ix
List of Tables	x
List of Figures	xi
Introduction	1
Research Outline	16
Chapter 2. Lost along the way: The significance of evolution in reverse	36
Is evolution reversible?	38
The continuum of reversibility	40
How does evolution back up?	44
The future of reversibility	46
Appendix 1. Glossary	48
Appendix 2. The debate over evolution in reverse	49
Appendix 3. Experimental systems in reverse evolution	51
Appendix 4. Regressive evolution in <i>Astyanax mexicanus</i>	52
Appendix 5. Studies of genomics in reverse	54
Chapter 3. Phylogenetic relationships within the Mysidae	59
Methods	63
Results	69
Discussion	74
Conclusions	79
Chapter 4. Decapod phylogeny and divergence times	94
Methods	98
Results	109
Discussion	113
Conclusions	118
Chapter 5. The Molecular Evolution of Invertebrate Opsins	129
Methods	133
Results	142
Discussion	145
Conclusions	161
References	163

LIST OF TABLES

Table 1-1. List of crustacean photoreceptor spectral sensitivities.....	18
Table 3-1. Mysidae taxonomy and GenBank accession numbers.....	81
Table 3-2. Parameter estimates for models of nucleotide substitution.....	83
Table 3-3. Likelihood topology tests.....	84
Table 3-4. Model comparisons using Bayes Factors.....	85
Table 4-1. Decapoda taxonomy and GenBank accession numbers.....	119
Table 4-2. Decapod fossil calibrations ages and taxonomy	121
Table 4-3. Likelihood and Bayesian topology tests	122
Table 4-4. Comparison of divergence times for major decapod lineages	123
Table 5-1. Opsin taxonomy, GenBank accession numbers, and λ_{\max}	153

LIST OF FIGURES

Figure 1-1.	3-D image of bovine rhodopsin	31
Figure 1-2A.	Invertebrate opsin phylogeny.....	33
Figure 1-2B.	Vertebrate opsin phylogeny	34
Figure 2-1.	Examples of convergence due to subterranean habitats	56
Figure 2-2.	Eye development in surface vs. cave <i>Astyanax mexicanus</i>	57
Figure 2-3.	Evolutionary history of aphids and their endosymbiont bacteria	58
Figure 3-1.	Previous hypothesis of Mysidae phylogenetic relationships	86
Figure 3-2.	Consensus tree of model-based phylogenies	87
Figure 3-3.	Consensus tree of eight most parsimonious trees	89
Figure 3-4.	Bayesian analyses of different numbers of rate matrices.....	91
Figure 3-5.	Rate parameter convergence graphs for BMCMC-bp analyses.....	93
Figure 4-1.	Previous hypotheses of decapod relationships.....	124
Figure 4-2.	Decapod divergence time chronogram	125
Figure 4-3.	Divergence time estimates using different sets of calibrations.....	127
Figure 5-1.	2 dimensional diagram of the opsin protein domains	155
Figure 5-2.	MSP absorbance curves for estimated mysid spectral sensitivity	156
Figure 5-3	ML amino acid phylogeny of invertebrate opsins	157
Figure 5-4.	Invertebrate opsin TreeSAAP sliding window analysis	158
Figure 5-5	3D opsin structure illustrating amino acids under selection	159

CHAPTER 1.

INTRODUCTION

Visual pigment research has long been of interest to a number of biological disciplines, including sensory ecologists, visual physiologists, biochemists, and molecular evolutionists. Composed of a chromophore bound to an integral membrane protein (opsin), visual pigments are phenotypically characterized by the wavelength of maximal absorption (λ_{\max}). Although there are a number of morphological and physiological methods of controlling the λ_{\max} spectral sensitivity, the underlying molecular mechanism is the interaction between the particular amino acid sequence of the opsin protein and the type of chromophore. Because both the amino acid sequence and the phenotype (λ_{\max}) can be measured directly, visual pigments have become model systems for studying genotype / phenotype interactions.

While a plethora of studies have looked at structure / function relationships in vertebrate opsin from an evolutionary perspective, fewer studies have investigated similar issues in invertebrates. Outside of insects, the only invertebrate taxa in which opsin sequences have been explicitly investigated are horseshoe crabs (Smith et al. 1993), molluscs (Brown and Brown 1958; Hall et al. 1991; Hubbard and St. George 1958; Morris et al. 1993), and crustaceans (Crandall and Cronin 1997; Crandall and Hillis 1997; Oakley and Huber 2004; Sakamoto et al. 1996). Furthermore, those few studies undertaken in invertebrate systems suggest different mechanisms of spectral tuning compared to the more broadly studied vertebrate system. This dissertation research is focused on expanding our knowledge of opsin evolution in invertebrate systems, particularly from non-insect taxa. The Crustacea were chosen as a focal group because of

the considerable amount of research previously devoted to their visual systems (Cronin 2005). The significant amount of crustacean physiological data available provide a solid background for understanding opsin evolution, but the lack of molecular studies (i.e. opsin sequence data) has made it difficult to assess the importance of molecular mechanisms relative to other types of visual system spectral control (i.e. eye structure, optical filters, screening pigments).

This study is aimed at collecting the phylogenetic systematic, physiologic, and molecular data necessary to understand visual pigment evolution in invertebrates, focusing on Crustaceans. In this introduction, I will review visual pigment structure, function (i.e. spectral tuning), evolutionary history, and finally what is known of crustacean visual pigment physiology and genetic variation. By reviewing these topics, I will illustrate the knowledge gap of crustacean visual systems at the molecular level and highlight the particular questions being addressed in this dissertation.

What is a visual pigment and how are they ‘tuned’?

A visual pigment is composed of two components, a chromophore (a form of vitamin A) covalently bound to an opsin protein (Figure 1-1). Although, there are many ways to modify the spectral sensitivity of a photoreceptor, including photoreceptor structure (e.g. tiered retinas in stomatopods, Marshall and Land 1993), utilizing different chromophores (Bowmaker 1995; Kito et al. 1986; Matsui et al. 1988; Suzuki et al. 1984), changes in opsin gene expression in response to the changing visual demands of different life stages (Beaudet and Hawryshyn 1999; Bowmaker 1995; McFarland and Loew 1994; Shand 1993; Shand 1994), differential expression of a common complement of opsin genes

among closely related species (Carleton and Kocher 2001), ocular filters (Cheroske et al. 2003; Cronin et al. 2001; Douglas and Marshall 1999), and screening pigments, all of these modifications build on the initial spectral sensitivity of the visual pigment as determined by the interaction between the photosensitive chromophore and the amino acid sequence of the opsin protein. Although several different chromophores have been documented from visual pigments, including 3-dehydroretinal (A2) (Jokela-Määttä et al. 2005; Kito et al. 1986; Suzuki et al. 1984), 3-hydroxyretinal (A3) (Seki and Vogt 1998), and 4-hydroxyretinal (Matsui et al. 1988), by far the most common chromophore is retinal (A1). Since most visual pigments utilize the same chromophore, the underlying variation in sensitivity is largely determined by the specific amino acid sequence of the opsin protein.

Opsin is a member of the G protein-coupled receptor (GPCR) superfamily, which consists of integral membrane proteins all with a similar structural motif that respond to environmental signals (ligands) and initiate signal transduction pathways (Filipek et al. 2003b). GPCRs most likely had an early evolutionary origin, as evidenced by their presence in the genomes of bacteria, yeast, plants, nematodes, arthropods, and chordates (Kroeze et al. 2003). These vital receptors are estimated to comprise significant portions of metazoan genomes (~1% in *Drosophila melanogaster*, ~ 5% in *Caenorhabditis elegans*, and ~2-6% in humans) (Bargmann 1998; Broek 2001; Fredriksson et al. 2003; Mirzadegan et al. 2003; Vassilatis et al. 2003). Their importance as receptors is stressed by the fact that anywhere from 1/3 (Robas et al. 2003) to 1/2 (Flower 1999) of currently marketed drugs target GPCRs. Opsins are the only group of GPCRs known to contain a bound ligand (the chromophore) and the only signal pathway initiated by a photon.

Rhodopsin is perhaps the most widely studied GPCR. Additionally, as the only crystallized GPCR (Palczewski et al. 2000), bovine rhodopsin has become the template for homology modeling of GPCRs in general (Filipek et al. 2003a; Filipek et al. 2003b; Oliveira et al. 2004). With a long history of serving as a model system for understanding structure and function issues of the GPCR superfamily, the affect of a large number of vertebrate amino acid residues have been investigated with respect to spectral tuning using site directed mutagenesis studies (Asenjo et al. 1994; Chan et al. 1992; Cowing et al. 2002a; Cowing et al. 2002b; Fasick et al. 2002; Fasick and Robinson 1998; Nathans 1990a; Neitz et al. 1991; Yokoyama 2000; Yokoyama 2002; Yokoyama and Radlwimmer 1999; Yokoyama and Radlwimmer 2001; Yokoyama et al. 2000; Yokoyama and Tada 2003).

‘Spectral tuning’ refers to the spectral adaptation of the λ_{max} of a visual pigment to a particular environment or visual task. Some of the first studies to investigate visual pigment spectral tuning focused on changes of λ_{max} in relation to spectral distribution of environmental light by demonstrating blue-shifted rhodopsin pigments from fish species of increasing depths (Crescitelli et al. 1985; Douglas et al. 1998; Lythgoe 1972; Lythgoe 1980; Partridge 1989; Partridge et al. 1988; Partridge et al. 1989). However, not until the isolation of the first complete opsin sequence from bovine rod opsin in 1983 (Nathans and Hogness 1983) were investigations of the underlying genetic mechanisms possible. Studies began to focus on exactly how specific amino acid replacements affect the spectral sensitivity of a visual pigment. Since then, a great deal of research has been devoted to identifying the members of the opsin gene family found in vertebrates (Bowmaker 1998; Bowmaker et al. 1991; Bowmaker et al. 1994; Yokoyama 2000) and

invertebrates (Briscoe and Chitkka 2001; Carulli et al. 1994; Carulli and Hartl 1992; Chang et al. 1996; Crandall and Cronin 1997; Popp et al. 1996; Sakamoto et al. 1996; Smith et al. 1997; Smith et al. 1993; Taylor et al. 2005; Towner and Gartner 1994; Towner et al. 1997) and investigating sites important to spectral tuning using comparative evolutionary, site directed mutagenesis, and homology modeling studies (Bellingham et al. 1998; Briscoe 2002; Britt et al. 1993; Carleton and Kocher 2001; Chang et al. 1995; Cowing et al. 2002a; Cowing et al. 2002b; Fasick et al. 2002; Feiler et al. 1992; Hunt et al. 2004; Neitz et al. 1991; Salcedo et al. 1999; Salcedo et al. 2003; Shi et al. 2001; Shi and Yokoyama 2003; Wilkie et al. 2000; Williams et al. 1992; Yokoyama and Radlwimmer 1999; Yokoyama et al. 2000; Yokoyama and Tada 2003). Spectral tuning has been more intensively researched in vertebrates, where site directed mutagenesis studies and in vitro expression systems allow researchers to measure the effects of single amino acid changes on λ_{\max} . In conjunction with comparative methods to identify potential sites of importance for further investigation, site directed mutagenesis has identified more than 18 amino acid residues important in the spectral tuning of the observed vertebrate spectral variants (Yokoyama 2002). In invertebrates, heterologous expression systems using *Drosophila* have been used to investigate spectral tuning via construction of chimeric rhodopsins molecules and site-directed mutagenesis (Britt et al. 1993; Salcedo et al. 2003); however, this methodology is difficult and time-intensive, and not many invertebrate opsin spectral variants have been investigated using this technique. Most commonly, studies in invertebrates have focused on using comparative methods and homology modeling to predict sites important to spectral tuning (Briscoe 2002; Chang et al. 1995; Morris et al. 1993). The combination of these methods indicate that while some

of the sites identified in vertebrate opsins are also important to invertebrate tuning, there are also potentially different sites responsible for the spectral diversity observed in invertebrates (Briscoe 2002).

Invertebrate vs. Vertebrate Opsin Evolution

In vertebrates, there are five major families of visual pigments (rhodopsin 1: Rh1, rhodopsin-like: Rh2, middle/long-wavelength sensitive: MWS/LWS, short-wavelength sensitive 1: SWS1, and short-wavelength sensitive 2: SWS2), which diversified before the separation of the major vertebrate classes (Bowmaker 1998). More recently, a number of additional opsin classes, both ocular and extraocular, have been identified that are not involved in image perception, including pinopsin expressed in the pineal of lamprey, teleosts, squamates, and birds (Forsell et al. 2001; Kawamura and Yokoyama 1996; Kawamura and Yokoyama 1997; Max et al. 1995; Okano et al. 1994; Yokoyama and Zhang 1997), vertebrate ancient opsin (VA) isolated from the teleost inner retina and pineal/deep brain structures (Kojima et al. 2000; Moutsaki et al. 2000; Philp et al. 2000; Soni and Foster 1997; Soni et al. 1998), retinal G protein-coupled receptors (RGR) expressed in Müller cells and retinal pigment epithelium (Hao and Fong 1996; Pandey et al. 1994; Shen et al. 1994), or have an unknown function such as peropsin (RRH, Sun et al. 1997), parapinopsin (Blackshaw and Snyder 1997), encephalopsin (Blackshaw and Snyder 1999), and melanopsin (Hattar et al. 2002). Phylogenetic analyses and intron position homology of this full complement of visual and non-visual vertebrate opsin genes indicate that vertebrate opsins arose from three ancestral lineages: a ‘classical opsin’ group, composed of visual (rod and cone opsins) and brain (pinopsin, VA-opsin,

parapinopsin, encephalopsin) variants, a clade of the RRH and RGR variants, and a melanopsin lineage (Bellingham et al. 2003). Within the ‘classical opsin’ group, pinopsins diverged from the visual pigment clade just prior to or just after the gene duplication separating LWS from SWS opsins, while VA opsins probably diverged early in the evolution of the vertebrate photopigments, preceding the initial visual pigment gene duplication (Bowmaker 1998).

Much less is known about the mode and tempo of invertebrate opsin evolution. Most of the evolutionary work related to visual pigments in invertebrates has focused on insect systems (Briscoe 2001; Briscoe 2002; Carulli et al. 1994; Carulli and Hartl 1992; Feiler et al. 1992; Feiler et al. 1988; Montell et al. 1987; O'Tousa et al. 1985; Salcedo et al. 1999; Smith et al. 1997; Spaethe and Briscoe 2004; Taylor et al. 2005; Towner and Gartner 1994; Zuker et al. 1987), and to a lesser extent on cephalopods (Brown and Brown 1958; Hall et al. 1991; Hubbard and St. George 1958; Morris et al. 1993; Suzuki et al. 1976). Outside of insects, much of the opsin variation in the invertebrate world remains uncharacterized and how the spectral variants from other major arthropod groups relate to these clades is unknown. Although there is similarity between vertebrate and invertebrates in phototransduction strategies and potentially in spectral tuning mechanisms, there are also differences in the underlying molecular machinery (Nakagawa et al. 1999; Zuker 1996), reflecting the hypothesis that the duplication of opsin genes occurred independently in the lines of descent leading to mammals and insects (Fryxell and Meyerowitz 1991; Pichaud et al. 1999; Zuker et al. 1987). Research of insect opsins have delineated at least four main spectral classes (Carulli et al. 1994; Feiler et al. 1992; Feiler et al. 1988; Montell et al. 1987; O'Tousa et al. 1985; Salcedo et

al. 1999; Zuker et al. 1987), with multiple gene duplications occurring within spectral classes for particular groups (Briscoe 1998; Briscoe 2000; Hill et al. 2002; Spaethe and Briscoe 2004). Furthermore, similar to vertebrates, a number of invertebrate extraocular photoreceptors have been documented, including stemmata, eyelet cells, nauplius eyes, frontal organs, intracerebral ocelli, and caudal photoreceptors (Andresen and Brown 1979; Brown and Brown 1973; Gotow 1989; Meyer-Rochow 2001; Sandeman et al. 1990; Wilkens and Larimer 1976). Recent research has indicated that some of the first extraocular opsin sequences isolated from insects are derived from visual opsin variants (Briscoe and White 2005; Shimizu et al. 2001). Future studies will be necessary to determine if any of this diversity of receptors is manifested at the genetic level by additional classes of opsins as in vertebrates.

Studies of the *Anopheles gambiae* (mosquito) genome suggest that there are a number of opsin genes present that do not fall into the classical insect spectral clades and may be non-visual receptors. Hill et al. (2002) found a total of twelve opsin genes in *A. gambiae*, two of which (GPRop11 and GPRop12) clustered with vertebrate non-visual opsins in phylogenetic analyses. Partial sequence homology with bee brain ESTs implies these genes may have non-visual functions similar to their vertebrate homologs. Furthermore, in their analyses, vertebrate melanopsin clustered with invertebrate visual pigments. These findings led Hill et al. (2002) to propose that the common ancestor of protostomes and deuterostomes had at least two non-visual opsin genes, one of which diversified producing invertebrate visual pigments while vertebrate pigments evolved from the other. Unfortunately, no attempts have been made to investigate this hypothesis further. Preliminary phylogenetic analyses of a taxonomic selection of all available visual, non-

visual, and uncharacterized opsin sequences from both vertebrates and invertebrates support this hypothesis by placing vertebrate melanopsin sequences basal to the main invertebrate visual pigment clade sister to platyhelminthes and scallop opsins (Figure 1-2A). In contrast to the findings of Hill et al. (2002), however, the *A. gambiae* genes GPRop11 and GPRop12 cluster as basal invertebrate sequences in this phylogeny, implying that early opsin evolution and diversification is still murky.

The vertebrate visual pigment clade (Figure 1-2B) is resolved similarly to other studies (Bowmaker 1998; Yokoyama 2000), with the RH1 and RH2 clades arising from the SWS clades, and the MWS/LWS and pineal opsin clades as the most basal lineages. Sister to the visual pigment+pineal clade are the parapinopsin and vertebrate ancient (VA) opsins. The most basal lineage in this clade is an opsin gene from the ascidian *Ciona intestinalis*, implying that vertebrate visual pigments arose from genes present in basal chordates. Falling in between the Chordate visual pigment clade and the rest of the characterized vertebrate extraocular opsins is an opsin from the *A. gambiae* genome, GPRop10. The placement of vertebrate melanopsins with the invertebrates and of GPRop10 with vertebrates fits with Hill et al.'s (2002) hypothesis that the vertebrate and invertebrate visual pigments arose from two different non-visual opsin genes, and show that the genomes of both groups still contain traces of these ancestral genes. Comparing the function and expression of these 'misplaced' basal groups will be an interesting avenue of future research, and may help clarify mechanistic differences between invertebrate and vertebrate photoactivation. In particular, as expression systems for invertebrates are still limited, investigations of melanopsin protein function relative to vertebrate visual pigments should be particularly productive.

The basal most lineages in the opsin phylogeny are mostly vertebrate extraocular opsins and sequences isolated from amphioxus, demonstrating that a number of opsin gene copies are present in chordate genomes that predate the invertebrate/vertebrate visual pigment diversifications. The only invertebrate representative in the basal group (*Platynereis dumerilii*) is an opsin from a ciliary type photoreceptor cell (c-opsin). As most invertebrate opsins are associated with rhabdomeric photoreceptor cells (r-opsin) and vertebrate visual pigments with ciliary types, the placement of this planarian c-opsin implies that the ancestral photoreceptor cell is of the c-type. Furthermore, the *A. gambiae* genes GPRop11 and GPRop12 have been hypothesized to belong to c-opsin type (Ardendt et al. 2004). The placement of these mosquito opsins basal to the rest of the invertebrate visual pigments implies that the r-opsins arose from the c-opsin lineage. Further research in basal animal lineage opsins will be of interest to investigating this hypothesis.

Finally, this phylogeny illustrates the disparity between the known gene family diversity and taxonomic representation of vertebrate versus invertebrate opsins, particularly the lack of sequence data from non-insect arthropods. Using our knowledge of insect visual pigment physiology and evolution as a foundation, the next stage in studies of invertebrate opsin evolution is to investigate opsins from other major groups within the Arthropoda.

Arthropoda Opin Evolution

The Arthropoda consist of four main taxonomic groups (Chelicerata, Crustacea, Myriapoda, Hexapoda). Of these four, there have been no opsin sequences isolated from

within the Myriapoda and opsin sequences from the Chelicerata are only known from studies of *Limulus polyphemus* (Smith et al. 1993). With slightly more representation, the available crustacean opsin sequences are from one crab (Sakamoto et al. 1996), one stomatopod (Brown 1996), ten crayfish (Crandall and Cronin 1997; Crandall and Hillis 1997), and two ostracod species (Oakley and Huber 2004). Based on the most recent arthropod phylogenetic studies, Chelicerata, Myriapoda, and Crustacea+Hexapoda are well-supported clades, although there is dispute about whether the Myriapoda are sister to the chelicerates or the crustacean/hexapod clade (Giribet et al. 2004; Wheeler et al. 2004). Investigations of the spectral sensitivities of photoreceptor classes across this phylogeny have led to the hypothesis that the ancestral Chelicerata possessed dichromatic vision, based on one UV and one green receptor, presumably to discriminate between ‘open space’ (i.e. high UV content) and dense habitat (i.e. reflected light lacking UV) (Pichaud et al. 1999). In comparison, it is hypothesized that ancestral Mandibulata (Crustacea, Hexapoda, and Myriapoda) gained an additional photoreceptor class to become trichromatic, possessing UV, blue, and green sensitive visual pigments (Chitkka 1996; Chitkka 1997). This implies that in both crustacean and hexapod species, missing photoreceptors represent secondary losses while more complex visual pigment systems are the result of independent gains. However, this hypothesis is based solely on the distribution of visual pigment spectral sensitivities and has never been critically tested using opsin molecular data. Furthermore, this hypothesis is based on the receptor spectral distribution of only two crustacean representatives, *Ligia exotica* (Isopoda) and *Daphnia magna* (Cladocera), which possess three or more photoreceptors, and no information is available for myriapods or basal hexapods. Since the phylogenetic

relationships of many of these groups are still contentious (e.g. arthropod and crustacean relationships, crustacean monophyly), deciphering which groups are ancestral versus derived is difficult. Given that most crustaceans have only two photoreceptor classes (Cronin 2005) compared with the ancestral UV, blue, and green receptor complement of pterygote insects (Briscoe and Chitkka 2001), it seems that the most crucial taxonomic groups for rigorously testing this hypothesis have yet to be investigated from either a physiological or a genetic perspective.

Evolution of visual pigments in crustaceans

The structure and design of arthropod compound eyes reflect on their function, and are influenced by the behavior, ecology, and evolutionary history of the species (Meyer-Rochow 2001; Schiff and Hendrickx 1997). The greatest diversity, by far, of eye designs and adaptations is found in aquatic invertebrates (Cronin 2005). In particular, the diversity of optical mechanisms in adult crustaceans involves more functional designs than are found in all other animals combined (Cronin 1986; Land 1981; Nilsson 1989). This diversity, in part, explains the long history of studies of visual systems, eye morphologies, and visual physiology in crustaceans. Yet very few molecular studies have been attempted to assess the importance of molecular mechanisms relative to other types of visual system spectral control (eye structure, optical filters, pigments, etc.), and most studies of crustacean visual systems have focused on characteristics of the compound eyes; very little is known about the other types of photoreceptors (Meyer-Rochow 2001).

With respect to visual pigments, most studies in crustaceans have involved measuring the λ_{\max} of species from various habitats or taxonomic groups (Table 1).

These investigations have shown that crustaceans are highly conservative in their spectral sampling capabilities, possessing only 1 or 2 photoreceptor types irrespective of habitat. This pattern suggests that phylogenetic constraints play a large role in the observed crustacean λ_{\max} diversity (Frank and Widder 1999), similar to observations for some deep-sea fish (Douglas and Partridge 1997; Partridge 1989; Partridge et al. 1992). The main photoreceptor used by crustaceans is sensitive to blue/green light, ranging in peak absorption from 480-540 nm (Crandall and Cronin 1997; Marshall et al. 2003; Marshall et al. 1999). If a second class of photoreceptor is present, it is invariably sensitive to UV/blue light (Johnson et al. 2002; Marshall et al. 1999). Exceptions to this pattern have been found in *Daphnia magna* where four photoreceptor classes have been documented (Smith and Macagno 1990), and in the isopod *Ligia exotica* where three spectral classes have been measured (Hariyama and Tsukahara 1993). The most notable exception to this pattern, however, is the Stomatopoda, which contain an unparalleled complexity in visual pigments, with some species capable of sampling up to 16 different spectra that span the ultraviolet and visible spectrum of light (~320-700 nm) in a single retina (Cronin and Marshall 1989; Cronin and Marshall 2004; Cronin et al. 2000; Cronin et al. 1994b; Cronin et al. 1994c).

The typical spectral classes of crustacean visual pigments, however, fit with the current hypothesis that the plesiomorphic condition for Mandibulata is trichromacy based on UV, blue, and blue/green classes of visual pigments, if the typical crustacean represents a derived and reduced case (Briscoe and Chitkka 2001; Chitkka 1996; Chitkka 1997). However, despite abundant physiological data, crustacean opsin sequence data with which to test this hypothesis are conspicuously lacking. Currently, the only

sequence data available with concordant physiological data for crustacean opsins are from one crab (Sakamoto et al. 1996) and several crayfish species (Crandall and Cronin 1997; Crandall and Hillis 1997). Moreover, the only probable short-wavelength crustacean opsins that have been molecularly characterized have been short fragments (~400bp) from two ostracod species lacking the corresponding physiological data (Oakley and Huber 2004).

In crustaceans, the closest to a model system for visual pigment studies is crayfish. Much of what is known about crustacean visual systems can be illustrated by a review of this body of research. Crayfish were the first crustaceans from which an opsin sequence was isolated (Hariyama et al. 1993). This first step led to studies investigating the spectral tuning and functional evolution of the main visual pigment opsin in crayfish (Crandall and Cronin 1997; Crandall and Hillis 1997). Although most studies have focused on the main retinular photoreceptors, crayfish represent the typical crustacean dichromatic state, where the eighth retinular cell has been identified as a violet receptor with a λ_{max} of 440nm, although the opsin for this visual pigment has yet to be genetically characterized (Cummins and Goldsmith 1981).

Crayfish photoreceptors possess two kinds of chromophores (A1 and A2), and there is seasonal variation of the A2 content, causing changes in the spectral sensitivity of the photoreceptors cells. Use of porphyropsin, the visual pigment based on the A2 chromophore, in an invertebrate was first characterized in *Procambarus clarkii* (Suzuki et al. 1984; Zeiger and Goldsmith 1989). Porphyropsin in crayfish is synthesized in response to lower temperatures, is broken down at higher temperatures in the presence of light, and the proportion varies with season, reaching its peak during the colder winter

months (Suzuki et al. 1985; Suzuki et al. 1984). With the exception of the squid *Watsenia scintillans* (Kito et al. 1986), rhodopsin (A1)/porphyropsin (A2) systems are only described from freshwater crustaceans, and are only well-studied in crayfish where it has been documented in *Procambarus clarkii*, *Cherax destructor*, and *Euastacus armatus*, but not in other crayfish or decapod species (Suzuki et al. 1985; Suzuki and Eguchi 1987; Suzuki et al. 1984; Zeiger and Goldsmith 1993). More recently, a rhodopsin/porphyropsin system has also been found in the freshwater mysid species, *Mysis relicta* (Jokela-Määttä et al. 2005).

Crayfish have also been scrutinized in terms of studies of extraocular photoreceptors, possessing both primitive photoreceptors in their caudal ganglion (Larimer 1966; Wilkens and Larimer 1976) and intracerebral ocelli (Bobkova et al. 2003). Caudal photoreceptors (CPRs) consist of a pair of light-sensitive neurons in the sixth abdominal ganglion and have maximal sensitivity of ~500 nm (Bruno and Kennedy 1962). Functional CPRs have been described from at least nine crayfish, including blind cave-dwelling species (Larimer 1966; Wilkens and Larimer 1976), and in at least ten additional decapod species (Wilkens and Larimer 1976). Among malacostracan crustaceans, intracerebral ocelli were first described in Isopoda (Martin 1971; Martin 1976), but more recently have been described from both amphipods (Frélon-Raimond et al. 2002) and crayfish (Hafner et al. 2003; Sandeman et al. 1990). Probes based on crayfish MWS/LWS pigments described by Crandall and Cronin (1997) have been localized in *Procambarus clarkii* using *in situ* hybridization to the retinular cells and potentially to these extraretinal intracerebral ocelli, suggesting that the same or a similar opsin is present in both photoreceptors, similar to the recently documented opsin

expression in lepidopteran adult stemmata (Briscoe and White 2005). However, these probes did not stain the caudal photoreceptor, indicating a delayed developmental expression, unsuitable methodologies, or the expression of a different photopigment.

RESEARCH OUTLINE

Because of a long history of research on visual system evolution and physiology, my dissertation research focuses on expanding the knowledge of opsin genetic variation within the Crustacea. However, even with the accumulated information on crustacean visual physiology in terms of λ_{\max} (Table 1), there are a number of issues remaining before opsin evolution in crustaceans (and invertebrates), can be properly investigated. One of these issues is understanding the development of the visual system. In particular, given that functional opsin genes have been isolated from cave-adapted crayfish species which no longer have functional eyes, questions exist related to how a structure (the eye) can be lost, but the integral protein (opsin) maintained (Crandall and Hillis 1997). In an attempt to examine these issues from the unique perspective of structural loss, chapter two examines what is known of the genetic and developmental control of eye degeneration in cave-adapted species in a general review of regressive and reverse evolution. These concepts are related to recent opsin research isolating non-visual photopigments from both vertebrates and invertebrates, and future research should be directed towards a better understanding of these extraocular receptors.

Another issue with understanding opsin evolution in crustaceans is the lack of consensus with regard to phylogenetic histories. Many crustacean lineages have not yet been investigated from a phylogenetic perspective and without hypotheses of

relationships, interpreting the ancestral versus derived states of physiological and gene evolution data are difficult. Therefore, chapters three and four are phylogenetic studies of two groups of crustaceans: the Mysida, which as an order has been virtually ignored from a phylogenetic perspective, and the economically important Decapoda, in which morphological phylogenetic hypotheses have been debated for centuries, but no attempts have ever been made at a comprehensive molecular phylogeny of the entire order.

Finally, while there are a large number of studies devoted to understanding visual pigment structure, function, and evolution in vertebrates (Asenjo et al. 1994; Bowmaker 1998; Fasick et al. 2002; Fasick and Robinson 1998; Merbs and Nathans 1992; Nathans 1990b; Nathans et al. 1989; Neitz et al. 1991; Shand 1993; Shand et al. 1988; Shi et al. 2001; Shi and Yokoyama 2003; Wilkie et al. 2000; Yokoyama and Yokoyama 1990; Yokoyama 2000; Yokoyama 2002; Yokoyama and Radlwimmer 1999; Yokoyama and Radlwimmer 2001; Yokoyama et al. 2000; Yokoyama and Shi 2000; Yokoyama and Tada 2003), relatively few studies have been devoted to invertebrates. Chapter five examines the selective influences in all available invertebrate opsins that have been both genetically and physiologically characterized. To add to the data available for crustaceans, opsin sequences were isolated and λ_{\max} was characterized from a select group of mysid and decapod species. This research will build a foundation for understanding opsin evolution across a larger diversity of invertebrates, for future studies comparing invertebrate opsin evolution and diversification with the more widely studied vertebrate opsin family, and for testing hypotheses of animal opsin evolution in general.

Table 1-1. List of crustacean photoreceptor λ_{max} from the literature and from this study. All measurements are from adults, unless the species name is followed by an N (nauplius), L (larvae), M (megalopa) or Z (zoea). More than one λ_{max} indicate species where multiple pigments have been characterized. For stomatopod visual pigments, λ_{max} values from different portions of the retina are listed in the following order: peripheral retina rhabdoms, midband row 1 distal rhabdom, midband row 1 proximal rhabdom, midband row 2 distal rhabdom, midband row 2 proximal rhabdom, midband row 3 distal rhabdom, midband row 3 proximal rhabdom, midband row 4 distal rhabdom, midband row 4 proximal rhabdom, midband rows 5 and 6 rhabdoms. Where depth ranges are not available, habitat abbreviations are: ST – subtidal; IT – intertidal; C – coastal; P – pelagic; MP – mesopelagic; BP – bathypelagic; DB – deep benthic; EST – estuarine; BW – brackish water; FW – freshwater; sTR – semi-terrestrial; TR - terrestrial. The abbreviations used for the method of measuring λ_{max} are as follows: BP – behavioral phototaxis; EON – extracellular/optic nerve; ERG – intercellular electrophysiology; EX – spectrophotometry of pigment extract; IC – intracellular electrophysiology; MSP – microspectrophotometry; VC – voltage clamp. Genbank accession numbers (<http://www.ncbi.nlm.nih.gov/>) are given for those species where visual pigments have also been characterized genetically. Sequences similar to opsin isolated from expressed sequence tag libraries are indicated by (EST) after the accession number.

	Habitat	method	λ_{max}	λ_{max} reference	Accession #
BRANCHIOPODA					
Anostraca					
Artemiidae					
<i>Artemia franciscana</i>					BQ605261 (EST)
<i>Artemia salina</i>		EON	410	(Hertel 1972)	
Diplostraca					
Daphniidae					
<i>Daphnia magna</i>		VC	348, 434, 525, 608	(Smith and Macagno 1990)	
MAXILLOPODA					
Copepoda					
Acartiidae					
<i>Acartia tonsa</i>		BP	450-520	(Stearns and Forward 1984)	
Cirripedia					
Balanidae					
<i>Balanus amphitrite</i>		ERG	532	(Hillman et al. 1973)	

<i>Balanus amphitrite</i>		MSP	532	(Minke and Kirschfield 1978)
<i>Balanus balanoides</i> (N)		BP	510-530	(Barnes and Klepal 1972)
<i>Balanus eburneus</i>		ERG	532	(Hillman et al. 1973)
<i>Balanus eburneus</i>		MSP	532	(Minke and Kirschfield 1978)
OSTRACODA				
Myodocopida				
Cypridinidae				
<i>Skogsbergia lernerii</i>				AF353374-
<i>Vargula hilgendorfi</i>				AF353339
				AF353338-
				AF353331
MALACOSTRACA				
Stomatopoda				
Squilloidea				
<i>Coronis scolopendra</i>	ST	MSP	494, 407, 436, 489, 518, 529, 533, 441, 468, 517	(Cronin et al. 1993)
<i>Lysiosquillia sulcata</i>	5-25m	MSP	499, 397, 434, 492, 516, 517, 538, 416, 461, 500	(Cronin et al. 1993)
<i>Pullosquilla litoralis</i>	sIT	MSP	509, 404, 425, 469, 509, 527, 540, 446, 455, 482	(Jutte et al. 1998b)
<i>Pullosquilla litoralis</i> (L)		MSP	446	(Jutte et al. 1998a)
<i>Pullosquilla thomassini</i>	1-37m	MSP	467, 405,	(Jutte et al. 1998b)

				445, 489, 509, ???, ???, 456, 452, 483	
<i>Pullosquilla thomassini</i> (L)		MSP	447	(Jutte et al. 1998a)	
Gonodactyloidea					
<i>Gonodactylellus affinis</i>	3-30m	MSP	500, 400, 424, 496, 521, 546, 541, 454, 474, 509	(Cronin et al. 2002)	
<i>Gonodactylaceus falcatus</i>		MSP	510, 400, 443, 513, 527, 532, 553, 443, 475, 518	(Cronin et al. 2000; Cronin et al. 1995)	
<i>Gonodactylaceus falcatus</i> (L)		MSP	499	(Cronin et al. 1995)	
<i>Gonodactylus smithii</i>	IT	MSP	517, 400, 440, 505, 528, 536, 552, 436, 469, 512	(Chiao et al. 2000)	
<i>Gonodactylopsis spongicola</i>	5-60m	MSP	506, 401, 444, 505, 525, 536, ???, 448, 474, 507	(Cronin et al. 2002)	
<i>Neogonodactylus curacaoensis</i>	2-20m	MSP	467, 400, 434, 494, 520, ???, ???, 435,	(Cronin et al. 1996)	

<i>Neogonodactylus oertstedii</i>	IT-3m	MSP	467, 511 528, 400, 430, 505, 525, 520, 551, 429, 460, 489	(Cronin and Marshall 1989)	(Brown 1996)
<i>Hemisquilla ensigera</i>	10-15m	MSP	501, 414, 451, 499, 510, 510, 535, 443, 473, 500	(Cronin et al. 1994a)	
<i>Haptosquilla trispinosa</i>	ST-25m	MSP	499, 400, 433, 508, 537, 539, 558, 422, 462, 510	(Cronin et al. 2002)	
<i>Pseudosquilla ciliata</i>	ST-25m	MSP	498, 400, 433, 498, 517, 535, 539, 425, 452, 510	(Cronin and Marshall 1989)	
<i>Odontodactylus brevirostris</i>	10-25m	MSP	490, 402, 457, 495, 524, 511, 535, 452, 460, 589	(Cronin et al. 1994a; Cronin et al. 1996)	
<i>Odontodactylus 'havanensis'</i>	20-35m	MSP	475, 407, 446, 485, 520, ???, ???, 428, 459, 501	(Cronin et al. 1996)	

<i>Odontodactylus scyllarus</i>	1-30m	MSP	503, 400, 430, 487, 509, 528, 546, 429, 451, 506	(Cronin et al. 1994a; Cronin et al. 1996)
Squilloidea				
<i>Squilla empusa</i>	1-150m	MSP	517	(Cronin 1985; Cronin et al. 1993)
<i>Squilla empusa</i>		MSP	507	(Cronin and Jinks 2001)
<i>Squilla empusa</i> (L)		MSP	509	(Cronin and Jinks 2001)
<i>Cloridopsis dubia</i>	IT	MSP	510	(Cronin et al. 1993)
Lophogastrida				
Lophogastridae				
<i>Gnathophausia ingens</i>	>400m	ERG	490, 520	(Frank and Case 1988b)
Mysida				
Mysidae				
<i>Archaeomysis grebnitzkii</i>		MSP	496	this study
<i>Bowmaniella</i> sp.		MSP	SW, 502	this study
<i>Hemimysis anomala</i>		ERG	393, 500	(Lindström 2000)
<i>Heteromysis formosa</i>		MSP	499	this study
<i>Holmesimysis costata</i>		MSP	512	this study
<i>Neomysis americana</i>		MSP	520	this study
<i>Neomysis integer</i>		ERG	525-535	(Lindström 2000)
<i>Neomysis mercedis</i>	FW/BW	MSP	521	this study
<i>Mysis mixta</i>		ERG	505-520	(Lindström 2000)
<i>Mysis relicta</i> sp.I		ERG	550-570	(Lindström 2000)
<i>Mysis relicta</i> sp.II		ERG	505-520	(Lindström 2000)
<i>Mysis relicta</i> sp.IV	FW	MSP	520	(Gal et al. 1999)
<i>Mysis relicta</i> sp.IV	FW	MSP	505, 520*	this study , (Jokela-Määttä et al. 2005)
<i>Praunus inermis</i>		ERG	520-530	(Lindström 2000)

<i>Praunus flexuosus</i>		ERG	505-515	(Lindström 2000)
Amphipoda				
Hyperiidae				
<i>Phronima sedenteria</i>	MP	ERG	470	(Frank and Widder 1999)
Isopoda				
Cirolanidae				
<i>Eurydice pulchra</i>				CO869196- CO157253 (EST)
Ligiidae				
<i>Ligia exotica</i>	TR	ERG	340, 460, 520	(Hariyama and Tsukahara 1993)
Talitridae				
<i>Talitrus saltator</i>	sTR	ERG	450	(Mezzetti and Scapini 1995)
Euphausiacea				
Euphausiidae				
<i>Euphausia superba</i>	MP	ERG	487	(Frank and Widder 1999)
<i>Euphausia superba</i>	MP	EX	485	(Denys and Brown 1982)
<i>Euphausia pacifica</i>	MP		462	(Kampa 1955)**
<i>Meganyctiphanes norvegica</i>	MP	ERG	490	(Frank and Widder 1999)
<i>Meganyctiphanes norvegica</i>	MP	EX	460-465	(Fisher and Goldie 1959)**
<i>Meganyctiphanes norvegica</i>	MP		462	(Fisher 1967)**
<i>Meganyctiphanes norvegica</i>	MP	ERG	460, 490, 515	(Boden et al. 1961)
<i>Meganyctiphanes norvegica</i>	MP	MSP	488	(Denys and Brown 1982)
<i>Nematobrachion boopis</i>	MP	ERG	488	(Frank and Widder 1999)
<i>Nematobrachion sexspinosus</i>	MP	ERG	478	(Frank and Widder 1999)
<i>Stylocheiron maximum</i>	MP	ERG	479	(Frank and Widder 1999)
<i>Stylocheiron maximum</i>	MP	EX	470	(Fisher and Goldie 1961)**
<i>Nematoscelis megalops</i>	MP	EX	465	(Fisher and Goldie 1961)**
<i>Thysanopoda acutifrons</i>	MP	EX	480	(Fisher and Goldie 1961)**
<i>Thysanopoda orientalis</i>	MP	ERG	478	(Frank and Widder 1999)

<i>Thysanoessa raschii</i>	MP	EX	460-465	(Fisher and Goldie 1961)**
Decapoda				
Dendrobranchiata				
Aristeidae				
<i>Plesiopenaeus armatus</i>	P	MSP	493	(Kent 1997)
Benthesicymidae				
<i>Bentheogennema intermedia</i>	P	MSP	494	(Kent 1997)
<i>Bentheogennema pasitheia</i>	P	MSP	500	(Kent 1997)
<i>Gennadas sp.</i>	P	MSP	495	(Kent 1997)
<i>Gennadas valens</i>	P	MSP	495	(Kent 1997)
Penaeidae				
<i>Funchalia villosa</i>	P	ERG	489	(Frank and Widder 1999)
<i>Penaeus duororum</i>	C	EX	516	(Fernandez 1965)
<i>Penaeus monodon</i>				AI770282, AI770242, AI770226, AI253885 (EST)
<i>Penaeus penicillatus</i>	C	ERG	480, 570	(Minjuan and Shujun 1990)
Sergestidae				
<i>Sergestes arcticus</i>	MP	ERG	495	(Frank and Widder 1999)
<i>Sergestes arcticus</i>	MP	extract	475	(Fisher and Goldie 1961)
<i>Sergestes corniculum</i>	MP	ERG	500	(Frank and Widder 1999)
<i>Sergestes curvatus</i>	P	MSP	493	(Kent 1997)
<i>Sergestes similis</i>	P	MSP	495	(Kent 1997)
<i>Sergestes similis</i>	MP	MSP	495	(Lindsay et al. 1999)
<i>Sergestes tenuiremis</i>	P	MSP	495	(Hiller-Adams et al. 1988)
<i>Sergia grandis</i>	MP	ERG	500	(Frank and Widder 1999)
<i>Sergia maximus</i>	P	MSP	495	(Kent 1997)
<i>Sergia phorcus</i>	P	MSP	495	(Kent 1997)
<i>Sergia robustus</i>	P	MSP	496	(Kent 1997)
<i>Sergia splendens</i>	P	MSP	497	(Kent 1997)

Pleocyemata

Caridea

Bresiliidae

<i>Rimicaris exoculata</i>	DB	ERG	500	(Johnson et al. 2002)
<i>Rimicaris exoculata</i>	DB	EX	500	(Van Dover et al. 1989)

Crangonidae

<i>Crangon allmani</i>	C	ERG	415, 525	(Johnson et al. 2002)
------------------------	---	-----	----------	-----------------------

Palaemonidae

<i>Palaemonetes paladosus</i>	EST	EX	539	(Fernandez 1965)
<i>Palaemonetes vulgaris</i>	EST	ERG	390, 540	(Wald and Seldin 1968)
<i>Palaemonetes vulgaris</i>	EST	MSP	496, 555	(Goldsmith et al. 1968)

Pandalidae

<i>Pandalus montagui</i>	C	ERG	515	(Johnson et al. 2002)
<i>Plesionika martia</i>	P	MSP	499	(Kent 1997)
<i>Stylopandalus richardi</i>	P	MSP	491	(Kent 1997)

Pasiphaeidae

<i>Parapasiphaea sulcatifrons</i>	P	MSP	501	(Kent 1997)
<i>Pasiphaea chacei</i>	P	MSP	509	(Kent 1997)
<i>Pasiphaea emarginata</i>	P	MSP	497	(Kent 1997)
<i>Pasiphaea multidentata</i>	P	ERG	497	(Frank and Widder 1999)
<i>Pasiphaea suspirosum</i>	P	MSP	501	(Kent 1997)

Oplophoridae

<i>Acanthephyra curtirostris</i>	P	ERG	510	(Frank and Case 1988a)
<i>Acanthephyra curtirostris</i>	P	MSP	485	(Hiller-Adams et al. 1988)
<i>Acanthephyra curtirostris</i>	P	MSP	485	(Kent 1997)
<i>Acanthephyra microphtalma</i>	P	MSP	482	(Kent 1997)
<i>Acanthephyra purpurea</i>	P	MSP	492	(Kent 1997)
<i>Acanthephyra smithi</i>	P	ERG	510	(Frank and Case 1988a)
<i>Acanthephyra smithi</i>	P	MSP	491	(Hiller-Adams et al. 1988)
<i>Acanthephyra stylorostratis</i>	P	MSP	489	(Kent 1997)
<i>Hymenodora frontalis</i>	P	MSP	495	(Kent 1997)

<i>Hymenodora glacialis</i>	P	MSP	500	(Kent 1997)
<i>Janicella spinacauda</i>	P	ERG	400, 500	(Frank and Case 1988a)
<i>Meningodora miccyla</i>	P	MSP	486	(Kent 1997)
<i>Meningodora vesca</i>	P	MSP	487	(Kent 1997)
<i>Notostomus elegans</i>	P	ERG	490	(Frank and Case 1988a)
<i>Notostomus gibbosus</i>	P	ERG	480	(Frank and Case 1988a)
<i>Oplophorus gracilirostris</i>	P	ERG	400, 500	(Frank and Case 1988a)
<i>Oplophorus spinosus</i>	P	ERG	400, 500	(Frank and Case 1988a)
<i>Oplophorus spinosus</i>	P	MSP	492	(Kent 1997)
<i>Systellaspis braueri</i>	P	MSP	411, 500	(Kent 1997)
<i>Systellaspis cristata</i>	P	MSP	414, 498	(Kent 1997)
<i>Systellaspis debilis</i>	P	MSP	400, 498	(Cronin and Frank 1996)
<i>Systellaspis debilis</i>	P	ERG	400, 500	(Frank and Case 1988a)
<i>Systellaspis debilis</i>	P	MSP	493	(Hiller-Adams et al. 1988)
<i>Systellaspis debilis</i>	P	MSP	417, 497	(Kent 1997)
Achelata				
Palinuridae				
<i>Jasus edwardsii</i>	C	ERG	472, 536	(Meyer-Rochow and Tiang 1984)
<i>Panulirus argus</i>	C	ERG	379, 510	(Cummins et al. 1984)
Anomala				
Diogenidae				
<i>Clibanarius vittatus</i>	EST	MSP	510	(Cronin and Forward 1988)
<i>Dardanus fucosus</i>	C	MSP	511	(Cronin and Forward 1988)
<i>Petrolisthes diogenes</i>	C	MSP	508	(Cronin and Forward 1988)
Coenobitidae				
<i>Coenobita clypeatus</i>	TR	MSP	508	(Cronin and Forward 1988)
<i>Coenobita rugosa</i>	TR	MSP	491	(Cronin and Forward 1988)
Paguridae				
<i>Pagurus annulipes</i>		MSP	493	(Cronin and Forward 1988)
<i>Pagurus longicarpus</i>	EST	MSP	515	(Cronin and Forward 1988)
<i>Pagurus pollicaris</i>	C	MSP	515	(Cronin and Forward 1988)

<i>Pagurus pollicaris</i>	C	MSP	516	(Lipetz and Cronin 1988)	
Galatheidae					
<i>Pleuroncodes planipes</i>	P	EX	523	(Fernandez 1973)	
Porcellanidae					
<i>Petrolisthes elongates</i>	C	ERG	536	(Ziedins and Meyer-Rochow 1990)	
Astacidae					
<i>Astacus fluviatus</i>	FW	MSP	530	(Hamacher and Kohl 1981)	
<i>Astacus leptodactylus</i>	FW	MSP	530	(Hamacher and Stieve 1984)	
Cambaridae					
<i>Cambarus hubrichti</i>	FW			(Crandall and Hillis 1997)	AF005385
<i>Cambarus maculatus</i>	FW			(Crandall and Hillis 1997)	AF005386
<i>Cambarellus schufeldtii</i>	FW	MSP	526	(Crandall and Cronin 1997)	AF003544
<i>Cambarellus ludovicianus</i>	FW	MSP	529	(Crandall and Cronin 1997)	AF003543
<i>Orconectes australis</i>					AF005387
<i>Orconectes virilis</i>					AF003545
<i>Orconectes rusticus</i>	FW	MSP	530-535	(Cronin and Goldsmith 1982; Goldsmith 1978)	
<i>Procambarus clarkii</i>	FW	MSP	530-533	(Goldsmith 1978; Zeiger and Goldsmith 1994)	S53494
<i>Procambarus clarkii</i>	FW	MSP	440	(Cummins and Goldsmith 1981)	
<i>Procambarus milleri</i>	FW	MSP	522	(Crandall and Cronin 1997; Cronin and Goldsmith 1982)	AF003546
<i>Procambarus orcinus</i>					AF005389
<i>Procambarus seminolae</i>					AF005388
Parastacidae					
<i>Engaeus cunicularius</i>	FW	MSP	522	(Crandall and Cronin 1997)	
Nephropidae					
<i>Homarus americanus</i>	C	MSP	515	(Bruno et al. 1977)	CN853478, CN854434 (ESTs)
<i>Homarus gammarus</i>	C	MSP	515	(Kent 1997)	

<i>Nephrops norvegicus</i>	C	MSP	498	(Kent 1997)
<i>Nephrops norvegicus</i>	C	MSP	498	(Lowe 1976)
<i>Nephrops norvegicus</i>	C	ERG	425?, 515	(Johnson et al. 2002)
Brachyura				
Bythograeidae				
<i>Bythograea thermydron</i>	DB	MSP	489	(Cronin and Jinks 2001)
<i>Bythograea thermydron</i> (L)		MSP	447	(Cronin and Jinks 2001)
<i>Bythograea thermydron</i>	BP	MSP	489	(Jinks et al. 2002)
<i>Bythograea thermydron</i> (M)	BP	MSP	479	(Jinks et al. 2002)
<i>Bythograea thermydron</i> (Z)	MP	MSP	447	(Jinks et al. 2002)
Calappidae				
<i>Calappa flammea</i>	C	MSP	486	(Cronin and Forward 1988)
<i>Calappa flammea</i>		MSP	483	(Lipetz and Cronin 1988)
<i>Hepatus epheliticus</i>	C	MSP	487	(Cronin and Forward 1988)
Cancridae				
<i>Cancer irroratus</i>	C	MSP	496	(Cronin and Forward 1988)
Gecarcinidae				
<i>Gecarcinus lateralis</i>	TR	MSP	487	(Cronin and Forward 1988)
<i>Gecarcinus lateralis</i>	TR	ERG	510	(Lall and Cronin 1987)
Geryonidae				
<i>Chaceon affinis</i>	DB	ERG	380?, 480	(Johnson et al. 2002)
<i>Geryon quinquedens</i>	DB	MSP	473	(Cronin and Forward 1988)
<i>Geryon quinquedens</i>	DB	MSP	470	(Lipetz and Cronin 1988)
Grapsidae				
<i>Leptograpsus variegatus</i>	EST	IC	484	(Stowe 1980)
<i>Sesarma cinereum</i>	EST	MSP	492	(Cronin and Forward 1988)
<i>Sesarma reticulatum</i>	EST	MSP	493	(Cronin and Forward 1988)
<i>Sesarma reticulatum</i>	EST	IC	508	(Scott and Mote 1974)
Homolidae				
<i>Paromola cuvieri</i>	DB	ERG	?, 470	(Johnson et al. 2002)
Majidae				

<i>Libinia dubia</i>	EST	MSP	489	(Cronin and Forward 1988)	
<i>Libinia dubia</i>		MSP	486	(Lipetz and Cronin 1988)	
<i>Libinia emarginata</i>	EST	MSP	493	(Hays and Goldsmith 1969)	
Ocypodidae					
<i>Uca pugilator</i>	TR	IC	508	(Scott and Mote 1974)	
<i>Uca pugilator</i>	TR	ERG	510	(Scott and Mote 1974)	
<i>Uca pugilator</i>	TR	EX	480	(Goldsmith 1972)	
<i>Uca pugnax</i>	TR	ERG	510	(Scott and Mote 1974)	
<i>Uca pugnax</i>	TR	IC	508	(Scott and Mote 1974)	
<i>Uca thayeri</i>	TR	ERG	430, 500- 540	(Horch et al. 2002)	
Portunidae					
<i>Arenaeus cribrarius</i>	C	MSP	498	(Cronin and Forward 1988)	
<i>Callinectes ornatus</i>	C	MSP	501	(Cronin and Forward 1988)	
<i>Callinectes sapidus</i>	EST	MSP	503	(Cronin and Forward 1988)	CV224458 (EST)
<i>Callinectes sapidus</i>	EST	IC	440, 508	(Martin and Mote 1982)	
<i>Callinectes sapidus</i>	C	MSP	504	(Cronin et al. 1995)	
<i>Callinectes sapidus</i> (M)	C	MSP	504	(Cronin et al. 1995)	
<i>Carcinus maenas</i>	EST	MSP	508	(Bruno and Goldsmith 1974)	
<i>Carcinus maenas</i>	EST	IC	440, 508	(Martin and Mote 1982)	
<i>Ovipales stephensonii</i>	C	MSP	505	(Cronin and Forward 1988)	
<i>Portunus spinimanis</i>	C	MSP	483	(Cronin and Forward 1988)	
<i>Portunus spinimanus</i>		MSP	479	(Lipetz and Cronin 1988)	
<i>Portunus trituberculatus</i>	C	ERG	513	(Weiyun and Minjuan 1990)	
<i>Scylla serrata</i>	EST	MSP	490	(Leggett 1979)	
Varunidae					
<i>Hemigrapsus edwardsii</i>	EST	EX	513	(Briggs 1961)	D50583- D50584
<i>Hemigrapsus sanguinensis</i>			480	(Sakamoto et al. 1996)	
<i>Hemigrapsus sanguinensis</i>	EST	ERG	360, 480	(Shukolyukov et al. 1984)	
Xanthidae					
<i>Eurypanopeus depressus</i>	EST	MSP	480	(Cronin and Forward 1988)	

<i>Menippe mercenaria</i>	EST	MSP	494	(Cronin and Forward 1988)
<i>Panopeus herbtii</i>	EST	MSP	493	(Fernandez 1973)
<i>Panopeus herbstii</i>	EST	MSP	491	(Lipetz and Cronin 1988)
<i>Panopeus obesus</i>	EST	MSP	492	(Cronin and Forward 1988)
<i>Pilumnus sayi</i>	C	MSP	489	(Cronin and Forward 1988)
<i>Rhithropanopeus harrisii</i>	EST	MSP	495	(Cronin and Forward 1988)

* - λ_{\max} based on porphyropsin template; ** - λ_{\max} may be contaminated by ommochrome pigments (Denys and Brown 1982)

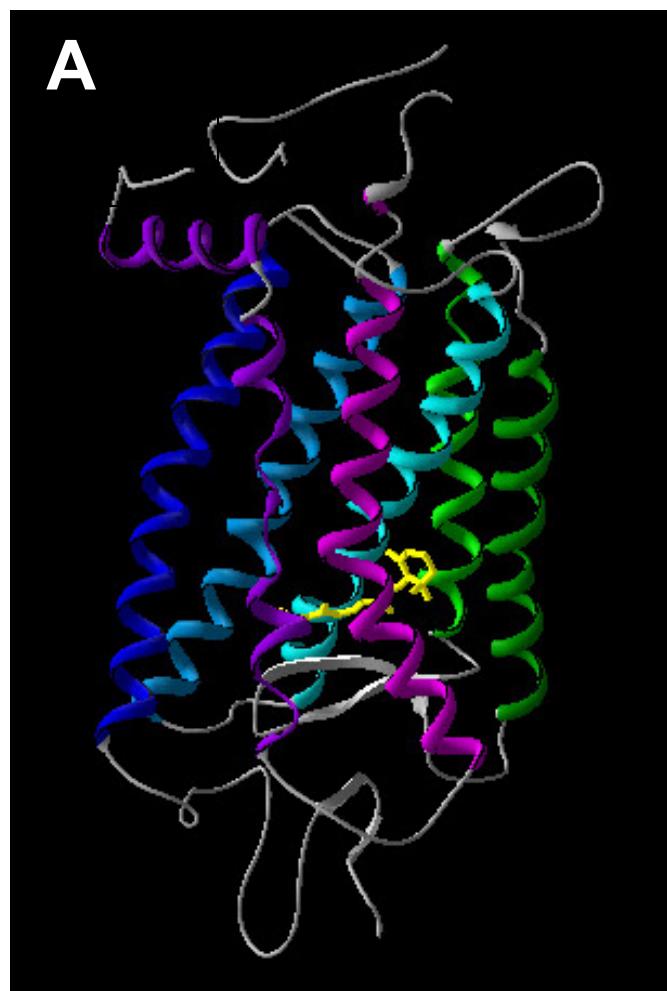
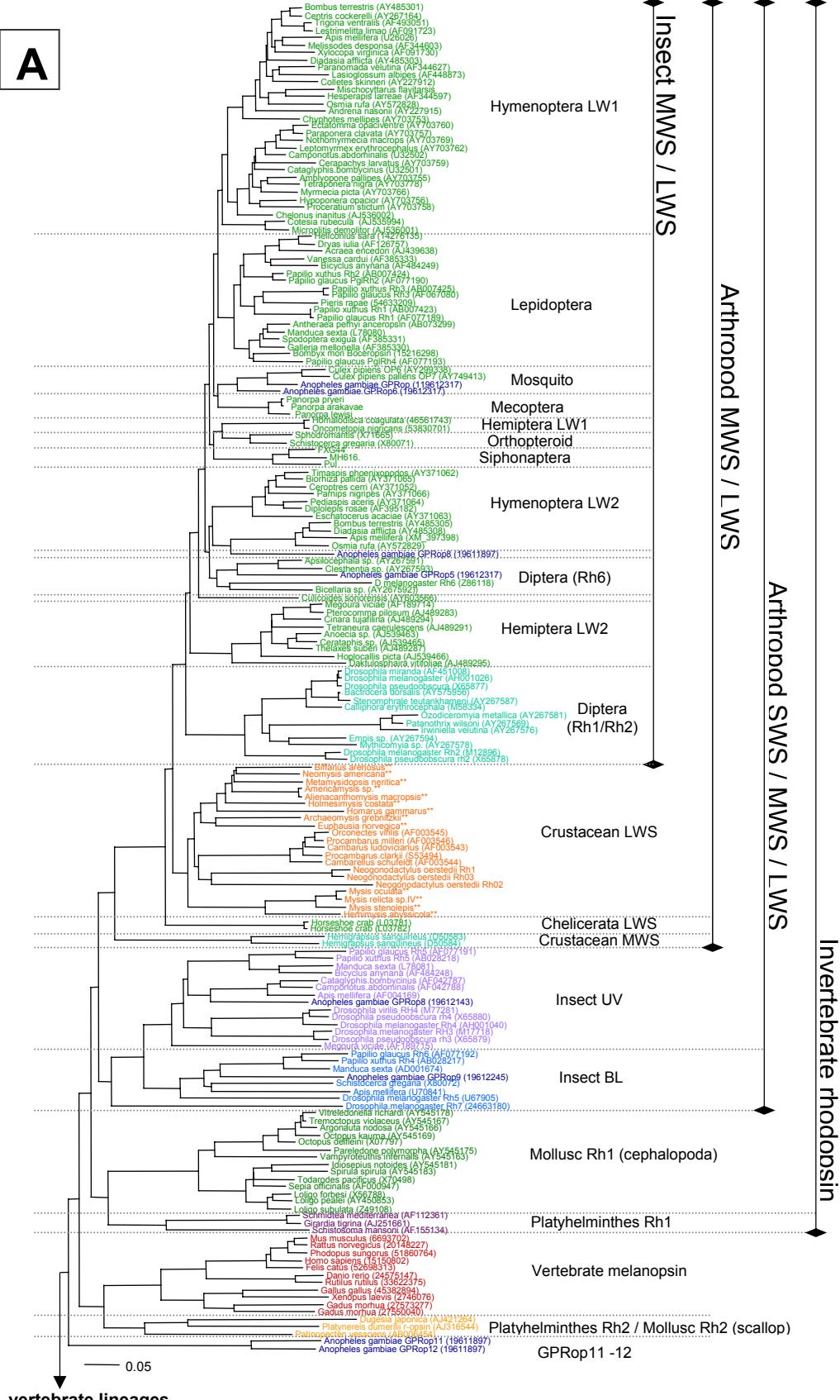


Figure 1-1. Three dimensional image of bovine rhodopsin produced using a high-resolution (2.6Å) bovine rhodopsin template (1L9H.pdb) from the Protein Data Bank (<http://www.rcsb.org/pdb/>) with the program Swiss-PdbViewer v.3.7 (<http://www.expasy.org/spdbv>; Guex and Peitsch 1997). The seven transmembrane domains are colored as follows: I – dark blue, II – medium blue, III – light blue, IV – light green, V – dark green, VI – pink, VII – purple. The chromophore is rendered in yellow and the non-transmembrane domains are grey. A) Bovine opsin oriented with the extracellular surface at the bottom and the cytoplasmic surface at the top of the image. B) View into the chromophore binding pocket from the cytoplasmic surface of the protein.



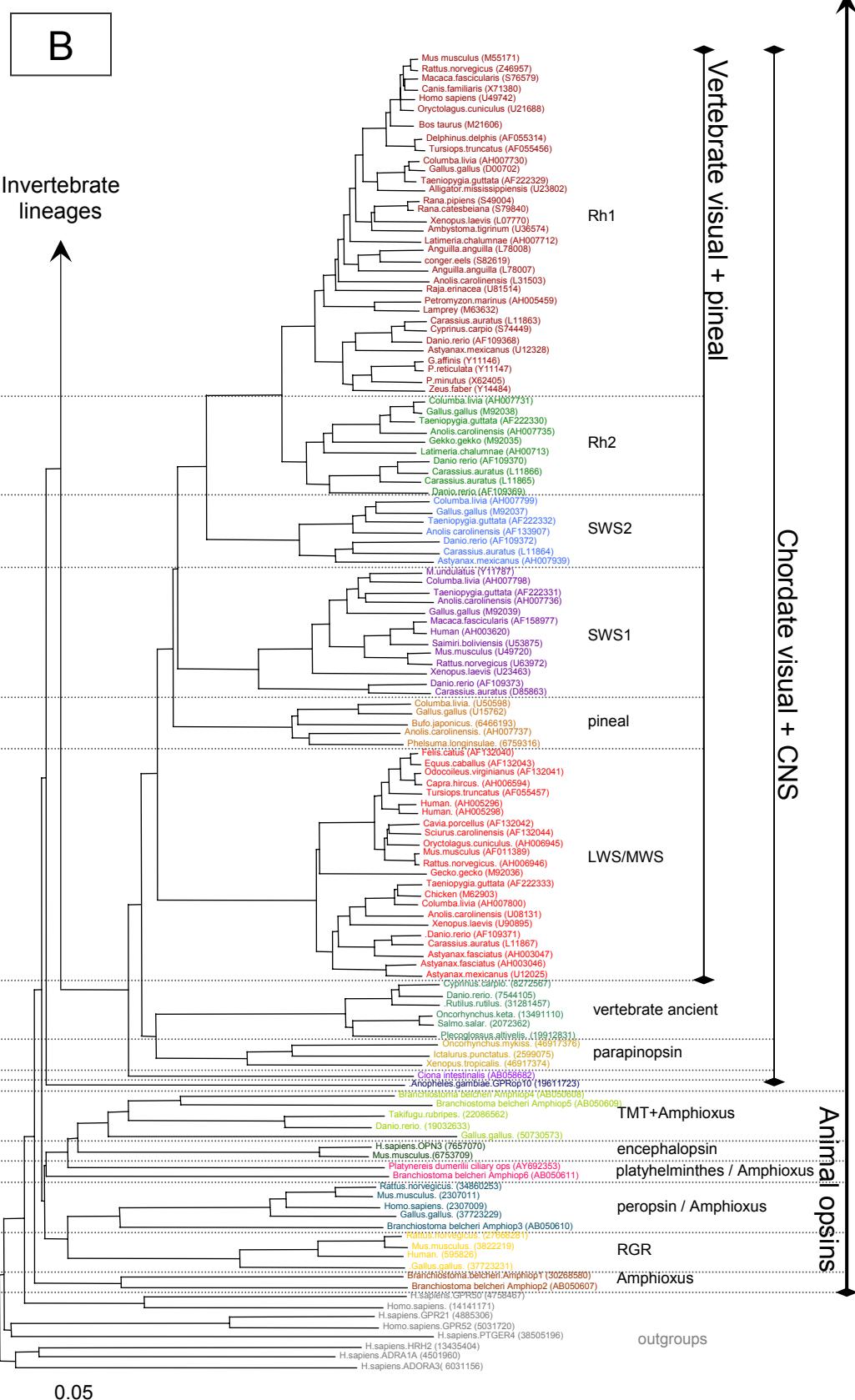


Figure 1-2. Neighbor-joining tree of amino acid data from a taxonomic selection of 310 animal opsins available from Genbank (<http://www.ncbi.nlm.nih.gov/>) combined with unpublished crustacean sequences from this research (indicated by ** following the species name). The accession numbers for each sequence from the database are given on the tree in parentheses after the species name. Major invertebrate and invertebrate clades are indicated on the tree. The tree is rooted to a selection of human GPCR lineages that are closely related to opsins based on previous phylogenetic analyses (Fredriksson et al. 2003).

CHAPTER 2.

LOST ALONG THE WAY:

THE SIGNIFICANCE OF EVOLUTION IN REVERSE

ABSTRACT¹²

Recently, researchers have begun to identify the prevalence of trait simplification, loss and reversal across all levels of biological organization. These studies have taken increasingly integrated approaches, incorporating phylogenetic, developmental and molecular methods, in the quest towards understanding the patterns and processes behind evolution in reverse. Here, we highlight the emerging interest in the reversibility of evolution by discussing a spectrum of studies examining both the genotypes and phenotypes of evolution in reverse. These integrative approaches have greatly increased our knowledge of the biological interactions that produce patterns of evolution in reverse and have led to promising new areas of research.

INTRODUCTION

‘Rudimentary, atrophied, or aborted organs. Organs or parts in this strange condition, bearing the stamp of inutility, are extremely common throughout nature’ (Darwin 1859).

¹ This chapter was published as: Porter, M.L. and K.A. Crandall. 2003. Lost along the way: the significance of evolution in reverse. *Trends in Ecology and Evolution* 18(10):541-547.

² Words in all caps throughout the chapter are defined in a glossary found in Appendix 1

Evolution in reverse is a widespread phenomenon in biology; however, many researchers are only just beginning to take notice of the significance and prevalence of trait loss and/or simplification (Wiens 2001). Part of this disregard is due to conflict among researchers about the validity of the concept of evolution in reverse. Many would argue that most commonly cited examples of REVERSE EVOLUTION (see Glossary) are actually *de novo* forms that have no relationship to ancestral states. Even if the concept is accepted, studies of the reversibility of evolution have been difficult to identify owing to confusion over what qualifies as ‘reverse evolution’. Terms such as simplification, REGRESSION, and REVERSION all refer to some form of reverse evolution (Appendix 2). In the strictest sense, reverse evolution has been defined as the reacquisition by derived populations of the same character states as those of ancestor populations (Teotónio and Rose 2001). But, in many natural systems, the character state(s) of the ancestor population is unknown, making reversions under these criteria unidentifiable. Additionally, evolution in reverse has been identified at various biological levels of organization, including phenotypes (structure, function, or behavior) and genotypes (gene deletions and back mutations) both within and among populations (Appendix 3). By restricting ‘reverse evolution’ to a process occurring only within populations, many cases might be misidentified as reverse evolution when, in fact, what is being observed is simply shifting allele frequencies, rather than the reversal of a fixed trait. Encompassing all of these related patterns, reverse evolution is an influential process in evolution, capable of forcing multiple diverged populations and species to converge on similar forms (Culver and Wilkens 2000; Teotónio and Rose 2002), overcoming evolutionary constraints that can impede diversification (Emlen 2001; Wake 1992; Whiting et al.

2003), and effectively ‘pruning’ unnecessary structures, functions and behaviors, enabling new evolutionary pathways to be explored (Borowsky and Wilkens 2002).

With the identification of new patterns and processes of evolution in reverse, several questions have become major areas for discussion. First, there has been considerable debate over how long a group of organisms must travel an evolutionary path before evolution becomes irreversible. Within the boundaries where evolution is reversible, the questions become more mechanistic: to what degree does evolutionary history constrain reverse evolution and what are the genetics behind reverse evolution? (Teotónio and Rose 2001). Here, we explore these questions by providing case studies that investigate the probability, potential mechanisms and evolutionary implications of reverse evolution.

Is evolution reversible?

One of the underlying issues of reversibility is whether evolution is actually reversible. In other words, can an organism retrace a previously traversed evolutionary pathway? Although many researchers would not argue with the potential to retrace pathways over short evolutionary time spans (i.e. several hundred generations) (Teotónio and Rose 2000; Teotónio and Rose 2002), the longer the path traveled, the more difficult it becomes to return. Therefore, the longer the time spans since diversifying from an ancestral state, the more researchers tend to accept the irreversibility of evolution (Teotónio and Rose 2001). Several recent studies focusing on opposite extremes of the reversibility spectrum with respect to time span have addressed this issue and have illustrated that evolution in reverse is achievable, both in short, experimentally controlled

studies of populations and over long evolutionary histories encompassing the diversification of large groups of species.

In *Drosophila melanogaster*, the effects of several hundred generations of evolution within a particular environment on fitness-related characters were undone in as little as 20 generations after the ancestral environment was re-imposed (Teotónio and Rose 2000). This study provided empirical proof that, over short time spans, evolution can be reversed for particular phenotypes. However, not all of the characters investigated reverted completely, if at all, demonstrating that phenotypic convergence on an ancestral form is not a universal possibility, even after only a few hundred generations of selection (Teotónio and Rose 2000; Teotónio and Rose 2002). Additionally, even over short time spans, caution must be taken when labeling a feature as having returned to an ancestral condition. It is not always obvious whether the reversal is an actual return to a primitive state, or a *de novo* convergent form approximating the primitive state. In microbial systems, changes in genotype can be directly linked to phenotypic changes and can be observed over time, making an ideal study system for CONVERGENT EVOLUTION. Studies of the reversal of antibiotic resistance have found that, once adapted to the selective regime imposed by a particular antibiotic, microbes are unlikely to return to the ancestral fitness conferred by sensitivity to antibiotics, even after removal of the antibiotic (Levin et al. 2000; Schrag et al. 1997). Instead, fitness is recovered by compensatory rather than back mutations of the changes conferring resistance, making the ‘re-evolved’ strain genetically distinct from the ancestor.

Even with the demonstration that evolution is reversible over several hundred generations, the possibility of evolution in reverse over longer evolutionary time periods

or evolution of more complex characters would still be rejected by many. Yet, recent phylogenetic investigations of the Phasmatodea (stick insects) indicate that it is possible to ‘re-evolve’ complex structures, as demonstrated by the multiple, independent re-acquisitions of wings in a group where the most recent common ancestor had lost wings. Furthermore, detailed studies of wing morphology and flight-specific musculature and innervation suggest that the reacquired wings are a re-expression of basic insect wings, rather than of convergent *de novo* forms. The studies of these insects illustrate that the basic blueprints for complex developmental structures can remain largely intact even over large evolutionary spans (i.e. radiations of higher level taxonomic groups), although the specific mechanisms of loss and/or reacquisition might differ between lineages (Whiting et al. 2003).

The continuum of reversibility

Although evolution in reverse is often described only as a return to an ancestral character state, a much broader spectrum of processes fall into this category, from traits that are completely or partially reversible to a recent ancestral state, to those complex traits termed ‘regressive’ that have been reduced or even lost. The difference between these trends is the evolutionary distance between the extant organism and the ancestral state that it is returning towards. Because ancestral states are generally unknown entities, the most commonly studied form of evolution in reverse is phenotypic regression. Although many studies have been devoted to characterizing the observed patterns of regressive evolution in nature, investigations have begun, more recently, to look at the underlying genetics of these systems. Here, we discuss several examples of systems that have

investigated not only phenotypic, but also the genetic, molecular and developmental mechanisms of evolution in reverse to illustrate the similarities and disparities between different forms of reversibility.

Phenotypic regression

The most frequently studied form of evolution in reverse is phenotypic regression, the vestigialization of morphological, physiological, or behavioral traits. Some of the most commonly cited examples include the structural simplifications of parasites, loss of limbs in snakes, flightlessness in birds and insects, and the loss of photosynthetic ability in parasitic plants. More recent work on sexual selection indicates that losses of elaborate male traits are taxonomically widespread, with several well studied groups (e.g. tanagers and dabbling ducks) containing loss:gain ratios as high as 5:1 for elaborate traits (Omland 1997; Omland and Lanyon 2000; Wiens 1999).

Perhaps the most powerful example, however, is observed in cave-dwelling organisms. The worldwide convergence of form found in the cave environment, exhibited in structural, functional and behavioral regressive changes across diverse taxonomic groups has fascinated and perplexed biologists since Darwin (Figure 2-1) (Culver and Wilkens 2000; Nevo 1999). Termed ‘troglomorphy’, this suite of changes includes reduction in pigment and eye size, hypertrophy of nonoptic sensory organs and a reduced metabolic rate (Borowsky and Wilkens 2002; Culver et al. 2000). One of the best studied cases is the teleost *Astyanax mexicanus*, a fish species that includes both eyed surface and eyeless cave-dwelling populations (Jeffery 2001; Jeffery and Martasian 1998; Yamamoto and Jeffery 2000). Although adults lack eyes to varying degrees, cavefish embryos

develop small optic primordia, which degenerate during ontogeny (Yamamoto and Jeffery 2000) via a series of steps involving both molecular and developmental mechanisms (Appendix 4). Furthermore, although several of the major developmental steps involved in cavefish eye degeneration are understood, whether selective or neutral forces underlie these steps is still ambiguous. Studies of the interaction between the expression of the master eye control gene, *Pax6*, and midline-signaling genes, such as *sonic hedgehog (shh)*, imply that the constructive changes in enhanced midline signaling activity might be important factors in controlling *Pax6* expression and therefore cavefish eye degeneration (Jeffery 2001; Strickler et al. 2001).

Genomics in reverse

At the molecular level, studies of evolution in reverse have investigated both reversions, in the form of microbial evolution in relation to antibiotic resistance and alternating host environments, and the regression of entire genomes. In microbial evolution, many have questioned whether environmentally induced genetic transformations can return to an ancestral genotype. Investigations of viral adaptation have provided some of the few clear cases of genotypic reversion (Crill et al. 2000; Massey et al. 2001; Nielsen et al. 2001).

When the bacteriophage ϕ X174 was switched between *Escherichia coli* and *Salmonella* hosts, original fitness levels were recovered by reversion of two to three substitutions in the major capsid gene controlling host recognition, rather than by compensatory mutations (Crill et al. 2000). The natural host range expansion of parvovirus shows the same pattern as the experimentally controlled bacteriophage host alternations. An originally feline parvovirus jumped hosts to canines in the early 1970s as a result of up to

five substitutions in the capsid protein (Hueffer et al. 2003; Simpson et al. 2000). After host range expansion, the canine parvovirus regained the ability to infect felines owing to changes in the same region of the capsid gene (Ikeda et al. 2000; Ikeda et al. 2002).

At the genome level, evolution in reverse takes the form of ‘streamlining’ or molecular noise suppression, where the transmission of redundant information is diminished by reduction, inactivation, or elimination of unimportant genetic material (Gil et al. 2002). The plethora of parasitic and symbiotic intracellular microorganisms and the multiple, independent occurrences of endosymbiotically derived plastids provide molecular equivalents to organismal habitat shifts inducing rudimentation. The establishment of permanent parasitic or symbiotic relationships usually leads to massive gene loss, resulting in the smallest known genomes for cellular organisms (Ochman and Moran 2001). These gene-level deletions occur because of a relaxation of constraints for functions that are no longer needed or that are already performed in the new host environment, host partitioning of population structure and the inaccessibility of foreign DNA as a source of gene acquisition (Appendix 5) (Andersson et al. 1998; Moran and Wernegreen 2000; Ochman and Moran 2001; van Ham et al. 2003). Studies of DNA in ‘nongreen’ (e.g. parasitic) plants show that the plastid genome is being streamlined, exhibiting an extreme reduction in gene content and an increase in the rate of evolution of the remaining genes (DePamphilis et al. 1997). In the nonphotosynthetic plant *Epifagus virginiana*, only 17 of the 30 tRNA genes and 21 of the 79 protein genes normally found in angiosperm plastid DNA remain in the genome (Lohan and Wolfe 1998). In the case of microbial genome shrinkage, massive gene losses become irreversible owing to the

specialized environment and the lack of encounters with foreign DNA (Ochman and Moran 2001).

How does evolution back up?

Many mechanisms have been proposed for affecting the reversibility of evolution, including both impeding (lack of genetic variation, small population sizes, EPISTASIS and novel genotype-by-environment interactions) and facilitating (PLEIOTROPY, GENETIC HITCHHIKING and mutation) factors (Borowsky and Wilkens 2002; Teotónio and Rose 2001). However, the emerging picture of the processes responsible for evolution in reverse is one of contingency. In short-term experimental studies of *Drosophila*, complete reversions are not universal and the incomplete reversions are not due to either lack of genetic variation or epistatic influences, because hybrids exhibit similar patterns of reversibility (Teotónio and Rose 2000). Instead, the degree of reversion is highly sensitive to environmental conditions and, at least partly, dependent on previous selective histories (i.e. genotype-by-environment interactions) (Teotónio and Rose 2002). However, the process of reverse evolution can drive the convergence of multiple populations with different life-history and genetic changes to a common character state (Teotónio and Rose 2002), similar to the widespread convergence of form observed in troglomorphic taxa.

In subterranean-adapted organisms, developmental and linkage studies point to pleiotropy as a significant process involved in reverse evolution (Borowsky and Wilkens 2002; Strickler et al. 2001). In quantitative trait loci (QTL) studies of the cave-dwelling *A. mexicanus*, close linkages were found between a regressive and a constructive trait in

two cases. These close linkages suggest that the regressive evolution of pigmentation and eye loss might be influenced by either pleiotropy or genetic hitchhiking (Borowsky and Wilkens 2002). Concordantly, the potential interaction of the *shh* gene with *Pax6* also indicates pleiotropy as an important mechanism (Jeffery 2001; Strickler et al. 2001).

In all of these examples, the most important factor is the interaction of traits and their underlying genetic backgrounds with the environment: in the *Drosophila* studies, returning to an ancestral environment after the evolution of divergent genetic backgrounds led to the differential reversibility of traits; in studies of *A. mexicanus*, selection pressure leading to adaptations to the subterranean environment (e.g. enhanced midline signaling) appear to interfere with the expression of genes involved in the development of the eye. In a particularly elegant study of the costs of horn production in *Onthophagus* beetles, Emlen (2001) demonstrated that there is a tradeoff between the development of exaggerated horns and the reduction of neighboring morphological structures (antennae, eyes, or wings). Because the *Onthophagus* beetles inhabit different physical environments, the position of the exaggerated horns in each species is correlated with the reduced structure that imposes the smallest cost.

By contrast, studies of asexual lineage evolution, where there are more direct correlations between genotypic and phenotypic changes, seem to be more influenced by epistatic interactions (Burch and Chao 1999; Elena and Lenski 2001). Studies of microbial antibiotic resistance lead to a return towards ancestral phenotypic characters through epistatic compensatory mutations rather than via back mutations (Maisnier-Patin et al. 2002; Rokyta et al. 2002). These epistatic intermediate-fitness compensatory mutations are more common than are higher fitness revertants, owing to the higher rate of

compensatory mutations and the characteristic population bottleneck dynamics of parasitic and symbiotic microbes (Levin et al. 2000).

The future of reversibility

The studies highlighted here illustrate the importance of reverse evolution as an influential evolutionary phenomenon. Ignoring patterns of reversibility might have important consequences for evolutionary analyses, particularly in phylogenetic reconstructions. Studies of dabbling ducks and orioles caution against using these types of character, particularly with parsimony methods that minimize homoplasy, in constructing phylogenies in general and against equally weighted gain:loss ratios in parsimony-based ancestral state reconstructions (Cunningham et al. 1998; Omland and Lanyon 2000). Acknowledging the reality of evolution in reverse, however, necessitates incorporating the often confusing terminology encompassing reverse phenomena into a more comprehensive framework for investigating this process. Further detailed genetic studies of a wide array of examples of reverse evolution are therefore needed to explore the interaction of ancestral evolutionary distance (time span) and reversibility. The next step in many studies of evolution in reverse is to identify the underlying genetic changes accompanying observed phenotypic reverersions and/or regressions. Once the crucial genetic transformations in regressed systems have been identified, these candidate loci can be used for detecting the operational selective forces (McClellan 2000; McClellan and McCracken 2001; Yang and Bielawski 2000). In studies of reversion, the direct comparison of ancestral and derived experimental populations is impossible in most natural settings (Teotónio and Rose 2000; Teotónio and Rose 2002). Additional natural

systems could be investigated using molecular methods, such as network approaches, which have been developed to deal specifically with intraspecific data, providing insights into the patterns involved in population-level evolutionary processes (Posada and Crandall 2001).

Furthermore, with genotype-by-environment interactions identified as one of the crucial factors affecting trait reversibility, several studies are already devoted to investigating the importance of the genetic architecture component of this interaction. But this interaction also needs to be approached from the opposite perspective, investigating the environmental effect in reverse evolution. By researching the occurrence of reversibility across a wide range of ecological types, the similarities and disparities across systems in terms of the effects of factors such as epistasis, pleiotropy, genetic variation and compensatory mutations can be investigated. Similarly, additional experimental studies of the effects of adaptation to varied ecological regimes on the reversion of similar genotypes are warranted.

Acknowledgements

We thank W.R. Elliott, H.H. Hobbs III, J. Jasper, W.R. Jeffery, J. Krejca, M. Pérez-Losada and A. Whiting, and anonymous referees for helpful comments and/or the use of their photographic images. This work was supported by grants from the Cave Conservancy Foundation, the Cave Research Foundation, and NSF Grants DEB-0120718 and DEB-0206537.

APPENDIX 1. GLOSSARY

Convergent evolution: a process in which changes occur from different ancestral character states to the same descendent character state in independent evolutionary lineages (Crandall 2002).

Epistasis: the nonreciprocal interaction of nonallelic genes, for example where one gene masks the expression of another (King and Stansfield 1997).

Genetic hitchhiking: the spread of a neutral allele through a population because it is closely linked to a beneficial allele and therefore is carried along as the gene that is selected for increases in frequency (King and Stansfield 1997).

Pleiotropy: the phenomenon in which a single gene is responsible for a number of distinct and seemingly unrelated phenotypic effects (King and Stansfield 1997).

Regression: relative to either the ancestral condition or closely related species, characters that are atrophied or degenerate and often without visible function as a result of an evolutionary change in lifestyle. Continued degeneration might lead to character loss. (also termed vestigialization) (Muller 2002).

Resident genome: genome of a bacterium that lives in close, often intracellular, association with a eukaryotic host (Moran and Wernegreen 2000).

Reverse evolution: the change of a character state to a state similar in appearance to an ancestral state, encompassing patterns associated with both reversion and regression.

Reversion: in reference to fixed character states, when a derived state evolves to a state present in ancestral lineages.

APPENDIX 2. STRAYING FROM THE PATH: THE DEBATE OVER EVOLUTION IN REVERSE

Discussing reverse evolution as an influential evolutionary phenomenon is fraught with problems, the least of which is the debate about whether it actually exists. Furthermore, although the concepts of regression, reversal and loss are not contentious, the idea of these being interrelated processes is. However, we contend that reverse evolution is a phenomenon much like speciation, where different processes can result in a similar pattern of evolution. We therefore unite the processes of reversion, regression and loss under the term of reverse evolution and endeavor to outline their similarities and differences.

For our purposes, we define reverse evolution as a change in character state to one resembling an ancestral state. We recognize that, owing to diverse evolutionary histories, reacquisition of an exact ancestral state is improbable in many cases. However, even the approximation of an ancestral state, although perhaps novel in form, indicates an interrelated set of evolutionary processes. Although many terms have been used to indicate some form of reverse evolution, most of the observed patterns can be encompassed by three types of change: regression, loss and reversion. Regression is related to the ideas of simplification or vestigialization, and includes a gradual reduction in functionality whilst still retaining some form of the original feature. After a reduction in utility is underway, the character can remain in some vestigial form or can be lost completely. The difficulty for many in accepting this as evolution in reverse is that the reduced characters usually only resemble an ancestral state. Furthermore, is the loss of a structure the same state as an ancestral state lacking that structure? Although a structure

might be lost, constraints might preserve developmental or genetic pathways that are not present in the ancestor (Crandall and Cronin 1997; Whiting et al. 2003).

In contrast to the reduction and loss of a trait, characters that evolve to known ancestral states are easier to accept as examples of reverse evolution. Not even this classification is straightforward, however, because reversion to an ancestral state is often accomplished using mechanisms that are not present in ancestral populations (Maisnier-Patin et al. 2002; Rokyta et al. 2002). Phenotypic similarity might mask mechanistic or developmental differences, making the classification of evolution in reverse dependent on the level of organization being studied. However, although the similarity to ancestral forms can vary from exact features to mere approximations, the novel pathways and forms used to accomplish these similarities are what make studies of evolution in reverse worthwhile. Reverse evolution offers organisms alternative routes for moving around functional constraints and evolutionary dead-ends.

APPENDIX 3. EXPERIMENTAL SYSTEMS IN REVERSE EVOLUTION

Because reverse evolution as a pattern of evolution has been identified across various levels of biological organization (genotype and phenotype), taxonomic diversity (microorganisms to vertebrates), and evolutionary groups (populations and species), by necessity, different terms and methodologies have been employed. Although these studies deal with similar issues, they are not necessarily directly comparable in terms of the form and function of reverse evolution. For example, owing to the differing reproductive strategies (asexual versus sexual) and genetic architecture, it would not be expected that microbial systems and vertebrate systems exhibit similar mechanisms of regressions and reversals. Also, studies of microbial systems have the unique advantage of having better characterized genomes, and changes at the molecular level can often be quantified directly at the phenotypic level. Perhaps the most important advantage to microbial systems, however, is the ability to control environmental factors and to produce replicates of a system of interest for comparative purposes (Burch and Chao 1999; Crill et al. 2000; Elena and Lenski 2001). This advantage can also be recognized in other laboratory systems (e.g. *Drosophila*) (Teotónio and Rose 2002). Although these tightly controlled experimental systems enable specific genetic interactions with the environment to be investigated and can provide indications of patterns and mechanisms, they might not be directly comparable to more natural systems, where replications under different environmental influences might not exist. In more natural systems, it becomes more crucial to find the appropriate controls for comparisons, for example populations of eyed surface fish versus eyeless cave populations in *Astyanax mexicanus* (Jeffery 2001).

APPENDIX 4. THE GENETIC AND DEVELOPMENTAL MECHANISMS BEHIND REGRESSIVE EVOLUTION IN *ASTYANAX MEXICANUS*

One of the best studied systems of phenotypic regression is the Mexican tetra *Astyanax mexicanus*. Because *A. mexicanus* contains both epigean and at least 29 different populations of cave-adapted forms, studies of drastic morphological changes in isolated populations of the same species are possible (Jeffery and Martasian 1998). These changes include both constructive changes (i.e. increased complexity of feeding apparatus and the mechanosensory system of cranial neuromasts) and regressive changes (i.e. loss of eyes, pigmentation and aggressive behavior) (Jeffery 2001). Of this suite of characters, particular attention has been paid to eye degeneration in the cave-adapted form of *A. mexicanus*, whose embryos form eye primordia that later degenerate and sink into the orbit (Strickler et al. 2002). Based on morphological, biochemical and phylogenetic studies, the different cave populations comprise at least two genetically distinct lineages with similar eyeless phenotypes, resulting from independent invasions from surface populations (Dowling et al. 2002). Additionally, crosses between geographically isolated cavefish populations can produce progeny with a greater degree of eye development than that exhibited by either parent, indicating that different cave populations have evolved different mechanisms of degeneration (Jeffery 2001). Within a single population, studies have shown that the genetics of regression are multifactorial, with at least three quantitative trait loci mapped to reduced eye size (Borowsky and Wilkens 2002).

Developmental studies have shown that both reduced expression of the *Pax6* gene, a crucial controller of eye development, during early eye development and apoptosis of the lens contribute to eye regression. In an elegant study of eye degeneration, *A. mexicanus*

lens tissue was reciprocally transplanted between developing surface and cavefish embryos. In cavefish eyes receiving a surface fish lens, eye structure was recovered, whereas in surface fish with a cavefish lens, degeneration was observed (Yamamoto and Jeffery 2000). This study illustrates that the lack of an inductive signal emanating from the lens is a major cause of eye regression in cave populations (Jeffery and Martasian 1998; Yamamoto and Jeffery 2000). The current understanding of the developmental steps in eye regression are the following: (1) *Pax6* expression is reduced at the anterior midline; (2) potentially as a result of this altered expression, a smaller lens and optic vesicle/cup are formed; (3) cavefish lenses undergo apoptosis instead of differentiation; (4) in the absence of lens signaling, further eye structures fail to develop; and (5) the eye collapses into the orbit and is covered by a flap of skin (Figure 2-2) (Jeffery 2001).

APPENDIX 5. APHIDS AND *BUCHNERA*: STUDIES OF GENOMICS IN REVERSE FROM A MUTUALISTIC ENDOSYMBIONT

One of the best studied systems of regressive genomic reduction comes from the coevolved mutualistic system of aphids and their primary endosymbiont bacteria, *Buchnera*. Evolved from a single bacterial infection of an ancestral aphid 100–250 million years ago (Martinez-Torres et al. 2001), the symbiotic relationship between aphids and *Buchnera* is obligate for both species and neither can reproduce without the other (Martinez-Torres et al. 2001). As a result of this long association, *Buchnera* exhibit characteristic features of ‘RESIDENT’ GENOMES, including severe genome reduction, extreme adenine-thymine bias and fast sequence evolution at all loci (Gil et al. 2002). In particular, the regressive genome reduction in *Buchnera* shares similarities with other reductive genomes, including the loss of loci encoding DNA repair, recombinase functions and nonessential amino acid biosynthesis pathways (Shigenobu et al. 2000; van Ham et al. 2003). Most of the reduction in genome size (65–74%) occurs soon after the establishment of the symbiosis but before the diversification of the major lineages of extant aphids, suggesting that genome reduction proceeds at an exponentially decreasing pace (Figure 2-3) (van Ham et al. 2003). This is also supported by the conserved genetic architecture of *Buchnera*, with no chromosome rearrangements or gene acquisitions having occurred within at least the past 50–70 million years (Tamas et al. 2002). However, the reductive process is ongoing as evidenced by lineages with further genome reductions, such as the recent 25-kb deletion of an already greatly reduced *Buchnera* genome from the aphid *Cinara cedri* (Gil et al. 2002).

Rather than adaptive changes, these major reductions are attributed to degenerate genome evolution via a mutational bias favoring deletions, reflecting a lack of effective selection for gene maintenance (Ochman and Moran 2001). The combination of rampant deletions and lack of selection is exacerbated by a suite of factors associated with the specialized host environment, including genetic isolation, small effective population sizes and the loss of DNA repair mechanisms (Ochman and Moran 2001; van Ham et al. 2003). Genetic isolation and the loss of recombination pathways effectively reduce the probability of gene acquisition to zero, whereas small effective population sizes enable an irreversible accumulation of deleterious mutations owing to drift. In *Buchnera* species, these reductive processes have led to some of the smallest genomes yet described (Figure 2-3) (Gil et al. 2002).

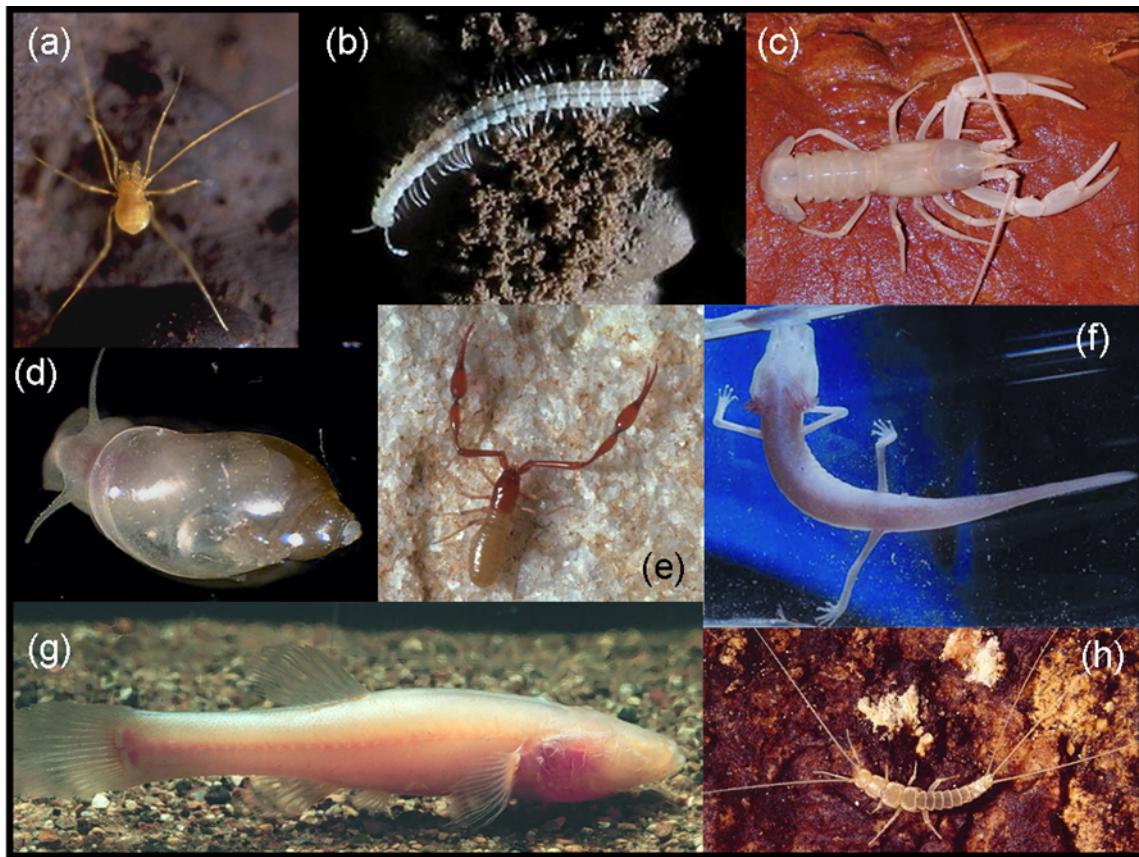


Figure 2-1. Examples of convergence across diverse taxonomic groups owing to similar subterranean habitats. **(a)** harvestman, *Texella reyesi*; **(b)** millipede, *Speodesmus* n.sp.; **(c)** crayfish, *Orconectes stygocaneyi*; **(d)** snail, *Physa spelunca*; **(e)** pseudoscorpion, *Australinocreagris grahami*; **(f)** cave salamander, *Eurycea rathbuni*; **(g)** cave fish, *Amblyopsis spelaea*; and **(h)** silverfish, *Texoreddellia texensis*. Reproduced with permission from W.R. Elliott (a, c, e), W.R. Elliott and J. Ivy (h), J. Jasper (b), J. Krejca (f), H.H. Hobbs III (g) and M.L. Porter (d).

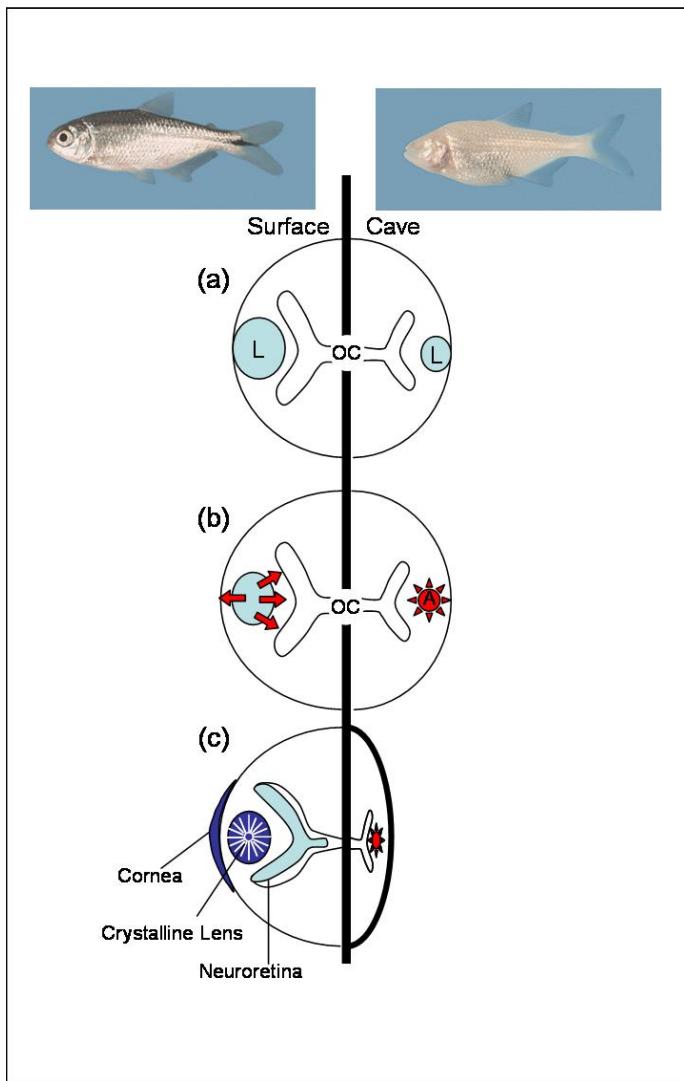


Figure 2-2. Eye development in surface versus cave populations of *Astyanax mexicanus*. (a) Smaller lens and optic cup form in cavefish embryos; (b) cavefish lenses undergo apoptosis, whereas surface fish functional lenses exhibit normal signaling; (c) surface fish lens signaling plays a role in inducing development of eye structures, including the cornea, iris, pupil and retinal photoreceptors, whereas in the absence of signaling, cavefish eyes degenerate. Abbreviations: L, lens; OC, optic cup; A, apoptosis; Reproduced with permission from (Jeffery and Martasian 1998) and W.R. Jeffery.

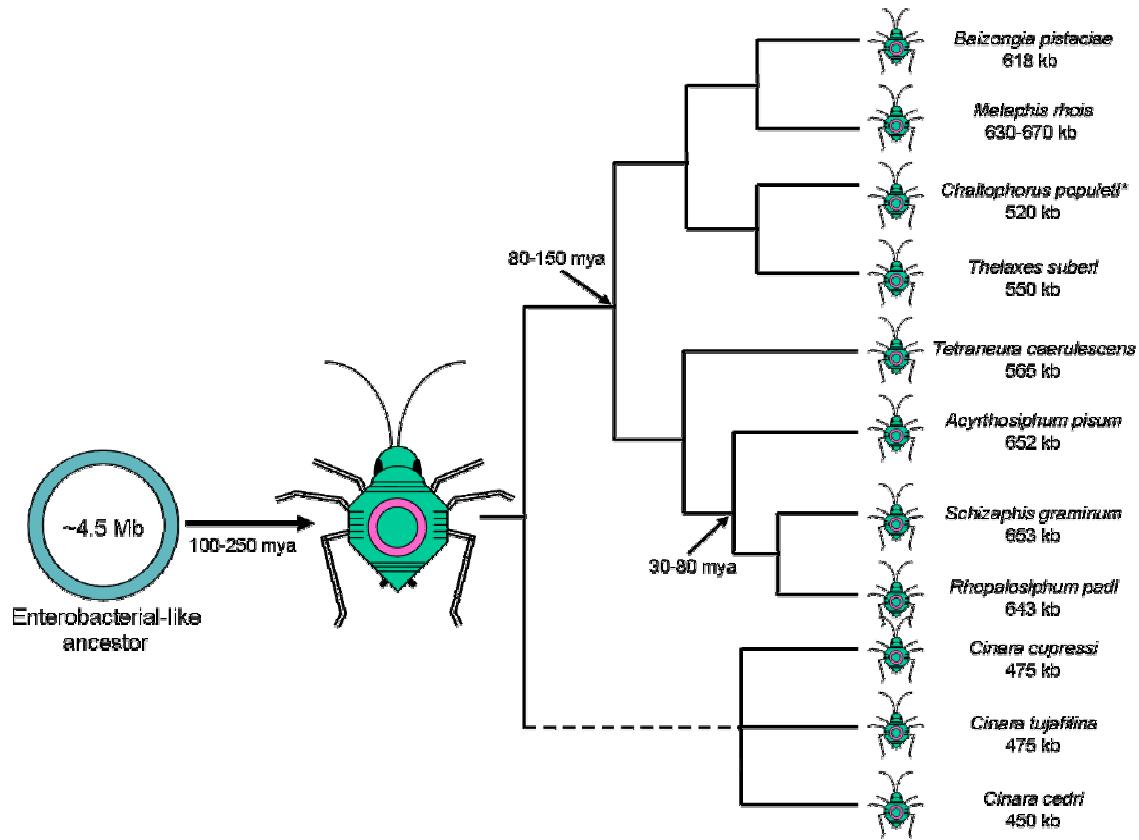


Figure 2-3. Evolutionary history of aphids and their endosymbiont bacteria, *Buchnera*. Divergence times are indicated for selected nodes (Data from Martinez-Torres et al. 2001; Moran et al. 1993; van Ham et al. 2003). *Buchnera* genome size is given under the name of each host species (Data from Gil et al. 2002; van Ham et al. 2003; Wernegreen et al. 2000). The phylogeny is based on maximum likelihood analyses of *Buchnera* 16S rDNA sequences presented in (Martinez-Torres et al. 2001). The dashed line indicates positioning on the tree based on other taxa from the same family, and from analyses of aphid sequences. * *Chaitophorus populeti* is placed on the tree based on representative species from the same genus.

CHAPTER 3.

PHYLOGENETIC RELATIONSHIPS WITHIN THE MYSIDAE

(MYSIDA, CRUSTACEA)

ABSTRACT

In the present study, the phylogenetic relationships within the Mysidae (Mysida, Crustacea) were inferred using 16S mtDNA and 18S and 28S rDNA. These genes were sequenced from 77 species within the family Mysidae, representing three of six subfamilies (Gastrosaccinae, Mysinae, Siriellinae) and four of seven tribes within the Mysinae (Erythropini, Heteromysini, Leptomysini, Mysini). The sequence data from all three genes were concatenated, aligned, and the ambiguous regions removed, for a final total dataset of 3048 bp. Phylogenies were reconstructed using maximum parsimony, maximum likelihood, and mixed model Bayesian methods coupled with Markov Chain Monte Carlo inference. Because the multiple ribosomal gene regions used as molecular markers are expected to contain variability in the pattern of evolution between stem and loop regions, the affect of partition choice on phylogeny is explored using newly developed Bayesian methods that detect pattern heterogeneity and assign sites to partitions during the analysis. These results were compared with the more common method of defining each gene as a partition *a priori*. The phylogenies revealed that most genera represented in our analyses and the Siriellinae are monophyletic ($BP \geq 92$; $pP \geq 0.99$). The Gastrosaccinae ($BP \geq 98$; $pP = 1.00$) species form a strongly supported clade but are paraphyletic due to the inclusion of species taxonomically placed within the Mysini. Although not monophyletic, the tribes Leptomysini and Mysini both form two well supported clades. The Mysinae is polyphyletic in all reconstructions, with

monophyly rejected by tree topology tests ($P < 0.001$). Investigations of pattern heterogeneity within these data indicate that increasing the estimated number of rate matrices decreases the log-likelihood scores; however, different numbers of matrices displayed different patterns of parameter mixing and convergence. Also, evolutionary model rate parameters that have reached stationarity appear to produce more resolved trees with better supported topologies.

INTRODUCTION

The order Mysida (i.e. mysid or opossum shrimp) contains over 1000 described species. This biodiversity is distributed across all latitudes throughout the waters of the world – subterranean, fresh, brackish, coastal, oceanic, and surface to deep sea (>7000m) (Wittmann 1999). Although the first mysidacean species was described in 1776 (Müller 1776), over two hundred years later there is still a large potential paucity in the taxonomic knowledge of this group, with extrapolations suggesting upwards of 4000 species worldwide (Wittmann 1999). With approximately three-fourths of the estimated species yet to be discovered, gaining an understanding of the evolutionary history of this group is crucial to taxonomic placement of newly described species and interpretations of new character states. As with many crustacean groups, however, the taxonomic scheme of the Mysida has a contentious history, and is still under debate. Originally mysids were grouped with the leptostracans, and later with the Euphausiacea in a taxon called the ‘Schizopoda’ (Schram 1986). Subsequently, mysids were moved to their own order, the Mysidacea, consisting of the suborders Lophogastrida and Mysida and moved to the Peracarida (Boas 1883). More recently, many crustacean researchers (Martin and Davis

2001; Schram 1986; Watling 1981; Watling 1983; Watling 1999) have raised both the Lophogastrida and Mysida from subordinal to ordinal rank and/or suggested phylogenetic affinities to members of the Eucarida (Jarman et al. 2000; Watling 1999). However, although the phylogenetic affiliations and taxonomic rank are still under debate, the monophyly of the Mysida is not contested.

The Mysida is composed of four families, with the largest proportion of the currently described species (and presumably those yet to be discovered) belonging to the family Mysidae, subfamily Mysinae (~990 and 750 species, respectively). Phylogenetic studies within Mysida are still in their infancy, and investigations below the ordinal level are scarce. Completed studies have either focused on the relationship between the Mysida and Lophogastrida and their placement within the Malacostraca (Jarman et al. 2000) or on the relationships among species within a single genus (i.e. *Pseudomma*, Meland 2004; Meland and Willassen 2004; *Mysis*, Väinölä 1986; Väinölä 1990; Väinölä 1992; Väinölä et al. 1994). The first phylogenetic study to investigate relationships among the biodiverse Mysidae assessed relationships among 25 species representing three of six subfamilies and two of seven tribes using 18S rDNA sequence data (Remerie et al. 2004). This study found significant incongruence between phylogenetic and taxonomic structure within the Mysidae. In particular the paraphyly of the species-rich Mysinae illustrates the need for taxonomic revision of the group (Figure 1).

The first goal of this study is to continue the phylogenetic investigations of the biodiverse Mysidae initiated by Remerie et al. (2004). Our study builds on this work by increasing both the number of species and molecular markers used for phylogenetic analyses. We refine the current phylogenetic hypotheses by using 16S mtDNA and 18S

and 28S rDNA to assess the relationships among 77 species of Mysida, including representatives of two additional Mysinae tribes (Heteromysini and Erythropini). Because we are analyzing multiple ribosomal gene regions, however, variability in the pattern of evolution between stem and loop regions is expected (Higgs 1998; Hillis and Dixon 1991; Savill et al. 2001; Schöniger and von Haeseler 1994). Recent advances within the Bayesian statistical framework have allowed phylogenetic analyses to incorporate heterogeneity across sites in the pattern of gene-sequence evolution by partitioning the data such that different models of evolution are assigned to different groups of sites. As currently implemented in one of the most commonly used Bayesian phylogenetic programs (MrBayes v3.04b, Ronquist and Huelsenbeck 2003), partitions in these ‘mixed model’ methods are chosen based on *a priori* information about how different portions of a sequence evolves, i.e. different gene markers, ribosomal stems and loops, or protein-coding first, second, and third positions; however, for some sequences *a priori* information on potential partitions may not be available, and pre-defined partitions may miss significant within-partition variability (Pagel and Meade 2004). In newly developed Bayesian mixture models, only the number of partitions is predefined and pattern heterogeneity is detected by summing the likelihood over different models, weighted by its probability (Pagel and Meade 2004). Moreover, this method simultaneously detects both rate and pattern-heterogeneity. As pattern-heterogeneity has previously been described from the ribosomal gene markers we are using, accurately modeling the pattern-heterogeneity is integral to phylogenetic reconstruction (Savill et al. 2001). Therefore the second goal of this research is to investigate the effect of defining different numbers of partitions on Mysidae phylogeny relative to more traditional

approaches (maximum parsimony, maximum likelihood, and *a priori* Bayesian mixed models). This study will be one of the first to use these methods in an organismal phylogenetics study.

METHODS

Taxon Sampling and Outgroup Choice

Sequences were collected from 76 species within the family Mysidae representing three of six subfamilies, four of seven tribes, and 34 genera (Table 1). All specimens were either collected by the authors using hand nets and light traps, or were provided by colleagues. Identifications were accomplished with the help of Mauchline (1980) and the Mysida Home Page (<http://tidepool.st.usm.edu/mysids/>) and references therein, or were provided by the collectors. Identifications made by the authors were verified either by comparison with previous genetic studies (Remerie et al. 2004) or by independent identification by others studying Mysida. Voucher specimens for species used in this study were deposited at the Monte L. Bean Museum, Brigham Young University. Relative to outgroup choice, the phylogenetic affinities of the Mysida are uncertain. Various studies have argued for either the monophyly (De Jong-Moreau and Casanova 2001; Richter and Scholtz 2001; Schram and Hof 1998; Wheeler et al. 2004; Wills 1998), paraphyly (Watling 1999), or polyphyly (Jarman et al. 2000) of the Mysida and Lophogastrida. Furthermore, Mysida affinities with the Peracarida (Richter and Scholtz 2001; Wills 1998) are also under debate, with a number of studies arguing for placing the Mysida basal to some portion of the Eucarida (Casanova et al. 2002; De Jong et al. 2004; Jarman et al. 2000; Kobusch 1998; Watling 1983; Watling 1999). Preliminary analyses

of a selection of potential outgroup taxa (Euphausiacea, Decapoda, Lophogastrida) indicate that levels of divergence make ingroup and outgroup sequence alignments difficult. Because the monophyly of the Mysida is generally uncontested, more appropriate outgroup taxa for the Mysidae would be representatives from the remaining families, Lepidomysidae, Petalophthalmidae, or Stygiomysidae; however these are all relatively small taxonomic groups (9, 34, and 6 species respectively) from restricted habitats (cave-adapted or deep-sea) that makes obtaining specimens difficult. Fortunately, a single representative of Stygiomysidae was provided to us by T. Iliffe; therefore, *Stygiomysis cokei* was used as the outgroup for all phylogenetic analyses.

DNA Extraction, PCR, and Sequencing

All specimens were stored in 70-95% ethanol and kept at 4°C until extracted. Genomic DNA was extracted using Qiagen DNeasy kits (Qiagen) following the manufacturer's instructions. After the initial lysis, the undigested exoskeletons were removed from the buffer solution, rinsed in water, dried in an ethanol series of increasing concentration, and mounted on glass slides in Canada balsalm for vouchers. Polymerase chain reaction (PCR, Saiki et al. 1988) products for the complete 18S rDNA (~2000 bp, Whiting 2002; Whiting et al. 1997) and 28S rDNA (~2500 bp, Whiting 2002; Whiting et al. 1997) nuclear genes, and the partial 16S (~460 bp, Crandall and Fitzpatrick Jr. 1996) mitochondrial gene were amplified using one or more sets of general primers from the literature. Standard PCR conditions (5 µl 10X *Taq* buffer, 6–8 µl 25mM MgCl₂, 8 µl 10mM dNTPs, 5 µl each of two 10 mM primers, 1.25 U *Taq*, ~20 µl double distilled water) were used on a Perkin-Elmer 9700 machine under the following conditions: an

initial denaturation at 96°C for 3 min followed by 40 cycles of 95°C for 1 min, 46°C for 1 min, and 72°C for 1min, followed by chain extension at 72°C for 10 min. PCR products were visualized by agarose (1.2%) gel electrophoresis and were purified using the Millipore Montage purification system. Sequences were generated in both directions on an ABI Prism 3730 capillary autosequencer using the ABI big-dye Ready-Reaction kit and following the standard cycle sequencing protocol but using 1/16th of the suggested reaction volume.

Model Selection

Model selection is an important issue in phylogenetic inference. Although the most common method of model justification is hierarchical likelihood ratio tests (Posada and Crandall 1998), this method of model selection may not be optimal due to limitations related to the nested model structure required for comparing hypotheses (Pol 2004; Posada and Buckley 2004). Other methods, including Akaike Information Criteria (AIC), have recently received a great deal of attention for the ability to evaluate both nested and non-nested sets of models, as well as allowing for model averaging procedures and assessments of model selection uncertainties (Adachi and Hasegawa 1996; Hasegawa 1990; Shimodaira 2001; Shimodaira and Hasegawa 1999b). In our analyses, model selection for maximum likelihood and Bayesian analyses followed the procedures outlined by Posada and Buckley (2004) for AIC. ModelTest v3.6 (Posada and Crandall 1998) was used to identify the model with the highest AIC weight for each gene individually and for the combined dataset.

Phylogenetic Analyses

Nucleotide sequences were aligned using Clustal X (Thompson et al. 1997) with the default parameters and refined by eye. Because many regions of the 16S, 18S, and 28S gene segments used for analysis are extremely divergent among the ingroup taxa and therefore difficult to align reliably, GBlocks v0.91b (Castresana 2000) was used to eliminate poorly aligned positions and divergent regions of the Clustal X alignment (GBlocks parameters used for 16S / 18S / 28S: minimum number of sequences for a conserved position = 36/51/43; minimum number of sequences for a flanking position = 45/65/60; Maximum number of contiguous nonconserved positions = 8/20/10; minimum length of a block = 5/5/10; allowed gap positions = with half).

Phylogenetic analyses of combined datasets have been shown to reveal hidden support for relationships in conflict among analyses of individual markers (Gatesy et al. 1999); therefore, the GBlocks-pruned datasets from each gene region were concatenated into a single combined dataset consisting of 3048 bp. The combined dataset was used to reconstruct phylogenies using equally weighted maximum parsimony (MP) and maximum likelihood (ML) heuristic searches in PAUP* v4b10 (Swofford 2002) and Bayesian methods coupled with Markov chain Monte Carlo inference as implemented in MrBayes v3.04b (BMCMC-MB, Ronquist and Huelsenbeck 2003). MP/ML analyses were run using 50,000/50 random addition (RA) replicates and tree bisection-reconnection (TBR) branch swapping. Confidence in the resulting relationships was assessed using the nonparametric bootstrap procedure (Felsenstein 1985) with 25,000/100 bootstrap replicates, TBR branch swapping, and 10/1 RA replicates used to calculate bootstrap proportions (BP). For BMCMC-MB analyses, three independent analyses were

run with each consisting of six chains. Each Markov chain was started from a random tree and run for 6.0×10^6 cycles, sampling every 1000th generation. Mixed models were used (Ronquist and Huelsenbeck 2003), with partitions designated as the three individual genes and models chosen as described above. Model parameters were treated as unknown variables with uniform default priors and were estimated as part of the analysis. To confirm that our Bayesian analyses converged and mixed well, we monitored the fluctuating value of likelihood and all phylogenetic parameters graphically, and compared means and variances of all likelihood parameters and likelihood scores from independent runs using the program Tracer v1.2 (Rambaut and Drummond 2003). All sample points prior to reaching stationary were discarded as burn-in. The posterior probabilities (pP) for individual clades obtained from separate analyses were compared for congruence and then combined and summarized on a majority-rule consensus tree (Huelsenbeck and Imennov 2002; Huelsenbeck et al. 2002).

Investigating dataset pattern heterogeneity

All of the sequence data we are using in this study are from nuclear and mitochondrial ribosomal genes, which are known to contain pattern-heterogeneity minimally in the stem versus loop structures (Higgs 1998; Hillis and Dixon 1991; Savill et al. 2001; Schöniger and von Haeseler 1994). In order to assess the effect of pattern-heterogeneity on mysid phylogeny given the ribosomal-based sequence data, we use Pagel and Meade's (2004) procedures as implemented in the program BayesPhylogenies (BMCMBP). In BMCMBP, Pagel and Meade (2004) implement a 'mixture model' for detecting pattern heterogeneity across sites by summing the likelihood over different models with varying

probabilities to the same site in the alignment, each weighted by its probability as estimated from the data. To determine the optimal number of partitions, a dataset of interest is analyzed under a Bayesian framework for sequentially increasing numbers of partitions. For each analysis, the average likelihood and the average standard deviation of the rate parameters are assessed graphically. This information is combined with the calculation of Bayes factors (Nylander et al. 2004; Pagel and Meade 2004), which penalizes more complex models, to determine the number of partitions that best represents the pattern-heterogeneity contained in the dataset. Sequential analyses were run by increasing the number of estimated rate matrices (Q). Each partition analysis was run once using four chains for 3.0×10^6 generations, sampling every 1000th generation with a gamma rate distribution of four categories. Each analysis was treated as above to determine convergence, mixing, burn-in, and clade pP . After burn-in, average likelihood scores and rate parameter standard deviations of each analysis were calculated and examined graphically, and Bayes factors (interpreted as in Kass and Raftery 1995) compared for each sequential model comparison. The effect of different numbers of partitions on the resulting topology was also assessed using tree statistics measuring the percentage of well-supported branches (# branches with $pP \geq 0.95$ / total number of branches in a rooted bifurcating tree), and resolution (Colless 1980) as calculated in the program RadCon (Thorley and Page 2000). The topology based on the optimal number of partitions (with partition number indicated by a number at the end of BMCMC-BP) was also compared to the previously described ML, BMCMC-MB, and MP analyses using tree topology tests (described below).

Testing Alternative Hypotheses

We tested all of our phylogenetic hypotheses against the only previously published phylogenetic hypothesis of mysid relationships (Remerie et al. 2004). Because there are differences between the taxon sampling of the *a priori* hypothesis and our dataset, the alternative topology was constructed in MacClade (Maddison and Maddison 2003) by rearranging only the branches representing the lineages in conflict. Additionally, we tested the monophyly of any taxonomic unit (subfamily, tribe, and genus) that was reconstructed as non-monophyletic in our analyses. Likelihood topology tests were conducted using our molecular data and the Shimodaira and Hasegawa (1999b, SH) test as implemented in PAUP*. Goldman et al. (2000), Buckley (2002), and Strimmer and Rambaut (2002) have pointed out that the SH test may be subject to a certain type of bias such that the number of trees included in the confidence set tends to be very large as the number of trees to be compared increases, which makes the test conservative. However, as these authors recognized and Shimodaira (2002) concluded, the SH test is still safe to use and is a good option when the number of candidate trees is not very large and more data are accumulated, as is the case with our study. Ten thousand replicates were performed for every topology test resampling the partial likelihoods for each site (RELL model).

RESULTS

Mysid Phylogenetics

We obtained 77 new complete and partial 18S and 28S, and 66 partial 16S gene sequences (Table 1). Of particular importance is the fact that 12 of the 76 species we

sampled were new species. Many of these new species are the result of intensive sampling of a single region by a collaborator, which emphasizes the potentially large number of undescribed species within the Mysidae given thorough sampling. AIC model selection for each gene individually and the concatenated dataset resulted in very similar models, with variation in parameter estimates (Table 2). Of the 20 genera represented by more than one species, 17 were monophyletic with high branch supports ($\text{BP} > 92$; $\text{p}P > 0.99$) (Figure 2), and three were not recovered as monophyletic in at least one phylogeny (*Gastrosaccus*, *Mysidopsis*, and *Neomysis*). Topology tests of these three genera show that enforcing either *Mysidopsis* or *Neomysis* monophyly produced significantly worse trees ($P < 0.001$). Trees with a monophyletic *Gastrosaccus* clade were not significantly worse than the ML tree ($P = 0.20$; Table 3) and in fact this genus was recovered as monophyletic in MP trees.

In comparison to the only previous ML phylogenetic hypothesis of Mysidae relationships, our ML tree is a significantly better topology using the SH test ($P < 0.01$). Among the analyses we performed the largest differences between tree reconstruction methods are observed between the MP tree and the explicitly model-based ML and BMCMC methods (Figures 2, 3). This is reflected by likelihood topology tests where ML and BMCMC produce trees that are not significantly different, but ML and MP trees are borderline significantly different ($P < 0.04$). Although there were topological differences, there was general agreement in the relative arrangement of the major taxonomic clades among all reconstruction methods employed. Within the family Mysidae, there are six subfamilies, three of which are represented in our analyses. One of these subfamilies, Siriellinae, is monophyletic with high branch support ($\text{BP} = 100$; $\text{p}P$

=1.0). The Gastrosaccinae are paraphyletic, but only because of the inclusion of a single Mysinae taxa, *Mesopodopsis africana*, in the well supported clade (BP \geq 98; pP = 1.0) and removing *M. africana* from the Gastrosaccinae does not produce significantly worse trees ($P = 0.36$). Consisting of seven tribes, four of which are represented here, the third subfamily Mysinae is polyphyletic. The majority of the Mysinae species included in these analyses, representing the tribes Mysini, Leptomysini, and Heteromysini, form a single clade that is weakly supported in all analyses (BP<50; pP>0.77). However, in the explicitly model-based methods (ML, BMCMC-MB, BMCMC-BP.6Q) the Erythropini are placed as the most basal lineage within the Mysidae. In contrast, in the MP phylogeny the Erythropini are placed sister to the rest of the Mysinae (Heteromysini+Leptomysini+Mysini), with two Mysini genera (*Antromysis* and *Mysidium*) placed basal to that group. In all analyses, the taxon *Australerythrops* n. sp. is not recovered with the remaining Erythropini species, thereby making the Erythropini polyphyletic. The instability of this taxon, in particular, may be due to the ~1100 bp of missing sequence data. Monophyletic arrangements of these clades, however, were significantly worse topologies (Mysinae $P < 0.001$; Erythropini $P < 0.001$)

Within the clade Mysinae, all analyses recover two monophyletic groups, corresponding to the Mysini-A and Mysini-B clades of Remerie et al. (Remerie et al. 2004). The Leptomysini were also recovered in two well supported clades, here referred to as Leptomysini-A and Leptomysini-B, with Leptomysini-A consisting of the genera *Americamysis*, *Dioptromysis*, *Leptomysis*, *Metamysidopsis*, and *Mysidopsis* and Leptomysini-B of *Prionomysis*. However, in all analyses, the Mysini species *Stilomysis grandis* and *Kainomatomysis* n. sp. fall within and the genus *Antromysis* is placed basal

to the Leptomysini-A clade. These stray Mysini species cause the well supported Leptomysini-A clade ($\text{BP}=100$; $\text{p}P=1.0$) to be paraphyletic, although a monophyletic Leptomysini-A is not a significantly worse topology ($P = 0.06$). Additionally, the relative arrangement of the two Mysini and the two Leptomysini clades varies between methods (Figures 2, 3). Monophyly was rejected for both Mysini and Leptomysini by likelihood topology tests ($P < 0.001$ for both tests).

BMCMB a priori vs. pattern heterogeneity mixture models

In order to investigate the pattern heterogeneity in our dataset, phylogenetic analyses defining *a priori* data partitions as implemented in BMCMB-MB mixed models were compared to likelihood-summed pattern heterogeneity mixture models from BMCMB-BP. Model comparisons between BMCMB-MB and BMCMB-BP three partition mixture models using Bayes factors show strong support for the BMCMB-BP model, implying that there was significant pattern variability within our pre-defined gene partitions. Analyses of sequentially increasing partition numbers in BMCMB-BP further illustrate this point, showing that log-likelihood score continues to improve up to six rate matrices, with the Bayes factors suggesting strong evidence for the more complex model at each step (Figure 4A, Table 4). However, overall branch support and resolution do not follow this same pattern, showing that the BMCMB-BP 3Q and 5Q analyses are the best topologies using these measures having similar values to the BMCMB-MB tree (% branch support = 78.7; resolution = 97.3). It is also clear from these analyses that the number of defined rate matrices characterizing the pattern heterogeneity affects the inferred relationships and their support. With increasing number of partitions, some

relationships disintegrated, such as the sister relationship of the Siriellinae and the Gastrosaccinae (BMCMC-BP.1Q), while the support of other clades (e.g. Mysini-A), varied ($pP > 0.95$ for 2Q and 3Q; $pP < 0.95$ with 1Q, 4Q, and 5Q; Figure 4B). Not surprisingly, the ML analysis resulted in a tree topology that was most similar to the BMCMC-BP single partition phylogeny.

Simulation studies have shown that when sufficient rate matrices have been estimated, adding matrices will lead to poorly estimated parameters and small weights for the excess matrices (Pagel and Meade 2004). Although the improvement in log likelihood score has started to slow by the 6Q analysis, the calculated Bayes Factors indicate that six rate matrices are still insufficient for the pattern heterogeneity contained in our three gene concatenated dataset, as adding rate matrices still significantly improves the log-likelihood scores. Interestingly, the increase of rate parameter standard deviation mimics the overall slow-down in log-likelihood improvement as the number of rate matrices increases, up to the 4Q analysis. However, in the 5Q analysis the standard deviation of the rate parameters decreases followed by a drastic increase in the 6Q analysis, independent of the likelihood score improvement (Figure 4A), indicating that the rate parameters have been poorly estimated in the 6Q analysis. Investigations of the convergence and mixing of the rate parameters show that while the 1Q parameters all converge rapidly (within the first few generations), the 2Q and 4Q parameter estimates exhibit rampant fluctuations, explaining the increase in rate parameter standard deviation of these analyses. Although the 3Q rate parameter standard deviations are higher than those calculated from the 2Q analysis, the estimates generally reach stationary by the burn-in used to calculate the phylogeny (1×10^6 generations). The 6Q tree has the lowest

branch support and resolution of any of the inferred topologies (Figure 4) and the highest rate parameter standard deviation. This is due to rate parameters not reaching stationary until near the end of the analysis ($\sim 2.5 \times 10^6$ generations; Figure 5), implying that more generations (and longer burn-ins) are required to properly sample the distribution. Interestingly, the rate parameter stability over the generations sampled to create the phylogeny is roughly correlated with the tree support and resolution, with the 3Q and 5Q models containing the best estimated parameters, and the highest tree support and resolution among all the partition analyses. All of these pieces of information imply that although continuing to add rate matrices to the model may still improve the likelihood score and thereby more accurately model the pattern heterogeneity in our dataset, it is at the expense of the rate parameter estimations. Much longer chain generations will be required to reach stable parameter distributions. Therefore, from the analyses that we completed, the 5Q pattern heterogeneity model is preferred as the best combination of log-likelihood score and parameter estimates.

DISCUSSION

Mysid Phylogenetics

Overall, although the different reconstruction methods vary in topology, there is agreement in the relative placement of the major clades. The Gastrosaccinae and Siriellinae and the Mysinae tribe Erythropini are more basal lineages within the Mysidae, while the Mysinae tribes Mysini, Leptomysini, and Heteromysini are more derived. The basic arrangement of taxonomic units in these analyses, in particular of the Mysinae tribes, are more congruent with taxonomy than the previous ML hypothesis (Remerie et

al. 2004) and likelihood tests suggest that our model-based trees, based on increased taxa and gene sampling, are significantly better topologies. Nevertheless, our results do agree with previous findings in recovering two main Mysini clades. Remerie et al. (2004) defined these clades as ‘Mysini-A’ including genera *Diamysis*, *Hemimysis*, *Limnomysis*, *Paramesopodopsis*, *Praunus*, and *Schistomysis*, and ‘Mysini-B’ composed of genera *Acanthomysis*, *Holemesimysis*, and *Neomysis*. Our analyses extend these two main groups by adding the genera *Anisomysis*, *Idiomysis*, *Katamysis*, *Mysidium*, *Mysis*, and *Taphromysis* to Mysini-A (Figure 2, 3). However, the phylogenetic arrangement of these two clades relative to each other is still uncertain; model based methods place them as sister forming a main ‘Mysini’ clade while the MP analysis is similar to the Remerie et al. (2004) arrangement with Mysini-B sister to the Leptomysini. The only higher taxonomic groups that are monophyletic are the subfamily Siriellinae and the tribe Heteromysini; however, likelihood topology tests suggest that the monophyly of the Gastrosaccinae is equivalent to the ML hypothesis of non-monophyly. Even though Gastrosaccinae and the Mysinae tribes Erythropini, Leptomysini, and Mysini are not monophyletic, the phylogenies show that the species represented here follow this basic taxonomic scheme, with only a few taxonomically misplaced species causing the discrepancies (Figure 2, 3). While higher level taxonomy appears to be incongruent, taxonomy at the genus-level appears to be robust, with 17 of the 20 multi-represented genera monophyletic. However, the poor support observed at deeper branches in our analyses may also reflect a lack signal in our data for higher-level relationships. Collecting additional genetic data will be essential to resolving the taxonomic / phylogenetic incongruencies observed in this study.

Our analyses further support those of Remerie et al. (2004) in suggesting that the Mysidae are sorely in need of taxonomic revision. In particular, both studies suggest that Mysinae is the most incongruent with taxonomy, with the polyphyletic arrangement of this group necessitating major revisions. However, Mysinae is the most species-rich subfamily in the Mysidae (~750 of ~990 total species), and taxonomic revision is not a trivial matter. In this case, our phylogenetic analyses provide a framework for where initial efforts should be focused. For example, in the Mysini-B clade the genus *Neomysis* is polyphyletic as are the *Mysidopsis* species in the Leptomysini clade. However, as pointed out by Remerie et al. (2004), in each of these clades all of the genera were formerly described as a single taxonomic unit. Within the well-supported Mysini-B clade ($BP = 100$; $pP = 1.00$) all of the species were originally described as *Neomysis*, and were subsequently moved to new genera by various taxonomists (i.e. *Holmesimysis*, *Alienacanthomysis*, Holmquist 1981). Similarly, in the ‘*Mysidopsis*’ clade ($BP = 98$; $pP = 1.00$) the intercolated genera *Americamysis* and *Metamysidopsis* were formerly contained within the genus *Mysidopsis*. Both of these clades represent species groups where the phylogeny strongly supports the monophyly of the group, but generic taxonomy is not yet well defined. Similarly, the consistent placement of *Antromysis* spp., *Kainaomatymysis* n. sp., *Mesopodopsis africana*, and *Stilomysis grandis* outside of the two main Mysini clades pinpoint additional species in need of taxonomic scrutiny.

Pattern Heterogeneity

Evolutionary models for multigene data sets can be improved considerably by recognizing across-partition heterogeneity in model parameters (Nylander et al. 2004;

Pagel and Meade 2004). However, even when across-partition heterogeneity is significant, other model components seem to be even more important, particularly those that deal with within-partition rate variation (Nylander et al. 2004; Yang 1996). The program BayesPhylogenies can deal with both rate- and pattern heterogeneity without assigning sites to specific partitions. Particularly when there is no clear case for partitioning the data, summing over models may allow for unforeseen patterns of evolution to emerge. This was the assumption for our dataset, which consisted of three concatenated ribosomal genes where qualitative differences in the pattern of evolution across sites (i.e. stem vs. loop regions) have been previously documented (Higgs 1998; Hillis and Dixon 1991; Savill et al. 2001; Schöniger and von Haeseler 1994). While designation of stem and loop partitions may seem reasonable, Pagel and Meade (2004) demonstrated that this approach can miss significant within-partition variability and our treatment of these data – concatenation and removal of ambiguously aligned regions – complicate the *a priori* determination of site partitions based on secondary structure. Furthermore, there was strong evidence based on Bayes factors that the pattern-heterogeneity using three rate matrices was a better model for our data than *a priori* designation of each gene as a partition (Table 4). Increasing the number of rate matrices continued to improve the log-likelihood score beyond what was expected by the addition of the extra parameters. In our analyses, we investigated up to 6Q; the trend in log-likelihood score indicates that additional rate matrices would continue to improve the model. However, examination of the rate parameters indicates that at 6Q the parameters are poorly estimated, as evidenced by the large jump in standard deviation (Figure 4A). These two pieces of information taken together indicate that although additional rate

matrices would produce better log-likelihood scores, the 5Q phylogeny is the best combination of likelihood score and parameter estimates.

Comparisons of phylogenies generated using different numbers of rate matrices show topological difference. Although the major clades remain stable, the relationships among them change. For example, depending on the number of rate matrices the Leptomysini-B clade is sister to either Heteromysini (1Q, 3Q, 6Q) or Leptomysini-A (2Q, 4Q). With respect to Mysidae phylogeny, the ambiguity of relationships (Mysini-A, Mysini-B, Leptomysini-A, Leptomysini-B, Heteromysini) within the Mysinae is illustrated by the different arrangement of clades in every analysis. Previous studies have shown that more complex models are associated with more topological uncertainty (Nylander et al. 2004). In our dataset patterns of topological uncertainty measured by percent resolution are correlated with the degree of stability, mixing, and convergence observed in the rate parameter estimates. Analyses estimating 2Q and 4Q contained less resolution (93.3 and 94.7%, respectively) than analyses reaching stationary (1Q – 96% and 3Q – 97.3%). Although the 6Q estimates reached stationary, it was at the end of the analysis resulting in a lower resolution. All of these pieces of information imply that although continuing to add rate matrices to the model may still improve the likelihood score and thereby more accurately model the pattern heterogeneity in our dataset, it is at the expense of the rate parameter estimations. Much longer chain generations will be required to reach stable parameter distributions.

CONCLUSIONS

The Mysidae are important components of coastal and estuarine ecosystems, where seasonal densities can constitute huge biomasses (upwards of 50,000 individuals / m³ of water, Wittmann 1999). Due to their ubiquitousness in coastal areas, mysid species have also been increasingly used as bioindicators in ecotoxicology studies (Verslycke et al. 2004). Yet, given their importance in ecological and toxicological studies the Mysidae have been mostly overlooked in phylogenetic studies. Our study is only the second molecular phylogenetic study within the biodiverse Mysidae. The results of our phylogenetic investigations of Mysidae indicate that, with the exception of a few taxonomically misplaced species, the Gastrosaccinae, Siriellinae, and Heteromysini form strongly supported clades while the Mysinae, Mysini, Leptomysini, and Erythropini are polyphyletic. Given that our sampling from within these taxonomic groups is small relative to total genus/species numbers, the monophyly of all of these clades will continue to be tested as additional taxa become available for analyses. Furthermore, the addition of representatives from the missing subfamilies (Boreomysinae, 38 spp.; Rhopalophthalminae, 18 spp; Mysidellinae, 16 spp.) and tribes (Aberomysini, 1 sp.; Mancomysini, 4 spp.; Calyptommini, 3 spp.) will be necessary to generate a complete phylogenetic understanding of the family. Finally, for a group of crustaceans first described in 1776 (Müller 1776), the percentage of new species in our study (~15%) support the hypothesis that a large number of Mysidae species remain to be discovered and described. These phylogenetic studies serve as a foundation for these descriptive efforts by providing a framework for hypotheses of generic relationships and character evolution, and by highlighting areas of Mysidae taxonomy in particular need of revision.

Using a three ribosomal gene dataset from the Mysidae as a test case, investigations of pattern heterogeneity confirm Pagel and Meade's (2004) assertion that mixed models defined *a priori* miss significant within partition variability and that improved phylogenies can be obtained by accounting for both rate- and pattern-heterogeneity. As the size of genetic datasets is rapidly increasing in terms of the number of independent markers employed and total sequence lengths, investigations of pattern heterogeneity will be needed to address the issues of how increasing sequence complexity affect rate parameter estimation in terms of convergence and mixing and the number of generations necessary to reach stationarity. Moreover, there is a limit to the number of parameters that can be successfully included in a model, and preliminary observations have demonstrated this with a model containing 12 partitions and 121 substitution parameters (Nylander et al. 2004). Given that a dataset consisting of only three genes is not fully characterized by a model estimating 6Q, the ability to model complex data is a very real concern. Further investigations of the costs of estimating additional parameters versus the gain in likelihood score will be of interest to examining pattern heterogeneity in the large and complex datasets that are becoming the standard of phylogenetic studies.

Table 3-1. Taxonomy for gene sequences from Mysidae species included in this study. ‘X’ indicate obtained sequences while N/A designates gene sequences we were unable to acquire, ‘†’ denotes species missing more than 400 bp of the final alignment due to incomplete sequence data, ‘‡’ indicates species described in(Fenton 1985), and * species new species collected and being described by A. Connell.

Taxa	16S	18S	28S
Mysidae			
Gastrosaccinae			
<i>Archaeomysis grebnitzkii</i> Czerniavsky, 1882	X	X	X
<i>Bowmaniella brasiliensis</i> Bacescu, 1968	X	X	X
<i>Bowmaniella floridana</i> Holmquist, 1975	X	X	X
<i>Eubowmaniella simulans*</i> (W. Tattersall, 1915)	N/A	X	X
<i>Gastrosaccus brevifissura</i> O. Tattersall, 1952	X	X	X
<i>Gastrosaccus bispinosa</i> Wooldridge, 1978	X	X	X
<i>Haplostylus</i> sp.1	X	X	X
<i>Haplostylus</i> sp. 2	X	X	X
Siriellinae			
<i>Siriella</i> n. sp.1*	X	X	X
<i>Siriella</i> n. sp.2 *	N/A	X	X
<i>Siriella</i> n. sp.3*	X	X	X
<i>Siriella jaltensis</i> Czerniavsky, 1868	X	X	X
<i>Siriella jaltensis</i> Czerniavsky, 1868	X	X	X
<i>Siriella</i> sp.3	X	X	X
<i>Siriella</i> sp.1	X	X	X
<i>Siriella armata</i> (Milne-Edwards 1837)	X	X	X
<i>Siriella</i> sp.2	X	X	X
<i>Siriella chierchiae</i> Coifmann, 1973	X	X	X
Mysinae			
Erythropini			
<i>Amathimysis gibba</i> Brattegard, 1969	X	X	X
<i>Amathimysis trigibba</i> Murano and Chess, 1987	X	X	X
<i>Australerythrops</i> n. sp.*†	X	X	X
<i>Pleurerythrops</i> n. sp.1*	X	X	X
<i>Pleurerythrops</i> n. sp.2*	X	X	X
Heteromysini			
<i>Heteromysis norvegica</i> G.O. Sars, 1883	X	X	X
<i>Heteromysis</i> sp.1	X	X	X
Leptomysini			
<i>Americamysis almyra</i> (Bowman, 1964)	N/A	X	X
<i>Americamysis bahia</i> (Molenock, 1969)	X	X	X
<i>Dioptromyssis spinosa</i> Brattegard, 1969	X	X	X
<i>Leptomyssis lingvura lingvura</i> (Sars, 1866)	X	X	X
<i>Leptomyssis</i> n. sp. aff. <i>Heterophila</i> *	N/A	X	X
<i>Leptomyssis lingvura lingvura</i> (Sars, 1866)	X	X	X
<i>Leptomyssis lingvura adriatica</i> (Sars, 1866)	X	X	X
<i>Metamysidopsis munda</i> (Zimmer, 1918)	N/A	X	X
<i>Metamysidopsis neritica</i> Bond-Buckup and Tavares, 1992	X	X	X
<i>Metamysidopsis elongata</i> (Holmes, 1900)	X	X	X
<i>Metamysidopsis swiftii</i> Bacescu, 1969	X	X	X
<i>Mysidopsis eclipse</i> Brattegard, 1969	X	X	X
<i>Mysidopsis furca</i> Bowman, 1957	X	X	X

<i>Mysidopsis</i> n. sp.3	X	X	X
<i>Prionomysis</i> n. sp.1 [‡]	X	X	X
<i>Prionomysis</i> n. sp. 1 [‡]	X	X	X
<i>Prionomysis</i> sp. M139.5.1	X	X	X
Mysini			
<i>Alienacanthomysis macropsis</i> (W. Tattersall, 1932)	X	X	X
<i>Anisomysis marisrubri</i> Bacescu, 1973	X	X	X
<i>Anisomysis</i> n. sp.* [†]	X	X	X
<i>Antromysis cenotensis</i> Creaser, 1936	X	X	X
<i>Antromysis cubanica</i> Bacescu and Orghidan, 1971	N/A	X	X
<i>Antromysis</i> n. sp.*	X	X	X
<i>Hemimysis abyssicola</i> Sars, 1869	X	X	X
<i>Hemimysis maderensis</i> Ledoyer, 1989	N/A	X	X
<i>Holmesimysis costata</i> (Holmes, 1900)	X	X	X
<i>Holmesimysis sculpta</i> (W. Tattersall, 1933)	X	X	X
<i>Idiomysis zuluensis</i>	X	X	X
<i>Kainomatomysis</i> n. sp.*	X	X	X
<i>Katamysis warpachowskyi</i> Sars, 1893	X	X	X
<i>Limnomysis benedeni</i> Czerniavsky, 1882	X	X	X
<i>Mesopodopsis africana</i> * O. Tattersall, 1952	X	X	X
<i>Mysidium gracile</i> * (Dana, 1852)	N/A	X	X
<i>Mysidium</i> sp.1 [†]	N/A	X	X
<i>Mysidium columbiae</i> [†]	X	X	X
<i>Mysis oculata</i> (Fabricius, 1780)	X	X	X
<i>Mysis relicta</i> Lovén, 1862	X	X	X
<i>Mysis stenolepis</i> Smith, 1873	X	X	X
<i>Neomysis americana</i> (S.I. Smith, 1873)	X	X	X
<i>Neomysis integer</i> (Leach, 1815)	N/A	X	X
<i>Neomysis mercedis</i> Holmes, 1897	X	X	X
<i>Neomysis patagona</i> Zimmer, 1907	X	X	X
<i>Neomysis rayii</i> (Murdoch, 1885)	X	X	X
<i>Neomysis</i> sp.1	X	X	X
<i>Paramesopodopsis rufa</i> Fenton, 1985	X	X	X
<i>Praunus flexuosus</i> (Müller, 1776)	X	X	X
<i>Schistomysis spiritus</i> (Norman, 1860)	X	X	X
<i>Schistomysis kervillei</i> (Sars, 1885)	X	X	X
<i>Stilomysis grandis</i> (Goës, 1863)	X	X	X
<i>Taphromysis bowmani</i> Bacescu, 1961	X	X	X
<i>Taphromysis louisianae</i> Banner, 1953	X	X	X
OUTGROUPS			
Stygiomysidae			
<i>Stygiomysis cokei</i> Kallmeyer and Carpenter, 1996	X	X	X

Table 3-2. Parameter estimates for the best-fit models of nucleotide substitution. π_A , π_C , π_G , and π_T are the empirical base frequencies; r_{AC} , r_{AG} , r_{AT} , r_{CG} , and r_{CT} are the relative substitution rates among nucleotides ($r_{GT} = 1$); I is the proportion of invariable sites; and α is the shape parameter of the gamma distribution for the variation among sites.

Gene Partition	Model	π_A	π_C	π_G	π_T	r_{AC}	r_{AG}	r_{AT}	r_{CG}	r_{CT}	α	I
16S	GTR+I+ Γ	0.370	0.122	0.164	0.344	1.502	5.231	1.935	0.836	7.562	0.618	0.189
18S	GTR+I+ Γ	0.251	0.209	0.271	0.269	1.359	2.496	1.625	0.656	4.802	0.413	0.264
28S	GTR+ Γ	0.234	0.238	0.287	0.241	1.066	2.199	1.949	0.622	4.170	0.801	---
All genes	GTR+I+ Γ	0.260	0.205	0.266	0.269	1.258	2.795	2.093	0.712	4.816	0.565	0.191

Table 3-3. Likelihood topology tests of hypotheses regarding Mysidae phylogeny, including those topologies generated in this study under different methods (ML, MP, BMCMC), the previous hypothesis as shown in Figure 1, and for tests of monophyly where taxonomy is incongruent with the recovered molecular phylogenies. The differences in likelihoods ($\Delta\text{-lnL}$) and the corresponding P values are indicated. All comparisons are single tree comparisons, except for those with the MP trees, which include all eight most parsimonious topologies.

Topology	$\Delta\text{-lnL}$	P
Methods Comparisons		
ML	---	---
BMCMC-MB	4.73	0.45
BMCMC-BP.6	7.36	0.57
8 MPT	46.4-81.6	<0.01-0.04
Previous Hypothesis		
Figure 1	55.8	<0.01
Clade Monophyly		
Gastrosaccinae	53.6	0.36
<i>Gastrosaccus</i>	1.86	0.20
Mysinae	264.2	<0.001
<i>Erythropini</i>	463	<0.001
<i>Leptomysini</i>	128.7	0.06
<i>Mysidopsis</i>	146.6	<0.001
<i>Mysini</i>	435.2	<0.001
<i>Neomysis</i>	1026.9	<0.001

Table 3-4. Model comparisons of BMCMC-BP analyses with increasing numbers of estimated rate matrices. Bayes factors are calculated for each comparison.

Model comparison (M ₁ /M ₀)	Model likelihood		Bayes Factor	
	log _e f(X M ₁)	log _e f(X M ₀)	log _e B ₁₀	2log _e B ₁₀
BMCMC-BP.2 / BMCMC-BP.1	-54,540	-55,256	716	1432
BMCMC-BP.3 / BMCMC-BP.2	-54,219	-54,540	321	642
BMCMC-BP.4 / BMCMC-BP.3	-54,075	-54,219	144	288
BMCMC-BP.5 / BMCMC-BP.4	-53,961	-54,075	114	228
BMCMC-BP.6 / BMCMC-BP.5	-53,897	-53,961	64	128
BMCMC-BP.3 / BMCMC-MB	-54,219	-54,581	362	724

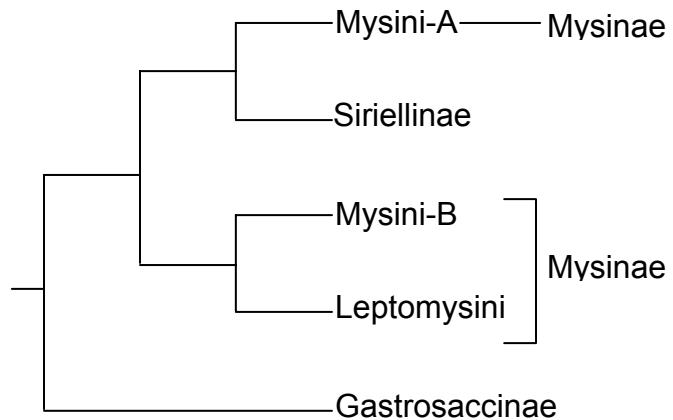


Figure 3-1. Previous maximum likelihood phylogenetic hypothesis of Mysidae relationships based on 18S rDNA from Remerie et al. (2004). The Mysini-A lineage includes the genera *Diamysis*, *Hemimysis*, *Limnomysis*, *Paramesopodopsis*, *Praunus*, and *Schistomysis*, while the Mysini-B lineage contains *Acanthomysis*, *Holmesimysis*, and *Neomysis*.

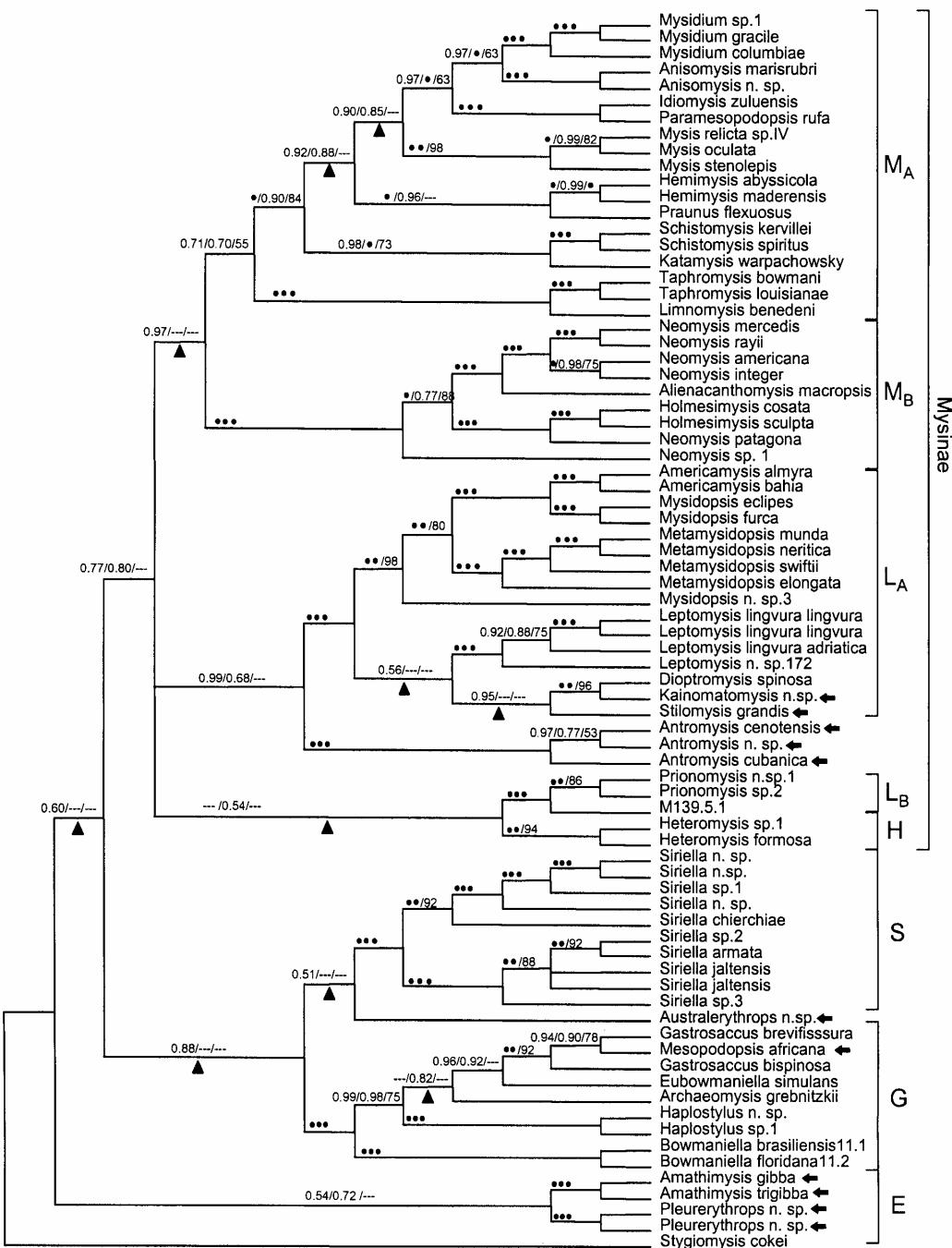


Figure 3-2. 50% majority rule consensus cladogram of the BMCMC-MB, BMCMC.6Q, and the ML explicit model based phylogenies. Triangles under branches indicate 67% tree consensus while branches with nothing indicate 100% consensus. The support values are indicated above each branch for BMCMC-MB pP / BMCMC-BP.6Q pP / ML BP. A circle in place of a support value indicates either pP = 1.00 or BP = 100. Arrows to the right of taxon names indicate species that are taxonomically incongruent with the reconstructed phylogeny. Major taxonomic groups indicated by: E – Erythropini; G – Gastrosaccinae; H – Heteromysini; L_A – Leptomysini-A; L_B – Leptomysini-B; M_A – Mysini-A; M_B – Mysini-B; S – Siriellinae.

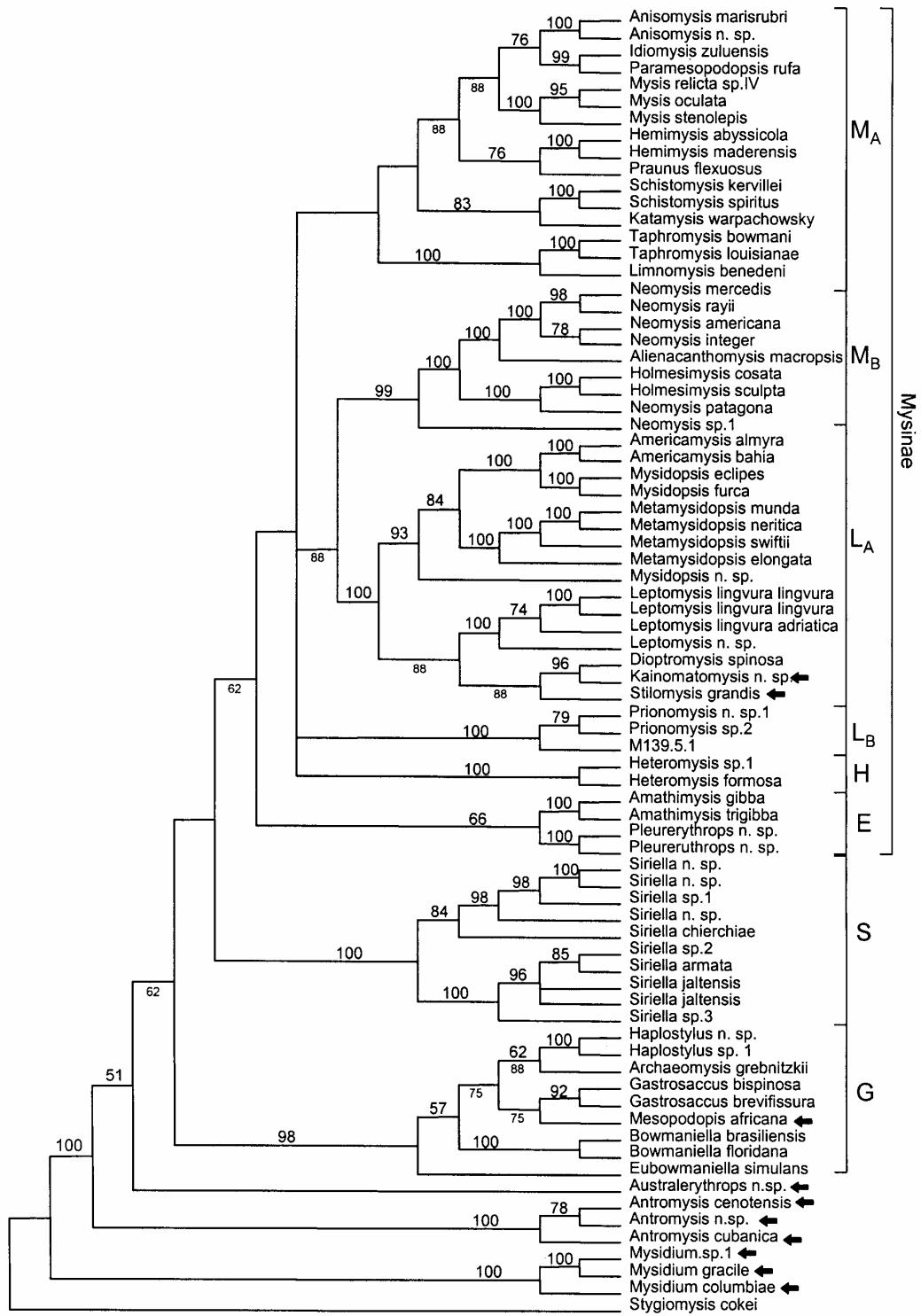
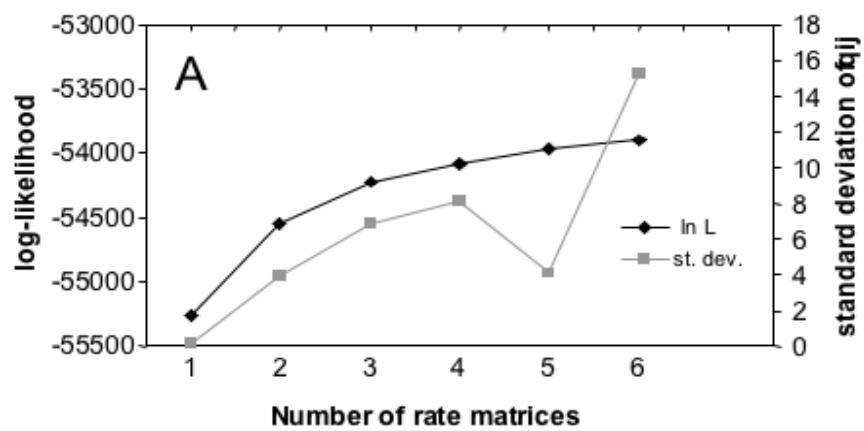


Figure 3-3. 50% majority rule consensus cladogram of eight most parsimonious trees (score=11640). Numbers below branches indicate where tree consensus was less than 100%. BP support values are placed above each branch. Arrows to the right of taxon names indicate species that are taxonomically incongruent with the reconstructed phylogeny. Major taxonomic groups indicated by: E – Erythropini; G – Gastrosaccinae; H – Heteromysini; L_A – Leptomysini-A; L_B – Leptomysini-B; M_A – Mysini-A; M_B – Mysini-B; S – Siriellinae.



B



Figure 3-4. BMCMC-BP analyses of sequentially increasing the number of estimated rate matrices. A) Graph of log-likelihood score and average standard deviation of the rate parameters for each analysis. B) Cladograms produced from each analysis. Dotted lines indicate branches where $pP < 0.95$. The overall branch support (BS) and resolution (RS) are indicated above each topology. Clades are labeled as follows: E – Erythropini; G – Gastrosaccinae; H – Heteromysini; L_A – Leptomysini-A; L_B – Leptomysini-B; M_A – Mysini-A; M_B – Mysini-B; S – Siriellinae.

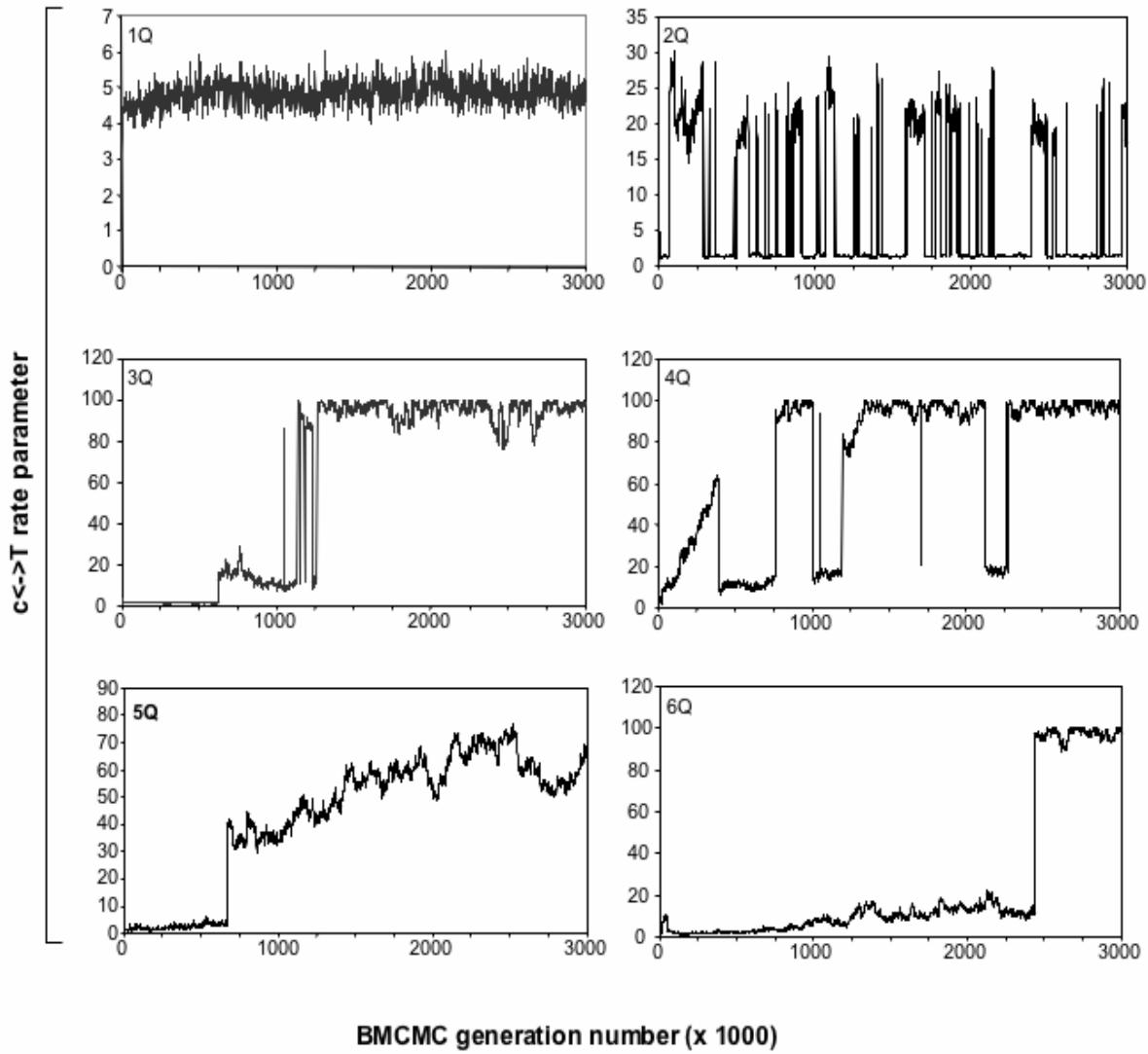


Figure 3-5. For each rate matrix analysis, a rate parameter was plotted against generation number to exemplify the observed pattern of convergence. In most cases, all estimated rate parameters in a particular analysis resembled the patterns shown here. The number of rate matrices for each analysis is indicated in each panel.

CHAPTER 4.

MODEL BASED MULTI-LOCUS ESTIMATION OF DECAPOD

PHYLOGENY AND DIVERGENCE TIMES

ABSTRACT¹

Phylogenetic relationships among the major decapod infraorders has never been estimated using molecular data, while morphological studies produce conflicting results. In the present study, the phylogenetic relationships among the decapod basal suborder Dendrobranchiata and all of the currently recognized decapod infraorders within the suborder Pleocyemata (Caridea, Stenopodidea, Achelata, Astacidea, Thalassinidea, Anomala, and Brachyura) were inferred using 16S mtDNA, 18S and 28S rDNA, and the histone H3 gene. Phylogenies were reconstructed using the model-based methods of maximum likelihood and Bayesian methods coupled with Markov Chain Monte Carlo inference. The phylogenies revealed that the seven infraorders are monophyletic, with high clade support values ($bp > 70$; $pP > 0.95$) under both methods. The two suborders also were recovered as monophyletic, but with weaker support ($bp = 70$; $pP = 0.74$). Although the nodal support values for infraordinal relationships were low ($bp < 50$; $pP < 0.77$), the Anomala and Brachyura were basal to the rest of the ‘Reptantia’ in both reconstructions, and Bayesian tree topology tests rejected alternate morphology-based hypotheses ($P < 0.05$). Newly developed multi-locus Bayesian and likelihood heuristic

¹ This chapter was submitted to *Molecular Phylogeny and Evolution* as: Porter, M.L., M. Pérez-Losada, and K.A. Crandall. Model based multi-locus estimation of decapod phylogeny and divergence times.

rate-smoothing methods to estimate divergence times were compared using eight fossil and geological calibrations. Estimated times revealed that the Decapoda originated earlier than 407 MYA, and that the radiation within the group occurred rapidly, with all of the major lineages present by 300 MYA. Node time estimation under both approaches is severely affected by the number and phylogenetic distribution of the fossil calibrations chosen. More consistent results were obtained by using both shallow and deep or clade-related calibration points.

INTRODUCTION

Estimated to contain upwards of 10,000 species, the decapods are the most species-rich group of Crustacea, including shrimp (Caridea, Stenopodidea, and Thalassinidea), crabs (Anomala and Brachyura), and crayfish and lobsters (Astacidea and Achelata; Bowman and Abele 1982). Accordingly, the decapods have been the subject of more published papers than have all other crustaceans combined, due in part to their species richness, economic importance, and morphological diversity (Martin and Davis 2001). Decapod species have served as laboratory model organisms in studies of physiology, morphology, and behavior for over a century (Huxley 1880). Hence, given the prevalence of decapods in the public and scientific mind, understanding the evolutionary history of this significant crustacean group seems crucial.

Currently decapod evolutionary studies are centered on fossil data and morphology-based phylogenies. The decapod fossil record begins in the Late Devonian, (Schram et al. 1978) with representation of almost all of the major lineages (Schram 2000). In particular, the Reptantia have the best fossil record, as well as the oldest, of the

decapods (see Glaessner 1969). However, although all the main extant taxonomic groups have fossil representatives, the decapod record through time is incomplete (Schram 2001). While the majority of the described decapod fossils extend into the Cretaceous (Schram 2001), a large gap exists between these and the earliest known fossils, *Palaeopalaeomon newberryi* (Late Devonian, Schram et al. 1978) and *Imocaris tuberculata* (Lower Carboniferous, Schram 1984). Recently, a number of studies of decapod relationships have incorporated both fossil and extant taxa into a phylogenetic framework to examine evolutionary relationships and patterns of diversity through time (Amati et al. 2004; Rode and Babcock 2003; Schram and Dixon 2003; Tshudy 2003). However, while these studies have made great progress in understanding the evolution of the decapod form and the phylogenetic affiliations of fossil taxa, they are limited to groups where well-preserved fossils make comparisons of morphological characters with extant taxa possible.

Molecular phylogenetic methods can overcome these issues by combining sequence data with fossil dates, allowing the estimation of divergence times across the entire gene tree of a group by incorporating fossils into the analysis as calibration points. In the past this has been accomplished assuming a molecular clock, that is, constancy of evolutionary rates across lineages (Zuckerkandl and Pauling 1965). Under this assumption, the estimated branch lengths can be converted into absolute divergence times using fossil calibration. However, most datasets appear to violate the clock model (Graur and Martin 2004), which can cause serious bias in divergence time estimation (e.g., Rambaut and Broham 1998; Soltis et al. 2002). Consequently, in the last few years several methods have been proposed within Bayesian (Thorne and Kishino 2002; Yang

and Yoder 2003) and likelihood (Yang 2004; Yang and Yoder 2003) frameworks that account for rate variation when estimating divergence times, and incorporate multiple gene loci and multiple fossil calibration points. Both of these methods assume a phylogenetic hypothesis of evolutionary relationships, which must be estimated separately.

Unfortunately, there are as many hypotheses of decapod phylogenetic relationships as there are experts with opinions (Schram 2001), with no consensus in sight (Fig. 1). Historically, the decapod crustaceans were divided into two groups based on mode of locomotion: the Natantia (the ‘swimming’ lineages) and the Reptantia (the ‘crawling’ lineages) (Boas 1880). However, early on the ‘Natantia’ were recognized as a paraphyletic group and accordingly the Decapoda were reorganized into the suborders Dendrobranchiata (penaid shrimp and their relatives) and Pleocyemata (all other decapods) by Burkenroad (1963; 1981). This taxonomic restructuring is supported by several defining morphological characters (i.e. dendrobranchiate gill structure and pleocyemate brooding of eggs on the female’s pleopods), and phylogenetic studies showing the ‘natant’ decapods to be a paraphyletic assemblage (Abele 1991; Abele and Felgenhauer 1986; Felgenhauer and Abele 1983). Most of the phylogenetic studies investigating the relationships among the major decapod lineages have been based on morphological characters, which, due to the extreme diversity of form makes it difficult to study the group as a whole (Schram 1986). Moreover, there has been a surprising paucity of molecular phylogenetic studies investigating ordinal level relationships in this group. Those molecular studies that have been completed have focused on only part of the order (i.e. the ‘Natantia’) and have not included adequate taxon sampling within the

Reptantia to evaluate the relationships of the major infraorders (Abele 1991; Kim and Abele 1990). From a molecular perspective, no attempts at a comprehensive phylogenetic assessment of either the Reptantia, or the entire order have ever been undertaken.

Even given the large number of conflicting hypotheses regarding decapod phylogenetic relationships, there appears to be general agreement on the monophyly of the suborder Pleocyemata and the informal ‘Reptantia’. Towards the goal of investigating the divergence times of the major decapod radiations, particularly for these two consistently monophyletic clades, we will first construct a model-based phylogeny of the major decapod infraorders. This will be the first study to use molecular data in order to evaluate relationships among all the Decapoda infraorders, and particularly within the Reptantia. The combination of our molecular phylogeny with multiple fossil calibration points will be used for divergence time estimation under Bayesian and likelihood approaches to provide insights into the timing of the major decapod evolutionary radiations and into the relative performance of these two different methods in real data analyses.

METHODS

Taxon Sampling

The most updated classification of the recent Crustacea (Martin and Davis 2001) was used to determine the taxonomy of the major lineages within the Decapoda with two exceptions. First, the infraorder ‘Palinura’, which historically included the polychelids, palinurids, and glypheoids has been shown to be polyphyletic, with the glypheoids

clustering within Astacidea, and the polychelids shown to be basal reptants (Amati et al. 2004; Dixon et al. 2003; Scholtz and Richter 1995; Schram and Dixon 2003). Therefore, we chose to use the term ‘Achelata’ as suggested by Scholtz and Richter (1995) and Dixon et al. (2003) to represent the extant families Scyllaridae, Synaxidae, and Palinuridae. Second, the ‘Anomura’ lineage as described by Borradaile (1907) included both anomuran crabs and thalassinids. The distinction of the thalassinids as a lineage separate from the Anomura has been documented in numerous studies (Crandall et al. 2000; Dixon et al. 2003; Schram 2001; Schram and Dixon 2003); therefore, following the resurrection by others (Dixon et al. 2003; Scholtz and Richter 1995), we chose to replace ‘Anomura’ with the Anomala of Boas (1880). Species used for these analyses included representatives from the Dendrobranchiata and from all of the major infraorders in the Pleocyemata (Table 1). All specimens were preserved in 95-100% ethanol and are housed in the crustacean collection at the Monte L. Bean Life Science Museum, Brigham Young University, Provo, Utah. Based on previous hypotheses of Eumalacostraca relationships, two species of Euphausiacea were used to root the tree (Christoffersen 1988; Dixon et al. 2003; Schram 1986).

DNA Extraction, PCR, and Sequencing

Tissue samples from each specimen were dried and used in previously described DNA extraction protocols (Crandall and Fitzpatrick Jr. 1996). Polymerase chain reaction (PCR, Saiki et al. 1988) products for the complete 18S rDNA, (~2000 bp, Whiting 2002; Whiting et al. 1997) partial 28S rDNA (~2000 bp, Whiting 2002; Whiting et al. 1997) and histone H3 (333 bp, Colgar et al. 1998) nuclear genes, and the partial 16S (~460 bp,

Crandall and Fitzpatrick Jr. 1996) mitochondrial gene were amplified using one or more sets of general primers from the literature. Standard PCR conditions ($5 \mu\text{l}$ 10X *Taq* buffer, $6-8 \mu\text{l}$ 25mM MgCl₂, $8 \mu\text{l}$ 10mM dNTPs, $5 \mu\text{l}$ each of two 10 mM primers, 1.25 U *Taq*, $\sim 20 \mu\text{l}$ double distilled water) were used on a Perkin-Elmer 9700 machine under the following conditions: an initial denaturation at 96°C for 3 min followed by 40 cycles of 95°C for 1 min, 46°C for 1 min, and 72°C for 1min, followed by chain extension at 72°C for 10 min. PCR products were visualized by agarose (1.2%) gel electrophoresis and were purified using the Millipore Montage purification system. Sequences were generated in both directions on an ABI Prism 3730 capillary autosequencer using the ABI big-dye Ready-Reaction kit and following the standard cycle sequencing protocol but using 1/16th of the suggested reaction volume.

Phylogenetic Analyses

Nucleotide sequences were aligned using Clustal X (Thompson et al. 1997) with the default parameters and refined by eye. Because many regions of the 16S, 18S, and 28S gene segments used for analysis are extremely divergent among the ingroup taxa and therefore difficult to align reliably, GBlocks v0.91b (Castresana 2000) was used to eliminate poorly aligned positions and divergent regions of the Clustal X alignment (GBlocks parameters used for 16S / 18S / 28S: minimum number of sequences for a conserved position = 26/26/26; minimum number of sequences for a flanking position = 40/36/43; Maximum number of contiguous nonconserved positions = 8/8/8; minimum length of a block = 6/5/5; allowed gap positions = with half). Phylogenetic analyses of combined datasets have been shown to reveal hidden support for relationships in conflict

among analyses of individual markers (Gatesy et al. 1999); therefore, the GBlocks-pruned datasets from each gene region were concatenated into a single combined dataset consisting of 3601 bp. Because the goal is to date the major decapod radiations using model-based estimation procedures, in order to be methodologically consistent, only model-based methods of tree reconstruction were employed. The combined dataset was used to reconstruct phylogenies using Maximum Likelihood (ML) heuristic searches in PAUP* v4b10 (Swofford 2002), and Bayesian methods coupled with Markov chain Monte Carlo (BMCMC) inference as implemented in MrBayes v3.04b (Ronquist and Huelsenbeck 2003). Model selection for ML and BMCMC analyses followed the procedure outlined by Posada and Buckley (2004) as implemented in ModelTest v3.6 (Posada and Crandall 1998). For ML searches, a GTR+Γ+I model (base frequencies = 0.2593 0.2165 0.2737; Rmat=0.9538 2.7863 2.0907 0.9950 4.2081; gamma shape parameter = 0.5303, proportion invariable sites = 0.3830) was chosen for the concatenated dataset; for BMCMC analyses, models GTR+Γ+I (18S, 28S, 16S) and TVM+Γ+I (H3) were implemented in MrBayes. ML searches (Felsenstein 1981) were run using 100 random addition replicates and TBR branch swapping. Confidence in the resulting relationships was assessed using the nonparametric bootstrap procedure (Felsenstein 1981) with 200 bootstrap replicates, using heuristic searches of one random addition with TBR branch swapping per replicate. For BMCMC techniques, four independent analyses were run with each consisting of four chains. Each Markov chain was started from a random tree and run for 3.0×10^6 cycles, sampling every 1000th generation. Model parameters were treated as unknown variables with uniform default priors and were estimated as part of the analysis. To confirm that our Bayesian analyses

converged and mixed well, we monitored the fluctuating value of likelihood and all phylogenetic parameters graphically and compared means and variances of all likelihood parameters and likelihood scores from independent runs using the program Tracer v1.2 (Rambaut and Drummond 2003). All sample points prior to reaching stationary were discarded as burn-in. The posterior probabilities (pP) for individual clades obtained from separate analyses were compared for congruence, and then combined and summarized on a 70% majority-rule consensus tree (Huelsenbeck and Imennov 2002; Huelsenbeck et al. 2002).

Testing Alternative Hypotheses

Alternative *a priori* phylogenetic hypotheses from the literature were tested under both likelihood and Bayesian frameworks. Likelihood topology tests were conducted using our molecular data and the Shimodaira and Hasegawa (1999a, SH) test as implemented in PAUP*. Goldman et al. (2000), Buckley (2002), and Strimmer and Rambaut (2002) have pointed out that the SH test may be subject to a certain type of bias such that the number of trees included in the confidence set tends to be very large as the number of trees to be compared increases, which makes the test conservative. However, as these authors recognized and Shimodaira (2002) concluded, the SH test is still safe to use and is a good option when the number of candidate trees is not very large and more data are accumulated. Ten thousand replicates were performed for every topology test resampling the partial likelihoods for each site (RELL model). Because there are differences between the taxon sampling of the *a priori* hypotheses and our dataset, alternative topologies were constructed in MacClade by rearranging only the branches representing

the infraordinal lineages in conflict. Bayesian topology tests were performed by calculating the pP of the set of trees containing the *a priori* hypothesis, as described in Huelsenbeck et al. (2002).

Reference Fossils

The decapod fossil record is continually being updated and reclassified, due to new discoveries and because many fossils are described from incomplete specimens causing uncertainty as to their phylogenetic affinities. Consequently, where possible fossil references for this study were taken from species where descriptions were based on nearly-complete specimens or where recent phylogenetic studies have placed fossil species relative to extant groups (Amati et al. 2004; Rode and Babcock 2003; Schram and Dixon 2003; Tshudy 2003). Additionally, the fossils chosen for calibration points in this study were chosen based on the precision of the estimated date of the oldest known representative for particular clades, across several levels of divergence relative to the taxa sampling of our phylogeny. Based on these factors and the ages of potential fossils relative to their placement on the phylogeny, a set of seven fossils were used as calibrations in our analyses (Table 2). Additionally, because the Bayesian method chosen for divergence time estimation (see below) requires at least one calibration to consist of an upper limit (maximum age), we set the split between the crayfish superfamilies Astacoidea and Parastacoidea as an upper limit of 185 MYA based on the splitting of Pangea (Crandall et al. 2000).

Although fossil burrows attributed to crayfish have been described from the Permian, it is often difficult to determine this association with certainty (Babcock et al.

1998; Hasiotis 2002). Therefore, with respect to crayfish lineages we have chosen to use only fossil records from descriptions of preserved animals (Imaizumi 1938; Van Straelen 1928). Furthermore, a number of marine Jurassic fossil lobster species have been assigned to the Astacidea, although their phylogenetic relationships are still being investigated (Amati et al. 2004; Schram and Dixon 2003). Because the majority of these species are marine, they represent ancestral lineages to the crayfish. In terms of calibrations, we have chosen the oldest described marine lobster affiliated with the Astacidea, but not specifically aligned with the Nephropoidea, to calibrate the infraorder Astacidea.

The oldest fossil ascribed to the decapods is the Late Devonian *Palaeopalaemon newberryi* Whitfield, 1880, which has been placed within the Reptantia by several authors due to astacidean-like features (Christoffersen 1988; Felgenhauer and Abele 1983; Schram et al. 1978), although at least one of these also cites the presence of characters with ‘natantian’ affinities (Felgenhauer and Abele 1983). A recent phylogenetic study incorporating both fossil and extant taxa surprisingly places *P. newberryi* in a polytomy with the Thalassinida and ‘Eurysternalia’ (Achelata, Anomala, and Brachyura) (Schram and Dixon 2003), although there has been no consensus as to its phylogenetic affiliations. Therefore, our use of this fossil to date the split between the ‘natant’ forms and the ‘Reptantia’ clade is conservative. The second oldest known decapod fossil is the Brachyuran *Imocaris tuberculata* from the Lower Carboniferous (Schram 1984). Although novel relative to other hypotheses of reptant relationships, our phylogenetic analyses place the Brachyura at the base of the reptant clade and we therefore use this fossil to calibrate the reptant node.

Divergence Time Estimation

Decapoda divergence times were estimated using the Bayesian method of Thorne and Kishino (2002, referred to as TK) and the likelihood heuristic rate-smoothing algorithm (AHRS) of Yang (2004). The former approach is an extension of Thorne et al. (1998) and Kishino et al (2001) Bayesian methods and the latter builds on Yoder and Yang (2000) and Yang and Yoder (2003) likelihood methods. These extended versions can accommodate multiple fossil calibration points and multiple genes, allow for missing taxa, and in the case of AHRS facilitate automatic assignment of branches to rate groups using a rate-smoothing procedure (Sanderson 1997; Sanderson 2002). As previously shown, simultaneous analysis of gene sequences from multiple loci and multiple calibrations is expected to improve estimates of divergence times and rate estimates (Pérez-Losada et al. 2004; Thorne and Kishino 2002; Yang 2004; Yang and Yoder 2003). The two approaches implemented here estimate branch lengths without assuming a molecular clock, and then estimate times and rates by minimizing the discrepancies in branch lengths and by minimizing rate changes over branches. Moreover, both methods make use of the rate-evolution model of Thorne et al. (1998) and Kishino et al. (2001), but the TK approach averages over the rates in the MCMC procedure while the AHRS approach optimizes rates together with divergence times. Another difference is that the AHRS does not need a prior for divergence times, which might be considered an advantage. There is some evidence that time estimation by the Bayes approach may be sensitive to the prior model of the divergence times (Yoder and Yang 2004). In contrast, in the TK method it is possible to specify fossil calibrations as lower or upper bounds on

node ages. The likelihood method does not deal with such constraints and uses only fixed node ages for fossil calibration. As a result, standard errors calculated for estimated divergence times are serious underestimates. The importance of accounting for uncertainties in fossil calibrations has been emphasized by Graur and Martin (2004). Nevertheless, the performance of the TK and AHRs methods in real data analysis has never been explored, as these methods are only beginning to be widely used (Yang 2004). A recent study published by our group compared several Bayesian and likelihood approaches using 18S rDNA sequences and single calibrations (Pérez-Losada et al. 2004). Here we have extended the comparison to the case of multiple genes and multiple calibration points.

Bayesian-based TK method. We used the multi-locus Bayesian method of Thorne and Kishino (2002) as implemented in the multidivtime package (<http://statgen.ncsu.edu/thorne/multidivtime.html>). The mean of the prior distribution for the time separating the ingroup root from the present (rttm) and the standard deviation (SD) of this prior distribution (rttmsd) were set to 6 (600 MY). Alternative values ranging from 5 to 7 were also tried but final estimates did not change much (± 10 MY). After inspecting the branch lengths estimated by estbranches for each gene, the evolutionary rate of the root node was given a gamma prior distribution with mean (rrate) and SD (rratesd) both equal to 0.027 substitutions at the average site per 100 MY. We chose this prior to obtain a distribution for the root that was simultaneously reasonable and relatively diffuse. The rrate and rratesd were estimated as suggested in the multidivtime manual. Prior distributions approximated under the MCMC approach included a burn-in period of 10^6 steps, after which 10^6 samples were collected every 100

accepted states; posterior distributions (less diffuse) included a burn-in period of 5×10^5 steps, after which 10^6 samples were collected every 100 accepted states. Default options were chosen for all the other parameters of the prior distribution and the MCMC procedure. Convergence was monitored by checking the proportion of successes (psuc) of times and rate changes proposed along the Markov chain. Four independent chains were run from different starting points. Parameters of the evolutionary model were estimated under the F84+ Γ model (Felsenstein 2003), the most complex model implemented. This model is less parameterized than the best-fit models selected by ModelTest (see above), however, previous studies (see Yang and Yoder 2003 and references therein) have shown that it is actually the rate variation among sites parameter that has the greatest effect on divergence time estimation. All the parameters within the model as well as the branch lengths were estimated separately for every gene.

Likelihood-based AHRs method. We used the likelihood heuristic rate-smoothing algorithm of Yang (2004) as implemented in PAML 3.14 (Yang 1997). Sequence data were analyzed using the same F84+ Γ model and parameters of evolution chosen for the Bayesian analysis. Likelihood analyses were performed using SmallDiff (small value used in the difference approximation of derivatives) values of 1e-6 and 0.5e-6. Only the results showing the best likelihood scores are reported here. Branches at each locus were automatically classified into four rate groups according to their estimated rates (default option). This assignment was then checked manually using UPGMA in PHYLIP v.3.6a (Felsenstein 2003) as described in Yang (2004). Rate distributions among the four categories were fairly homogeneous for all genes.

Calibrations. Calibration points for the divergence time analysis were taken from known fossils representing major decapod lineages (Table 2). Given that most fossils are dated to an age range, the midpoint of each range was used for the divergence time estimations, using the 1999 GSA Geologic Time Scale to determine dates. Fossil calibrations were accommodated as lower limits (minimum ages) or as fixed ages, depending on the estimation method used and introduced into the analysis as follows: 1) under the TK method calibrations were used in separate analyses as lower limits (except for the Astacoidea/Parastacoidea split which was treated as an upper limit) or as fixed ages; 2) under the AHRS method calibrations were treated as fixed ages. All minimum or fixed age calibrations were mapped to the node prior to the basal node of the clade of interest.

The most important factors affecting divergence time estimation using molecular data are the number and distribution of the calibration points on the tree (Lee 1999; Thorne and Kishino 2002; Yang and Yoder 2003; Yoder and Yang 2000), although some methods seem to be more sensitive than others (Pérez-Losada et al. 2004). To explore the relative performance of the Bayesian TK and likelihood AHRS approaches at estimating divergence times relative to calibration number and distribution, we performed multiple analyses using 14 calibration schemes, and compared these results to the chronogram estimated using all the calibrations. For these particular analyses all calibrations were treated as fixed ages. To construct the calibration schemes, the eight calibrations (seven fossil dates plus the Pangea split) were arranged chronologically from oldest to youngest and separate analyses were run where in each consecutive analysis, a fossil calibration was removed, one at a time in chronological order until only a single

fossil remained. This process was repeated twice, first starting with removing the oldest fossils so that progressively younger fossils remained and the reverse where younger fossils were removed first.

Although divergence times were estimated under both TK and AHRS methods and using multiple combinations of calibration points to explore their relative performance, our best estimate (see below) of the diversification of the Decapoda lineages (including 95% confidence intervals; CI) was calculated using the TK Bayesian method and treating the seven fossil calibrations as minimum ages and the Astacoidea-Parastacoidea split as a maximum age.

RESULTS

Decapod Phylogenetics

We obtained 35 new complete 18S, and 33 partial 16S, 42 partial 28S, and 46 partial H3 gene sequences (Table 1). Tree topologies reconstructed in both ML and BMCMC methods were not conflicting (SH test $P=0.41$), although the BMCMC phylogeny was less resolved and therefore only the ML tree is presented (Fig. 2). In both analyses, the Pleocyemata, Reptantia, and all of the major infraorders were recovered as monophyletic clades with strong nodal support in at least one framework (thick black or grey branches, Fig. 2). However, there is very little support for infraordinal relationships within the Pleocyemata. This is evident when comparing our placement of the stenopod lineage with previous morphological hypotheses; the ML tree recovered a caridean + reptant clade (*a priori* hypothesis Fig. 1B), but this is not a significantly different topology than Fig. 1C (stenopod + reptant clade; SH $P=0.51$, $pP=0.42$) or Fig. 1A using the SH test

($P=0.18$). However, a caridean + stenopod clade arrangement (Fig. 1A) is a significantly worse hypothesis in the BMCMC analysis ($P=0.03$; Table 3).

With respect to relationships within the reptant clade, both the Pleocyemata and Reptantia clades were recovered with strong support in at least one method. Second, the Astacidea is monophyletic, containing the monophyletic nephropoid and astacid lineages. Third, the Thalassinidea is sister to the Astacidea, with weak pP support in BMCMC analyses. Finally, contrary to all but the only other molecular study including representatives of the major reptant lineages (Fig. 1F, Crandall et al. 2000), our analyses place the Brachyura and Anomala as the basal reptant lineages. In comparisons with *a priori* hypotheses, this arrangement is found to be significantly better than hypothesis Fig. 1D using the SH test, and to Fig 1D, E, and F using Bayesian pP (Table 3).

Decapod divergence Time Estimation

A likelihood ratio test significantly rejected ($P < 0.001$) the null hypothesis that all genes, separately and combined, were evolving with rate constancy across the decapods, justifying the use of non-clocklike molecular methods to estimate divergence times. The decapod TK chronogram based on the single ML topology and treating the calibration points as minimum or maximum ages is shown in Figure 2. All the major clade estimates including 95% CI are also shown in Table 4. Multiple independent Bayesian runs produced similar mean estimates, although the 95% CI were larger than expected; however, by constraining the age of one of the backbone calibrations within the interval of its first paleontological occurrence, the analysis produced similar mean divergence time estimates, but the SD was reduced by half (data not shown). The TK analysis places

the origin of the Dendrobranchiata and Pleocyemata decapod lineages in the early Devonian (407 MYA). This implies that the stem line of the decapods emerged even earlier; however, we are unable to estimate this age given our taxon sampling. Based on this analysis, the radiation of the major decapod lineages occurred rapidly. The reptant lineage originated 387 MYA and all of the major reptant infraorders were present by the late Carboniferous, 87 MY later. The radiation of the extant taxa within each infraorder, however, occurred at different periods of time. The natant lineages have an early origin (387-393 MYA), however the caridean superfamilies Alpheoidea, Atyoidea, and Palaemonoidea radiate in the Triassic (237 MYA). Among the Brachyuran superfamilies sampled, the Majoidea has the oldest lineage (233 MYA). The Achelata originate 315 MYA, with radiation of the extant lineages (Palinuridae and Scyllaridae) occurring as early as 225 MYA. The Thalassinidea appear 300 MYA, with the radiation of the Callianasoidea occurring at least 152 MYA.

Within the Astacidea and Anomala, we have sampled all the extant superfamilies. Therefore the divergence time estimates for the radiation of these groups are more accurate. The anomalan lineage originated 329 MYA, with the extant superfamilies radiating between 203-287 MYA. The Astacidea lineage originated 300 MYA, with the divergence between the astacid lineages (Astacoidea, Parastacoidea) and the Nephropoidea occurring 264 MYA. Within the astacids, the radiation of the Parastacidae (~130 MYA) occurred earlier than the Astacidae (88 MYA) or the Cambaridae (97 MYA). The Nephropodidae radiated as early as 143 MYA, with the genus *Homarus* appearing ~23 MYA.

Divergence time methods comparison

Decapod divergence times estimated under the TK approach using calibrations as minimum node ages were different from those estimated under the TK and AHRS methods using fixed age calibrations (Table 4). For four of the nodes corresponding to the Decapoda, Pleocyemata, Stenopodidea, and Reptantia the time differences ranged between 4-23 MY in all comparisons, but for the other six nodes the differences ranged between 42-96 MY for the TK-TK comparison and 43-123 MY for TK-AHRS comparison. However, the estimates using fixed calibrations were more congruent with each other, regardless of method.

Time chronograms estimated under the Bayesian and likelihood approaches using four genes and 14 different combinations of eight calibrations are presented in Fig. 3. These comparisons illustrate that divergence time estimates can be severely affected by the number and distribution of the calibrations used across the tree. For example, in Fig. 3A and C, as older fossil calibrations are progressively removed from the analysis, the estimates of the entire backbone of the phylogeny are pulled towards younger dates, with differences as large as 156 (TK) -257 (AHRS) MY between node estimates based on eight calibrations vs. only the youngest calibration. While the opposite trend is observed when removing younger calibrations from the analysis, older calibrations produce more stable backbone estimates, and hence more stable estimates across the tree, with the largest differences observed as 108 (AHRS) – 121 (TK) MY. While neither method remained stable as calibrations were removed, different trends were observed between the

methods relative to removing younger versus older calibrations. The AHRS estimates seem to be less perturbed by the removal of younger fossils while the TK method appears to be better able to deal with the removal of older fossils. Finally, we observed a crown effect, where removal of calibrations from a specific clade affected estimates within that lineage, while estimates across the rest of tree remained relatively stable. For example, in Fig. 3D, when calibration C1 from the Caridea and C6 from the Brachyura are removed, only the estimates within these lineages are significantly overestimated; however, these overestimations remain stable as calibrations are removed from other areas of the phylogeny.

DISCUSSION

Decapod radiation

This study presents the first molecular phylogenetic hypothesis of the infraordinal relationships within the Decapoda. However, it is not the final answer to the long debate regarding decapod relationships; indeed, it appears to add yet another scheme to the already large set of hypotheses concerning decapod phylogenetic relationships. However, our results do support several relationships that seem to be stable based on both molecules and morphology, i.e. the monophyly of the suborder Pleocyemata and the informal ‘Reptantia’ (Crandall et al. 2000; Dixon et al. 2003; Schram 2001). Furthermore, the infraorders included in our analyses are all strongly supported as monophyletic; however, this is a hypothesis that will continue to be tested as additional taxa from underrepresented decapod groups (especially from within the Caridea and Brachyura) are added to the molecular dataset. Of particular interest are several lineages

not represented in our analyses due to difficulty in obtaining the necessary specimens. The taxonomy of these groups, including the polychelids, glypheoids, thaumastochelids, and entoplometopodids, have been revised several times based on recent morphological estimates of phylogeny (Amati et al. 2004; Dixon et al. 2003; Scholtz and Richter 1995; Schram 2001), and inclusion in molecular analyses may provide additional insights into their phylogenetic placement within the decapods.

While there is strong support for the monophyly of the infraorders, there is little support for the relationships among them, and in fact, determining these relationships is one of the biggest remaining issues/controversies in decapod systematics (Abele 1991). While in our analyses the monophyletic Astacidea as sister to the Thalassinidea and the placement of Achelata close to the Anomala contradict Scholtz and Richter's (1995) hypothesis, this general arrangement is similar to at least one other study (Dixon et al. 2003). It also mirrors conjectures by Schram (1986) that the thoracic endoskeleton anatomy of Thalassinidea indicates a closer relationship to the astacideans than to the anomalans, and the observation by Tudge and Scheltinga (2002) that the resemblance of *Aegla* (Anomala) spermatozoa is closer to *Jasus* (Achelata) than to Thalassinidea. Perhaps the most controversial result of these analyses, however, is the placement of the Brachyura and Anomala as the basal reptant lineages. In fact, the arrangement of reptant lineages based on our molecular data is the reverse of that recovered in several of the most recent morphological phylogenies, at least one of which also uses euphausiids as an outgroup (see Fig. 1E; Dixon et al. 2003; Schram and Dixon 2003). Although this seems troubling on the surface, the similar branching patterns between molecular and morphological hypotheses is encouraging; there only seems to be a difference in the

polarization of the characters between methodologies. Given the extreme diversity of decapod forms, this is perhaps not too surprising. In particular, the highly modified morphologies of the Brachyura and Anomala predispose these lineages to be placed as the more derived clades in morphology-based phylogenies. Further investigations, including combined molecular and morphology phylogenetic analyses and studies of outgroup choice relative to character polarization, are required to understand these differences. However, by using model-based phylogenetic tree topology tests (ML and BMCMC), a statistical comparison of our results with previous morphological estimates is possible. These topology tests indicate that although the particular arrangement of the carid and stenopod lineages relative to the reptant lineages is unclear, it is most likely that they are not sister to each other (Fig. 1A). Furthermore, hypotheses based on Scholtz and Richter's (1995) data are significantly worse than our phylogenetic hypothesis in both SH and Bayesian topology tests (Table 3).

With respect to the only previous molecular hypothesis of reptant relationships (Fig. 1F) and to the morphological baseline of decapod phylogeny established by Dixon et al. (2003, Fig. 1E), ML topology tests find no significant difference, while Bayesian methods show significant differences. These results indicate that BMCMC methods are much more sensitive to topological differences than ML methods. Furthermore, none of these studies exhibit strong nodal support for reptant relationships. As a more conservative test however, the SH test indicates that there is no clear consensus between molecular and morphological estimates of decapod phylogenetic relationships.

The basal position of the Brachyura in our phylogeny, although contradictory to most other morphology-based hypotheses of decapod relationships, matches the current

understanding of the fossil record well. Provisionally, the Brachyura contain the second oldest known reptant fossil, *Imocaris tuberculata* (Schram 1984), indicating a long evolutionary history. Schram (1986) noted Brachyuran radiation events in the Cretaceous and in the Eocene when many of the modern families of crabs are found for the first time. However, our analysis indicates that many of the modern families may have had a much earlier origin. Also of interest relative to previous hypotheses of decapod crustacean radiations are the dates of astacid divergence. Our estimated divergence time of the astacid lineage in the early Permian (264 MYA) matches well with fossil crayfish and burrows associated with Permian and Early Triassic deposits (Hasiotis and Mitchell 1993) and the hypothesis by Crandall et al. (2000) that crayfish have a Pangean origin.

Although we have estimated decapod divergence times without assuming a molecular clock and using multiple molecular markers and fossil calibration points, our analyses come with a number of caveats. There are inaccuracies associated with the fossil record and with phylogeny estimation that are not taken into account (Graur and Martin 2004). We assumed that the fossil ages are known with no error. Moreover, the methods we have utilized are heavily dependent on topology and our molecular ML phylogeny is significantly different than most morphological hypotheses; therefore, our estimates represent only a single hypothesis of decapod evolution from a larger, incongruent set. These alternative topologies would possibly generate different estimates for the crown nodes of the infraorders, but the two main conclusions of our analyses - that the Decapoda originated in the Devonian and have experienced a fast radiation with all of the major infraorders present by the late Carboniferous - would not change. Nevertheless, future advances in divergence time estimation methodologies could take

advantage of the Bayesian framework to account for uncertainties in topology estimation and fossil dating and use different priors for rates and divergence times, as those included in Aris-Brosou and Yang (2002). An extension of this Bayesian approach to include multiple genes and calibrations would be desirable.

Divergence time estimation method comparison

Our methods comparison further illustrates the potential pitfalls of divergence time estimations, where number and phylogenetic distribution of calibrations can severely affect estimates. Since fossils do not fix the ages of internal nodes but merely constrain them to be minimum ages (Smith 1994), it seems more appropriate to constrain nodes to lie within some interval rather than fix them to a particular time (Norell 1992). This is one of the strengths of the TK method versus the AHRS algorithm. However, TK estimates have large confidence intervals. Where possible, this problem may be alleviated by including multiple upper limits in the analysis. While each of the two methods compared appears to have different strengths relative to the calibrations used (young versus old), in general using a combination of both deep and shallow calibrations will provide better estimates across the entire phylogeny. Furthermore, where possible, using at least one calibration within each crown lineage will help alleviate clade specific inaccuracies.

CONCLUSIONS

Rapid diversification and radiation is characteristic of the Crustacea as a whole (Schram et al. 1978), and this is a trend readily apparent in our divergence time estimates of decapod lineages (Fig. 2). Major decapod radiation events have been proposed in the Eocene (Brachyura, Schram 1986), the Cretaceous (Feldman 2004), and the Triassic (macrurous forms, Schram 1986). Our molecular-based divergence time estimates are earlier than hypotheses based solely on the fossil record, with the radiation of the ‘natant’ infraorders occurring in the Devonian, the reptant infraorders in the Carboniferous, Anomalan diversification in the Permian-Triassic, and the Callianassoidea and Palaemonoidea in the Cretaceous. As decapod paleontological research is quickly expanding (Feldmann 2003), it will be most interesting to track the knowledge of decapod fossil date ranges relative to molecular-based divergence time estimations.

ACKNOWLEDGEMENTS

We thank D. Abed, E. Benavides, J. Buhay, J.M. Foster, T. Frank, S. Harrison, S. Nates, G. Poore, G. Svenson, M. Terry, C. Tudge, M. Wicksten, and J.D. Williams for assistance in collecting specimens. We also thank S. De Grave for identifications and S. Quinlan and T. O’Connor for help with labwork. We are very grateful to R. Feldmann and F. Schram for valuable comments concerning fossil decapods. This study was supported by grants from the National Science Foundation (DEB-0075600 and DEB-0206537).

Table 4-1: Taxonomy and GenBank accession numbers for gene sequences from Decapoda species included in this study. Sequences obtained from GenBank are indicated in bold. An ‘N’ designates gene sequences we were unable to acquire.

Taxon	Gene			
	16s	18s	28s	H3
Decapoda Latreille, 1802				
Dendrobranchiata Bate, 1888				
Penaeoidea Rafinesque, 1815				
<i>Penaeus semisulcatus</i> de Haan, 1844	DQ079731	DQ079766	DQ079809	DQ079698
Pleocyemata Burkenroad, 1963				
Stenopodidea Claus, 1872				
<i>Stenopus hispidus</i> (Olivier, 1811)	DQ079734	DQ079769	DQ079812	DQ079701
Caridea Dana, 1852				
Atyoidea de Haan, 1849				
<i>Atyoida bisulcata</i> (Randall, 1840)	DQ079704	DQ079738	DQ079774	DQ079661
<i>Typhlatya pearsei</i> Creaser, 1936	DQ079735	DQ079770	DQ079813	DQ079702
Alpheoidea Rafinesque, 1815				
<i>Lysmata debelius</i> Bruce, 1983	DQ079718	DQ079752	DQ079793	DQ079681
<i>Lysmata wurdemanni</i> (Gibbes, 1850)	DQ079719	DQ079753	DQ079794	DQ079682
Palaemonoidea Rafinesque, 1815				
<i>Creaseria morleyi</i> (Creaser, 1936)	DQ079710	DQ079746	DQ079784	DQ079671
<i>Cryphiops caementarius</i> (Molina, 1782)	DQ079711	DQ079747	DQ079785	DQ079672
<i>Macrobrachium potiuna</i> (Muller, 1880)	DQ079721	DQ079756	DQ079797	DQ079685
<i>Macrobrachium</i> sp.	DQ079720	DQ079754	DQ079795	DQ079683
<i>Palaemon elegans</i> Rathke, 1837	DQ079729	DQ079764	DQ079807	DQ079696
<i>Palaemonetes paludosus</i> (Gibbes, 1850)	N	DQ079755	DQ079796	DQ079684
‘Reptantia’				
Achelata Scholtz and Richter, 1995				
Palinuroidea Latreille, 1802				
<i>Jasus edwardsii</i> (Hutton, 1875)	DQ079716	AF235972	DQ079791	N
<i>Panulirus regius</i> De Brito Capello, 1846	DQ079730	DQ079765	DQ079808	DQ079697
<i>Scyllarus arctus</i> (Linnaeus, 1758)	DQ079732	DQ079767	DQ079810	DQ079699
Anomala				
Galatheoidea Samouelle, 1819				
<i>Aegla abtao</i> Schmitt, 1942	AY050067	AF439390	AY595965	DQ079658
<i>Uroptychus parvulus</i> (Henderson, 1885)	AY595926	AF439386	AY596097	DQ079703
<i>Munida subrugosa</i> (White, 1847)	AY050075	AF439382	AY596099	DQ079688
Hippoidea Latreille, 1825				
<i>Emerita brasiliensis</i> Schmitt, 1935	DQ079712	AF439384	DQ079786	DQ079673
Lomisoidea Bouvier, 1895				
<i>Lomis hirta</i> (Lamarck, 1810)	AY595928	AF436013	AY596101	DQ079680
Paguroidea Latreille, 1802				
<i>Lithodes santolla</i> (Molina, 1782)	AY595927	AF439385	AY596100	DQ079679
Astacidea Latreille, 1802				
Astacoidea Latreille, 1802				
<i>Astacus astacus</i> (Linnaeus, 1758)	AF235983	AF235959	DQ079773	DQ079660
<i>Cambarellus shufeldtii</i> (Faxon, 1884)	AF235986	AF235962	DQ079778	DQ079665
<i>Cambaroides japonicus</i> (de Haan, 1841)	AF235987	DQ079742	DQ079779	DQ079666
<i>Cambarus maculatus</i> Hobbs and Pfleiger, 1988	AF235988	AF235964	DQ079780	DQ079667
<i>Orconectes virilis</i> (Hagen, 1870)	AF235989	AF235965	DQ079804	DQ079693
<i>Pacifastacus leniusculus</i> (Dana, 1852)	AF235985	AF235961	DQ079806	DQ079695
Parastacoidea				
<i>Astacopsis gouldi</i> (Horwitz, 1991)	AF135969	DQ079737	DQ079772	DQ079659
<i>Cherax glaber</i> Riek, 1967	AF135978	DQ079745	DQ079783	DQ079670

Nephropoidea Dana, 1852				
<i>Acanthacaris caeca</i> (A. Milne-Edwards, 1881)	N	DQ079736	DQ079771	N
<i>Homarus americanus</i> H. Milne-Edwards, 1837	AF370876	AF235971	DQ079788	DQ079675
<i>Homarus gammarus</i> (Linnaeus, 1758)	DQ079714	DQ079749	DQ079789	DQ079676
<i>Nephrops norvegicus</i> (Linnaeus, 1758)	DQ079726	DQ079762	DQ079803	DQ079692
<i>Nephropsis aculeata</i> Smith, 1881	DQ079727	DQ079761	DQ079802	DQ079691
Brachyura Latreille, 1802				
Cancroidea Latreille, 1802				
<i>Cancer pagurus</i> Linnaeus, 1758	DQ079708	DQ079743	DQ079781	DQ079668
Grapsoidea MacLeay, 1838				
<i>Pachygrapsus marmoratus</i> (Fabricius, 1787)	DQ079728	DQ079763	DQ079805	DQ079694
Majoidea Samouelle, 1819				
<i>Maja squinado</i> (Herbst, 1788)	DQ079723	DQ079758	DQ079799	DQ079687
Potamoidea Ortmann, 1896				
<i>Geothelphusa</i> sp.	DQ079715	DQ079750	DQ079790	DQ079677
Portunoidea Rafinesque, 1815				
<i>Carcinus maenas</i> (Linnaeus, 1758)	DQ079709	DQ079744	DQ079782	DQ079669
<i>Macropipus puber</i> (Linnaeus, 1767)	DQ079722	DQ079757	DQ079798	DQ079686
<i>Necora puber</i> (Linnaeus, 1767)	DQ079724	DQ079759	DQ079800	DQ079689
Thalassinidea				
Callianassoidea Dana, 1852				
<i>Biffarius arenosus</i> (Poore, 1975)	DQ079705	DQ079739	DQ079775	DQ079662
<i>Callichirus major</i> (Say, 1818)	DQ079707	DQ079741	DQ079777	DQ079664
<i>Callianassa subterranea</i> (Montagu, 1808)	DQ079706	DQ079740	DQ079776	DQ079663
<i>Lepidophthalmus louisianensis</i> (Schmitt, 1935)	DQ079717	DQ079751	DQ079792	DQ079678
<i>Sergio mericeae</i> Manning and Felder, 1995	DQ079733	DQ079768	DQ079811	DQ079700
OUTGROUPS				
Euphausiacea Dana, 1852				
<i>Euphausia eximia</i> Hansen, 1911	DQ079713	DQ079748	DQ079787	DQ079674
<i>Nematoscelis</i> sp.	DQ079725	DQ079760	DQ079801	DQ079690

Table 4-2. Taxonomy and ages of fossils used as calibrations for divergence time estimations. Calibration C8 is 185 MYA, based on the splitting of Pangea.

Taxonomy	Species	Reference	Geologic Age (MYA)	Node #
Suborder Pleocyemata				
Infraorder Caridea				
Family Palaemonidae	<i>Palaemon antonellae</i>	(Garassino and Bravi 2003)	Early Cretaceous (Albian) (99-112)	C1
	<i>Alburnia petinensis</i>	(Bravi and Garassino 1998)	Early Cretaceous (Albian) (99-112)	C1
'REPTANTIA'	<i>Palaeopalaemon newberryi</i>	(Whitfield 1880)	Late Devonian (Famennian) (354-364)	C2
Infraorder Astacidea				
Family Chimaerastacidae	<i>Chimaerastacus pacifluvialis</i>	(Amati et al. 2004)	Mid Triassic (Upper Ladinian) (227-234)	C3
Superfamily Astacoidea				
Family Astacidae	<i>Astacus licenti</i>	(Van Straelen 1928)	Late Jurassic (144-159)	C4
	<i>Astacus spinirostris</i>	(Imaizumi 1938)	Late Jurassic (144-159)	C4
Infraorder Anomala				
Family Aeglidae	<i>Protaegla miniscula</i>	(Feldmann et al. 1998)	Early Cretaceous (Albian) (99-112)	C5
Infraorder Brachyura				
Family Cancridae	<i>Notocarcinus sulcatus</i>	Schweitzer et al. 2002	Mid Eocene (41.3-49)	C6
	<i>Imocaris tuberculata</i>	(Schram 1984)	Early Carboniferous (323-354)	C7

Table 4-3. Likelihood (S-H) and BMCMD topology tests of previous hypotheses of decapod relationships, as shown in Fig. 1A-F. For S-H tests, the difference in likelihoods ($\Delta\text{-lnL}$) and the corresponding P values are indicated. In BMCMD analyses, the number of trees (N) congruent with the previous hypothesis out of the posterior distribution of 11,400 trees is shown, with the corresponding posterior probability (pP) values.

Figure 1	S-H		BMCMD	
	$\Delta\text{-lnL}$	P value	N	pP
A	3.37	0.18	372	0.03
B	---	---	3013	0.26
C	0.51	0.45	4799	0.42
D	25.56	0.03	0	<0.001
E	6.28	0.26	15	0.001
F	6.16	0.17	1	0.00009

Table 4-4. Comparison of divergence times for major decapod lineages as estimated from the TK method incorporating calibrations as minimum ages, and the TK and AHRS methods using calibrations as fixed ages. Divergence times are taken from the crown node in each clade except for the Stenopodidea, where there is only a single representative included in this analysis. Because the Reptantia node contained a calibration, in the fixed age analyses this estimate is constrained to be 339 MYA; these calibration times are indicated in bold. Node numbers for each clade correspond to node numbers included on the chronogram in Fig. 2.

Taxon (Node)	Divergence Time (95% CI) MYA		
	TK	TK	AHRS
	<i>Minimum age</i>	<i>Fixed age</i>	<i>Fixed age</i>
Decapoda (90)	407 (374-460)	386 (366-414)	411 (410-412)
Pleocyemata (89)	393 (366-443)	370 (360-387)	411 (410-412)
Caridea (54)	237 (195-284)	150 (133-169)	139 (129-149))
Stenopodidea (89)	393 (366-443)	370 (360-387)	411 (410-412)
Reptantia (87)	356 (339-398)	339	339
Achelata (67)	225 (169-283)	171 (128-218)	159 (135-183)
Anomala (65)	287 (245-333)	191 (157-232)	244 (221-267)
Astacidea (83)	264 (227-304)	213 (195-229)	217 (207-226)
Brachyura (60)	233 (181-286)	141 (97-194)	110 (91-130)
Thalassinidea (71)	152 (105-205)	110 (78-146)	109 (95-123)

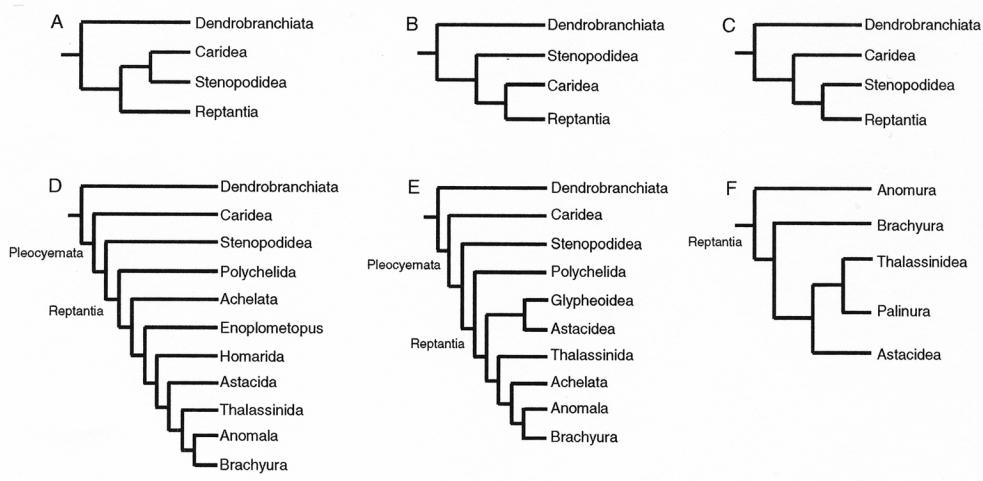


Figure 4-1. Previous hypotheses of decapod relationships, with A-C illustrating morphological hypotheses relative to ‘natant’ lineage relationships, D-E illustrating morphological hypotheses including ‘reptant’ lineage relationships, and F illustrating a molecular hypothesis of ‘reptant’ lineages only A) Burkenroad (1963; 1981); B) Christofferson (1988); C) Abele and Felgenhauer (1986), Abele (1991), Schram (1986); D) Schram (2001), based on reevaluation of data from Scholtz and Richter (1995); E) Dixon et al. (2003); F) Crandall et al. (2000).

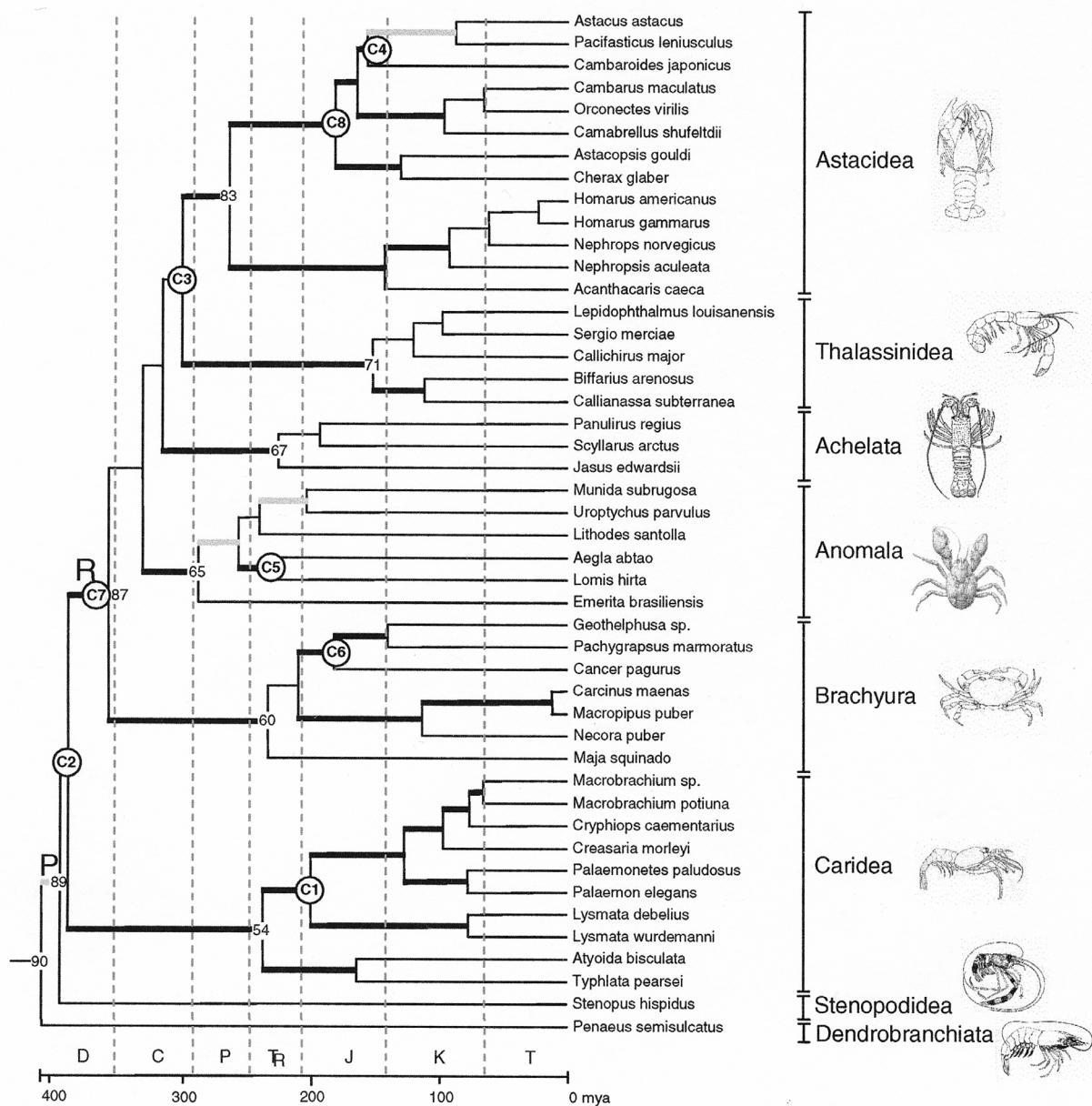


Figure 4-2. Decapod divergence time chronogram estimated using topology of ML tree. On branches with both ML bootstrap values >70% and BMCMC $pP > 0.95$, support is indicated by a thick black line; branches strongly supported by only one tree reconstruction method are indicated by thick grey lines. Fossil calibration nodes are indicated by C1-C8. Node numbers from divergence time estimations are included for reference on nodes of important decapod lineages. The decapod infraorders are delineated, and the nodes corresponding to the suborder Pleocyemata (P) and the informal Reptantia (R) are indicated on the phylogeny. The major geologic periods are also mapped onto the phylogeny, using the following standard symbols: D = Devonian, C = Carboniferous, P = Permian, T_R = Triassic, J = Jurassic, K = Cretaceous, T = Tertiary.

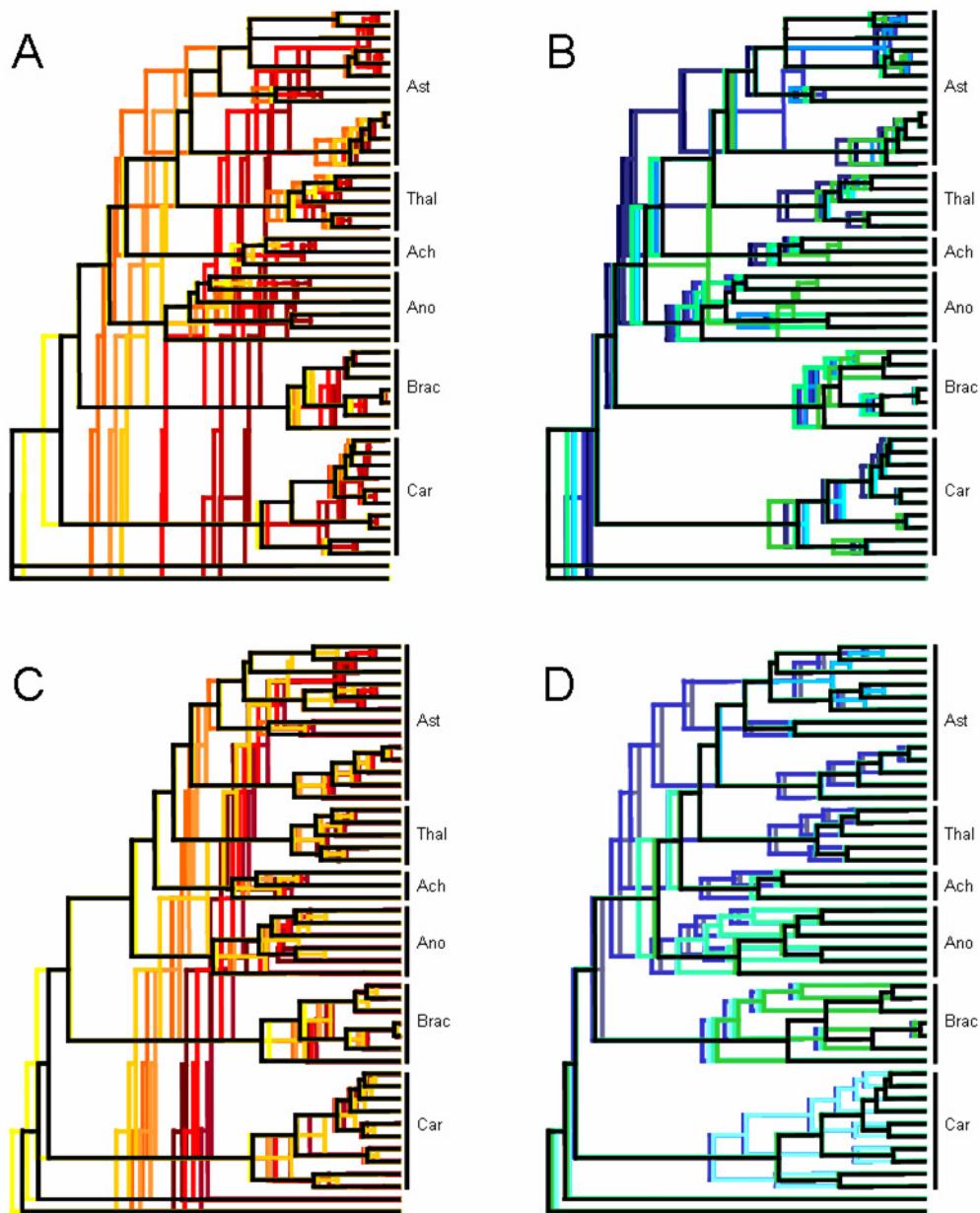


Figure 4-3. Comparison of divergence time estimates from Bayesian (TK) and Likelihood (AHRS) methods under 15 different calibration schemes. Calibrations are used as fixed ages in all analyses. In each panel, the best estimate chronograms based on all eight calibrations are black. Estimated chronograms from the successive removal of calibrations are mapped behind our best estimate, with the successive removal of deep (older) calibrations shown in reds/yellows and of shallow (younger) calibrations in blues/greens. For deep calibration removals, yellow indicate the analysis with removal of only one calibration, while dark red indicates removal of all but the youngest calibration. For deep node analyses, green denotes the analyses with removal of the first calibration while dark purple is the analysis with only the oldest calibration. The major decapod lineages are indicated as follows: Ast – Astacidea; Thal – Thalassinidea, Ach –Achelata, Ano – Anomala, Brac – Brachyura, Car – Caridea. A) AHRS estimates comparing chronograms from the successive removal of deep calibrations; B) AHRS estimates comparing chronograms from the successive removal of shallow calibrations; C) TK estimates comparing chronograms from the successive removal of deep calibrations; D) TK estimates comparing chronograms from the successive removal of young calibrations.

CHAPTER 5.

THE MOLECULAR EVOLUTION OF INVERTEBRATE OPSINS

ABSTRACT

Investigations of invertebrate opsin evolution have long been focused on insect visual pigments. Other invertebrate groups, in particular the crustaceans, have received little attention. Furthermore, few studies have explicitly investigated the selective influences across all known invertebrate opsins. In this study, we contribute to the knowledge of crustacean opsins by sequencing the opsin gene from six previously uncharacterized crustacean species (*Euphausia superba*, *Homarus gammarus*, *Archaeomysis grebnitzkii*, *Holmesimysis costata*, *Mysis relicta* sp. IV, and *Neomysis americana*). Furthermore, visual pigment spectral sensitivities were characterized using microspectrophotometry for *A. grebnitzkii* (496 nm), *H. costata* (512), *M. relicta* sp. IV (501), and *Neomysis americana* (520 nm). These novel crustacean opsin sequences were included in a phylogenetic analysis with previously characterized (both genetically and spectrally) invertebrate opsin sequences to determine the evolutionary placement of our opsin sequences relative to the well-established insect spectral clades (long-/middle-/short-wavelength sensitive). Phylogenetic analyses indicate these novel crustacean opsins form a well-supported clade with previously characterized crayfish opsin sequences, and form a sister-group to insect middle-/long-wavelength sensitive opsins. The reconstructed invertebrate opsin phylogeny was used to investigate selective influences within invertebrate opsin evolution using standard d_N/d_S ratio methods and more sensitive techniques investigating the amino acid property changes resulting from nonsynonymous replacements in a historical (i.e. phylogenetic) context. While the conservative d_N/d_S

methods did not detect any selection, four amino acid properties (coil tendencies, compressibility, power to be at the middle of an alpha helix, and refractive index) were found to be under destabilizing positive selection. Ten amino acid sites relating to these properties were found to be facing the binding pocket, within 4Å of the chromophore, with potential to affect spectral tuning.

INTRODUCTION

Visual pigment research has long been of interest to a number of biological disciplines, including sensory ecologists, visual physiologists, biochemists, and molecular evolutionists. Composed of a chromophore bound to an integral membrane protein (opsin), visual pigments are phenotypically characterized by the wavelength of maximal absorption (λ_{\max}). As most visual pigments contain the same chromophore, variation in opsin sequence is responsible for most of the observed variation in λ_{\max} . The ability to study the genetic mechanisms behind phenotypic variation has made opsin a model evolutionary system. In vertebrates, these studies often take the form of site-directed mutagenesis studies where the effect of changes at a single amino acid residue on λ_{\max} can be quantified (Asenjo et al. 1994; Cowing et al. 2002b; Nathans 1990b; Neitz et al. 1991; Wilkie et al. 2000; Yokoyama and Radlwimmer 1999; Yokoyama and Radlwimmer 2001; Yokoyama et al. 2000; Yokoyama and Tada 2003). With regard to spectral tuning, many of the hypotheses have involved the distribution of charged or polar residues relative to the chromophore binding pocket (Chan et al. 1992; Nathans 1990b; Neitz et al. 1991). In contrast, most studies of invertebrates have used comparative methods and homology modeling to identify sites potentially important in spectral tuning

(Briscoe 2001; Briscoe 2002; Chang et al. 1995; Chase et al. 1997; Crandall and Cronin 1997) while a brave few have investigated spectral tuning using *Drosophila* heterologous expression systems (Britt et al. 1993; Salcedo et al. 1999; Salcedo et al. 2003). These studies have confirmed that there are some similarities between spectral tuning sites in vertebrates and invertebrates (Briscoe 2001; Salcedo et al. 2003) but also illustrated that there are a number of unique residues affecting spectral tuning only in invertebrates. Given that the visual pigment genes diversified into the observed spectral clades independently, after the separation between invertebrates and vertebrates (Pichaud et al. 1999), some differences in functionality are expected. In fact, several significant differences in the mechanism of photoactivation have been documented, most notably differences in photoactivation related to the Schiff base. In vertebrates, a negatively charged counterion (E113) stabilizes the positive charge of the Schiff base (Nathans 1990a; Sakmar et al. 1989; Zhukovsky et al. 1992). However, in invertebrates this same site is a tyrosine (or in UV-sensitive opsins a phenylalanine). Studies have shown that this residue does not function as a counterion and the exact method of photoactivation is not yet completely understood in invertebrates (Nakagawa et al. 1999; Salcedo et al. 2003).

While a plethora of studies have looked at structure / function relationships in vertebrate opsin from an evolutionary perspective, fewer studies have investigated similar issues in invertebrates. Much less is known about the mode and tempo of invertebrate opsin evolution. Most of the evolutionary work related to visual pigments in invertebrates has focused on insect systems, delineating at least four main spectral classes: long wavelength sensitive (LWS), middle wavelength sensitive (MWS), and two

short-wave sensitive (SWS) groups (UV and blue) (Briscoe 2001; Briscoe 2002; Briscoe and Chitkka 2001; Carulli et al. 1994; Carulli and Hartl 1992; Feiler et al. 1992; Feiler et al. 1988; Montell et al. 1987; O'Tousa et al. 1985; Salcedo et al. 1999; Smith et al. 1997; Taylor et al. 2005; Towner and Gartner 1994; Zuker et al. 1987). Outside of insects, much of the opsin variation in the invertebrate world remains uncharacterized. Although there have been a large number of physiological studies of spectral sensitivities in other major arthropod groups, particularly within crustaceans, how these spectral variants relate to the defined insect clades is unknown. Furthermore, outside of insects, the only invertebrate taxa in which opsin sequences have been explicitly investigated are horseshoe crabs (Smith et al. 1993), molluses (Brown and Brown 1958; Hall et al. 1991; Hubbard and St. George 1958; Morris et al. 1993), and crustaceans (Crandall and Cronin 1997; Crandall and Hillis 1997; Oakley and Huber 2004; Sakamoto et al. 1996).

This research is focused in investigating opsin evolution in invertebrate systems. While a number of insect groups have been the focus of evolutionary studies (most notable of *Drosophila* and *Lepidoptera*; Briscoe 2000; Briscoe 2001; Briscoe 2002; Carulli et al. 1994) and representatives from known invertebrate spectral variants are often used to phylogenetically place novel opsins and form hypotheses about uncharacterized spectral sensitivities, few studies have focused on explicitly investigating the selective forces acting upon the diversity of invertebrate opsins. We add to the knowledge of non-insect opsins by isolating sequences from six additional crustacean species and by characterizing the λ_{\max} in those species without previous estimates. These data are added to the complement of invertebrate opsins that have been characterized both genetically and spectrally to investigate selective influences relative to spectral

variation. Traditional methods using d_N/d_S ratios are compared to more sensitive methods that investigate the relative change in amino acid properties resulting from nonsynonymous replacements in a historical (i.e. phylogenetic) context.

METHODS

Taxon sampling and outgroup choice

Opsin sequence data were collected from six crustacean species. Two of these species, *Homarus gammarus* and *Euphausia superba*, have previously been characterized with respect to spectrally sensitivities and expand the number of opsin sequences from within the Eucarida, adding to the five previously characterized crayfish sequences. Four species, *Archaeomysis grebnitzkii*, *Holmesimysis costata*, *Mysis relicta*, and *Neomysis americana*, represent the Mysida, an order of crustaceans from which λ_{\max} have been previously characterized (Gal et al. 1999; Lindström 2000) but no opsin sequences have been isolated. Finally, three sequences from the stomatopod *Neogonodactylus oerstedii* that were characterized for a Ph.D. thesis (Brown 1996) and have corresponding λ_{\max} data but have never been published were included in our analyses. With the addition of these data, we have added opsin sequence data from three orders of crustaceans (Euphausiacea, Mysida, Stomatopoda) to our analyses that have never been included in evolutionary analyses of invertebrate opsins. To these data we added any opsin sequence data for invertebrate species that represented at least half of the transmembrane spanning portion of the gene and had corresponding λ_{\max} values previously reported in the literature (Table 1). Outgroups were chosen from the vertebrate opsin clade, based on the hypothesis that duplication of opsin genes occurred independently in the lines of descent leading to

invertebrates and vertebrates (Pichaud et al. 1999). Bovine rhodopsin was included as the only GPCR to have been crystallized (Palczewski et al. 2000). The inclusion of this sequence in our alignment allowed for sites identified by selection detection methods to be mapped to the bovine protein structure (see below). Two representatives of pineal opsin and an ‘orphan’ human opsin paralog (GPR52) were included as outgroups, representing basal lineages to the vertebrate visual pigment clade (Bellingham et al. 2003; Fredriksson et al. 2003). Finally, human melatonin receptor 1A, a GPCR that has been phylogenetically placed close to the human visual pigment clade (Fredriksson et al. 2003), was chosen as a distant outgroup.

Microspectrophotometry (MSP)

For spectral analyses, live specimens were shipped overnight in dark conditions to the University of Maryland Baltimore County. In most cases, animals were used within a week of collection. All organisms were dark-adapted at least overnight, but more commonly for several days, before use. Eyes were removed under dim red light, mounted in tissue medium, and flash frozen. Frozen eye samples were sectioned immediately using a cryostat microtome to produce ~14 μm thick sections. Individual sections were mounted on cover slips and scanned under dim red light on a microscope for usable rhabdom structures. Suitable sections were mounted in Ringers buffer solution between coverslips sealed with a ring of silicone grease. The equipment and general procedure used for MSP have been described by Cronin (1985). Briefly, a linearly polarized scanning beam was placed within a single rhabdom. Scans were made from 400 to 700 nm, with measurements taken at 1-nm steps. Each dark-adapted rhabdom was

scanned twice to check for stability. If the two scans were identical, the second was saved as the direct absorption spectrum of the dark-adapted photoreceptor. The rhabdom was then exposed to two minutes of bright white light, followed by a second absorption scan. During bright-light treatments, the field diaphragm of the sub-stage illuminator was closed down to produce a spot of ~10 μm diameter at the level of the rhabdom, minimizing local heating of the preparation. For each photoreceptor, the rhodopsin absorption spectrum was taken as the difference between the initial, dark-adapted spectrum and the final, photobleached spectrum. The wavelength of maximum absorption (λ_{\max}) was estimated for each difference spectrum using a least squares procedure (see Cronin et al. 1994a). We compared all photobleach difference spectra to standard rhodopsin and porphyropsin templates derived by Stavenga et al. (1993) and subsequently averaged together those that closely resembled either template. The average spectrum were then fitted to the corresponding template again to determine a λ_{\max} value that best represents the spectra of the measured visual pigments. Results from 6-10 rhabdoms, representing two or more individuals, were obtained for each species.

DNA Extraction, PCR, Cloning, and Sequencing

All specimens were stored in 70-95% ethanol and kept at 4°C until extracted. Genomic DNA was extracted using Qiagen DNeasy kits (Qiagen) following the manufacturer's instructions. Polymerase chain reaction (PCR, Saiki et al. 1988) products for the opsin gene were amplified using a semi-nested degenerate PCR strategy. An initial PCR using primers LF1a: 5' TGG TAY CAR TWY CCI CCI ATG AA 3' and OPSRD: 5'CCR TAN ACR ATN GGR TTR TA 3' (Chang et al. 1996) with standard

conditions (2.5 μ l 10X *Taq* buffer, 4 μ l 10mM dNTPs, 2.5 μ l each of two 10 mM primers, 1.25 U HotMaster *Taq* (Eppendorf), ~ 12.5 μ l double distilled water) was run on a Perkin-Elmer 9700 machine for 35 cycles of 95°C for 30 sec., 48°C for 45 sec., and 70°C for 1:15 min., followed by chain extension at 72°C for 15 min. The first round PCR reactions were then diluted 1:10 with sterile water and used as template for a second round PCR of another 35 cycles using primers F1a and Scylla: 5' TTR TAI ACI GCR TTI GCY TTI GCR AA 3' (Taylor et al. 2005). Second round PCR products were visualized by agarose (1.2%) gel electrophoresis. Visible opsin DNA bands were excised and cleaned from the agarose gel using a GeneClean II kit (Bio 101). Purified opsin PCR products were cloned using the TOPO TA Cloning Kit (Invitrogen, San Diego, California), following manufacturer instructions. For each species examined, 3-10 clones were lysed in 50 μ l buffer (10mM Tris-HCl; 0.1 mM EDTA, pH 8.0) for 10 min at 96°C and inserts were PCR-amplified from lysed cells using plasmid-specific primer pairs M13(-20) (5' GTAAAACGACGGCCAGT 3') and M13(-24) (5' AACAGCTATGACCATG 3') and the following PCR conditions: denaturation at 94°C for 1 min, primer annealing at 55 °C for 1 min, chain extension at 72 °C for 3 min, for 30 cycles. Clone PCR products were purified using the Millipore Montage purification system and sequenced in both directions on an ABI Prism 3730 capillary autosequencer using the ABI big-dye Ready-Reaction kit and following the standard cycle sequencing protocol but using 1/16th of the suggested reaction volume.

Phylogenetic Analyses

Opsin sequences were either generated in the lab (see above section) or were downloaded from Genbank (<http://www.ncbi.nlm.nih.gov>; see Table 1). All of the species included in the analyses were from species where both the opsin sequence and the λ_{\max} have been previously characterized or were measured in this study. Only sequences that spanned more than half of the opsin transmembrane domains were used to reduce spurious arrangements caused by short sequences and to strengthen the analyses of selection (see below section). MacClade (Maddison and Maddison 2003) was used to create an initial sequence alignment which was then converted to amino acid sequences for alignment. Amino acid sequences were aligned in Clustal X (Thompson et al. 1997) using a Blosum log₂ weight matrix generated using MatrixGen (<http://matrixgen.sourceforge.net/>) with an aligned sequence of opsin GPCRs from the GPCR website (www.gpcr.com). The initial nucleotide sequence alignment was then adjusted to match the resulting amino acid alignment and fine-tuned based on structural information (Chang et al. 1995; Palczewski et al. 2000)

Phylogenetic trees were reconstructed from both the nucleotide and amino acid alignments. The nucleotide phylogeny was reconstructed using Bayesian methods coupled with Markov chain Monte Carlo (BMCMC) inference as implemented in MrBayes v3.04b (Ronquist and Huelsenbeck 2003). Because different codon positions have different functional constraints, the dataset was partitioned into first, second, and third codon positions for mixed model analyses. Model selection for each partition followed the procedure outlined by Posada and Buckley (Posada and Buckley 2004) for AIC as implemented in ModelTest v3.6 (Posada and Crandall 1998). Four independent

BMCMC analyses were run with each consisting of four chains. Each Markov chain was started from a random tree and run for 6.0×10^6 cycles, sampling every 1000th generation. Model parameters were treated as unknown variables with uniform default priors and were estimated as part of the analysis. To confirm that our Bayesian analyses converged and mixed well, we monitored the fluctuating value of likelihood and all phylogenetic parameters graphically and compared means and variances of all likelihood parameters and likelihood scores from independent runs using the program Tracer v1.2 (Rambaut and Drummond 2003). All sample points prior to reaching stationary were discarded as burn-in. The posterior probabilities (pP) for individual clades obtained from separate analyses were compared for congruence and then combined and summarized on a majority-rule consensus tree (Huelsenbeck and Imennov 2002; Huelsenbeck et al. 2002). A phylogeny based on the amino acid alignment was constructed in the program PHYML (Guindon and Gascuel 2003), which allows for the fast estimation of large datasets within a maximum likelihood (ML) framework. The best-fit model for the amino acid alignment was determined using ProtTest v1.2.6, which uses the PAL library in conjunction with PHYML to compute the likelihood for each of 64 candidate models of protein evolution. The fit of each of these models to our dataset was then determined using a second order AIC_c framework with sample size equal to the total number of characters (i.e. alignment length). Branch support values were estimated from 100 PHYML bootstrap replicates as bootstrap proportions (BP).

Investigating selective influences

The influence of selective forces on the evolution of invertebrate visual pigments was investigated using a suite of methodologies. First, selection was investigated using nonsynonomous to synonymous substitution rate ratios (d_N/d_S) as calculated in a likelihood framework in the CODEML module of PAML v3.14 (Yang 1997). Selection was determined by using likelihood ratio tests to evaluate nested site-specific models with and without incorporating selection (Yang et al. 2000). For site-specific models, models M1a (nearly neutral) vs. M2a (positive selection) and M7 (beta) vs. M8 (beta& ω) were tested as suggested in the PAML v3.14 manual. For site classes where $\omega > 1$, Bayes empirical Bayes calculations of posterior probabilities are implemented in models M2a and M8 to identify the particular sites under positive selection (Yang et al. in press). All models were run twice with starting omega values of less than and greater than 1 as to test for entrapment in local optima (Yang et al. 2000). Likelihood ratio tests, to determine whether particular models provided a significantly better fit to the data, were performed by comparing the likelihood ratio test statistic ($-2[\ln L_1 - \ln L_2]$) to critical values of the Chi square distribution with the appropriate degrees of freedom (Yang 1997).

d_N/d_S are a common measure of selective pressure in protein coding genes, where $\omega > 1$, $= 0$, or < 1 indicate positive selection, neutral evolution, and purifying selection, respectively. Many advances have been made in the estimation of this ratio that increase its power of detecting selection, including models that are either lineage- or site-specific (Anisimova et al. 2002; Bielawski and Yang 2004; Forsberg and Christiansen 2003; Nielsen and Yang 1998; Yang 1998; Yang and Nielsen 1998; Yang et al. 2000), and site-specific will be used to investigate the presence of positive selection across the entire

invertebrate opsin phylogeny. However, if adaptive evolution occurs at only a few time points and/or affects only a few amino acids, site-specific models may still lack power in detecting positive selection (McClellan et al. 2004). For protein coding genes such as opsin where strong structural and functional constraints lead to a large proportion of invariable residues, d_N/d_S methods are still likely to be very conservative. Furthermore, beyond detecting the presence of positive selection at either specific sites or lineages within a phylogeny, d_N/d_S methods do not provide information on the type of positive selection detected (directional or non-directional, stabilizing or destabilizing), have very little power to detect purifying selection, and offer little insight into how the identified selection affects the overall structure and function of the protein. Recent methods have taken the investigation of selection in protein coding genes further by addressing several of these issues. A different approach to detecting selection in amino acid sequences is to look at the magnitudes of property change of nonsynonymous residues across a phylogeny. Amino acid substitutions have a wide range of effects on a protein depending on the difference in physicochemical properties and location in the protein structure. This approach provides further resolution to differentiating between types of selective pressures with the ability to detect positive and negative stabilizing and destabilizing selection and offers insights into the structural and functional consequences of the identified residues under selection (McClellan et al. 2004). We used the program TreeSAAP v3.2 (Wooley et al. 2003) to test for selection on amino acid properties within our invertebrate opsin dataset. The potential magnitude of change for each property was divided into eight categories, with categories one and two indicating conservative change and seven and eight radical change. Because we are interested in the evolution of

spectral variation, we target sites identified to be under positive destabilizing selection, which we defined *a priori* as properties significantly different from neutral expectations for magnitude categories seven and eight. Using this criterion, identified sites were mapped onto the protein structure to investigate the structural/functional impact of the selection. Within TreeSAAP, thirty one amino acid properties are evaluated across a phylogeny using either the entire dataset or using a sliding window analysis. For our purposes, properties showing significantly more observed than expected numbers of change at the $P \leq 0.001$ level were first identified with an overall analysis of our data. The identified properties were then subjected to a sliding window analysis, investigating varying window sizes (15, 20, 30, and 40 codons in width) to determine the range that maximizes signal. The results of the sliding window analyses were used to identify regions in the protein that differ significantly from a nearly neutral model at a significance level of 0.001. Finally, the particular amino acid residues within each of these regions that contained positive destabilizing selection for each property were identified. Using a high-resolution (2.6 Å) bovine rhodopsin template (1L9H.pdb) from the Protein Data Bank (website) with the program Swiss-PdbViewer v.3.7 (<http://www.expasy.org/spdbv>; Guex and Peitsch 1997) we mapped the identified residues to the opsin protein structure using our alignment as a template. Throughout this paper, domains are labeled as indicated in Figure 1. All references to amino acid residues are given using bovine rhodopsin numbering to make inferences directly comparable to previous studies.

RESULTS

MSP and molecular characterization of novel crustacean opsin sequences

The quantified mysid λ_{\max} ranged from 496 to 520 nm (Figure 2). Most rhabdoms scans fit a Stavenga rhodopsin template. Several scans from *M. relicta* indicated the presence of a porphyropsin pigment (utilizing the A2 chromophore). However, to keep our analyses consistent, we present only the rhodopsin data; the porphyropsin data are presented elsewhere (Jokela-Määttä et al. 2005). The isolated crustacean opsin sequences were 284 amino acids in length, and spanned from 9 amino acids before the start of TMI to the middle of TMVII (AA site 299). Similar to genes previously sequenced from crayfish and cephalopods (Crandall and Cronin 1997; Crandall and Hillis 1997; Morris et al. 1993), none of the isolated crustacean sequences contained any introns. Within each species the clone sequence variability was less than 1%, and therefore clones were merged into a consensus sequence for each species for further analyses. When aligned with other opsin sequences, the new crustacean sequences exhibited the characteristic indel region of invertebrates in CL3 of ~14 amino acids.

Phylogenetic analyses of invertebrate opsins

For each codon position the best-fit models all corresponded to GTR+I+G (AIC_w : position 1 = 0.56; position 2 = 0.92; position 3 = 0.90). For the amino acid data, the best fit model was WAG+G+F ($AIC_{cw} = 0.77$, $\alpha = 1.1$) (Whelan and Goldman 2001). Phylogenetic analyses using these models with either nucleotide or amino acid data produced similar trees, with only minor differences in the relationships of tip taxa within well-supported clades. Both analyses placed the novel crustacean opsin sequences with

previously characterized crustacean sequences in a monophyletic clade sister to the insect MWS/LWS clade (Figure 3). Within each insect spectral clade, species cluster roughly by higher order taxonomy (i.e. Lepidoptera, Hymenoptera, Diptera). In contrast, although the crustacean clade is strongly supported ($BP = 98$; $pP = 1.00$) the currently sampled taxa exhibit very little taxonomic clustering with generally low branch support values. For example, neither the decapod species (crayfish lineage + *H. gammarus*) nor the four mysid species form monophyletic groups. The three genes sequenced from the stomatopod *G. oerstedii* do cluster, although branch support is low for the placement of *G. oerstedii* Rh2 ($BP < 50$, $pP = 0.89$). Interestingly, the *G. oerstedii* visual pigments do not cluster by spectral sensitivity, with strong support for the Rh1 (489 nm) + Rh3 (522 nm) clade ($BP = 100$, $pP = 1.00$). Furthermore, the opsin located in the *G. oerstedii* peripheral ommatidia (Rh2) that is most homologous in structure to other crustacean compound eyes contains the most divergent opsin sequence (40.9-48.3% amino acid difference relative to other crustacean LWS opsins).

The general topology of this tree demonstrates the presences of a monophyletic arthropod LWS clade containing representatives from the Hexapoda, Crustacea, and Chelicerata ($BP=100$; $pP = 1.00$). Curiously, within the Diptera there has been a gene duplication event leading to the diversification of a MWS clade, representatives of which have not yet been found in other insects. A similar situation exists within the Crustacea, with the only sequenced MWS opsin falling outside of the main arthropod LWS clade, indicating an earlier gene duplication event; however, sequence data from other crustaceans are necessary to confirm that this gene is not a copy unique to *H. sanguineus*. Sister to the arthropod LWS/MWS is a monophyletic insect SWS clade ($BP=100$; $pP =$

1.00), containing well-supported lineages representing both blue (BP=96, $pP = 1.00$) and UV (BP=99, $pP = 1.00$) spectral variants. Finally, sister to the arthropod visual pigment clade is the cephalopod lineage (BP=100, $pP = 0.99$).

Selective influences in invertebrate opsins

Analyses of d_N/d_S ratios using PAML did not detect any sites under positive selection, and none of the models incorporating selection parameters contained site classes with $\omega > 1$ or were significantly better than neutral models based on likelihood ratio tests. In contrast, TreeSAAP analyses identified four amino acid properties to be under positive destabilizing selection in our invertebrate opsin dataset ($P < 0.001$): coil tendencies (P_c), compressibility (K^0), power to be at the middle of the α -helix (α_m), and refractive index (μ). Evaluation of various sliding window sizes used indicated that a window of 20 amino acids provided the best signal:noise ratio; window sizes larger than 20 began to match domain size while with smaller windows regions under selection began to reduce to single peaks. Sliding window analyses using a 20 amino acid window size identified that selection on these properties occurred primarily in the transmembrane domains (Figure 4); However, historically (i.e. the number of changes counted across the phylogeny) the distribution of sites under selection within each domains differed between properties. For example, the majority of sites across the phylogeny exhibiting selection for refractive index were found in TMII (30.6%), TMIII (13.9%), CL3 (22.2%), and EL2 (19.4%) while those found for coil tendencies were in CL2 (32.3%), TMV (14.9 %), TMIII (11.5%) and EL2 (10.3%).

Using a high resolution model (1LH9.pdb) the sites identified to be in regions of the protein under destabilizing selection were mapped to the structure of bovine rhodopsin (Figure 5). For K^0 , sites were concentrated on the extracellular end of TM1 and in CL2 and scattered throughout TMIII, TMIV, and TMVI. Remarkably, every single identified amino acid change under selection identified for K^0 was a change to an alanine. Outside of the transmembrane domains, P_c and α_m also identified sites within CL2 and P_c , α_m , and μ sites within EL2. Using features of the program Swiss-PdbViewer v.3.7 (<http://www.expasy.org/spdbv>; Guex and Peitsch 1997) identified residues were evaluated for their proximity to the chromophore binding pocket. Using a conservative distance of 4Å to infer a residue potential interaction with the chromophore, ten sites were isolated from those identified by TreeSAAP analyses: 113 (K^0), 117 (K^0), 118 (K^0), 121 (P_c), 122 (K^0), 186 (α_m), 187 (α_m , μ), 189 (P_c , α_m), 207 (P_c , α_m), and 265 (K^0). These residues are clustered in two areas of the protein in TM3 (113-122) and EL2 (187-207). In addition to these sites, TreeSAAP analyses identified several residues - 90 (K^0 , α_m), 123 (P_c , K^0), 164 (K^0), and 274 (K^0 , α_m) – which are within ~10Å of the chromophore and have been identified in other studies of spectral tuning (Briscoe 2002; Salcedo et al. 2003; Wilkie et al. 2000).

DISCUSSION

Invertebrate opsin evolution

As the available sequence data from insects increased, most investigated opsins fell within three spectral clades – LWS, UVS, and blue sensitive. These genetic clades fit well with the abundance of physiological characterizations of spectral sensitivity in

insects, and led to the hypothesis that the ancestral visual system was trichromatic (Chitkka 1996; Chitkka 1997). However, while the ancestral state may have been trichromatic, the current diversification of opsin genes appears to be more complex. For example, the complement of six visual pigments described from *Drosophila* represent three gene duplications that most likely occurred only in the Dipteran lineage (two leading to the MWS Rh1 and Rh2 and one duplicating a UV gene leading to Rh3 and Rh4). More recent studies have identified similar gene complement expansion in insect LWS opsins within specific lineages (Briscoe 1998; Briscoe 2000) (Hill et al. 2002) and early within insect diversification (Spaethe and Briscoe 2004). While the expression patterns and spectral sensitivities of most of these additional LWS opsin genes have not yet been investigated, recent studies suggest that at least some of these copies may have an extraocular expression, and potentially be involved in circadian regulation systems (Briscoe and White 2005; Shimizu et al. 2001). Given the demonstrated high-copy number of LWS genes in insects, it is not difficult to imagine that a similar pattern of gene expansion will be documented in the crustacean LWS clade once more sequence data are obtained. Even given the low numbers of species and taxonomic groups included in the current study, the low taxonomic clustering observed in the crustacean clade suggests the presence of unidentified gene duplication events occurring at least before the divergence of the Mysida and Decapoda. Interestingly, however, the crustaceans analyzed form two main clades that are well supported by BMCMC analyses ($pP > 0.98$) in which the species are roughly divided by spectral sensitivities. With the exception of *G. oerstedii* Rh1 (489 nm), these two clades are roughly divided by spectral sensitivities, with one shorter wavelength clade (496 – 501 nm) composed of *M. relicta*

sp.IV, *A. grebnitzkii*, and *E. superba* and one longer-wavelength clade (512-533 nm) with the remaining species. A particularly interesting crustacean group to conduct further investigations of opsin evolution will be the stomatopod crustaceans, here represented by the species *N. oerstedi*. Physiologically, stomatopods have been characterized to contain up to 16 different visual pigments that span the ultraviolet to visible spectrum of light (Cronin and Marshall 1989; Cronin and Marshall 2004; Cronin et al. 2000). The three genes included in our analyses represent only a small fraction of this diversity, and indicate that at least some of the similarity in visual pigment λ_{\max} observed within the stomatopod retina is due to convergent evolution rather than symplesiomorphy. Characterization of additional LWS opsin sequences will be required to elucidate the validity of the observed patterns of evolution in crustaceans.

Another feature of invertebrate opsin evolution that is infrequently discussed is the large indel present in the CL3 domain. This indel is consistent in all characterized invertebrate opsins, with the most conserved stretch taking the form of R(E/D)QAKKM(N/G) in arthropods and AAMAKR(L/I)N in cephalopods. As the function of the CL3 domain has been linked to G-protein docking and signal transfer, and a relatively large proportion of sites were identified to be under destabilizing positive selection for K^0 (16.1%) and μ (22.2%) (Figure 4), this conserved region should be of interest to future investigations of opsin functionality between invertebrate and vertebrates.

Selective forces in invertebrate opsins

The ability to study the genetic mechanisms behind phenotypic variation has made opsin a model evolutionary system for studying the selective influences leading to genetic adaptation. The combination of genetic, physiologic, and biochemical studies have shown a direct correlation between the environment and visual pigment spectral sensitivity (Crescitelli et al. 1985; Douglas et al. 1998; Lythgoe 1972; Lythgoe 1980; Partridge 1989; Partridge et al. 1988; Partridge et al. 1989) and demonstrated the effects of single amino acid changes on tuning this spectral sensitivity (Asenjo et al. 1994; Cowing et al. 2002b; Nathans 1990b; Neitz et al. 1991; Wilkie et al. 2000; Yokoyama and Radlwimmer 1999; Yokoyama and Radlwimmer 2001; Yokoyama et al. 2000; Yokoyama and Tada 2003), making opsins one of the clearest, and best-studied, instances where adaptation at the molecular level can be quantified. The observed variation in visual pigment spectral sensitivity is the result of selection and adaptation on the opsin gene and therefore an excellent system for testing methods of detecting the selection known to have acted upon this gene family. In our analyses, estimated d_N/d_S ratios did not detect any evidence of selection; however, these results only confirm that the opsin gene is a generally conservative protein-coding gene that requires the use of alternative criteria to investigate molecular adaptations (McClellan et al. 2004). Using method that evaluates only nonsynonymous replacements with respect to the relative change in a suite of amino acid properties, four properties were identified to be under positive destabilizing selection across the invertebrate phylogeny (P_c , K^0 , α_m , μ). Interestingly, none of the identified properties measure aspects of polarity or charge, two of the properties hypothesized to be important in residues affecting vertebrate spectral tuning (Chan et al.

1992; Nathans 1987; Nathans 1990b; Neitz et al. 1991). Instead, the identified properties relate to structural aspects (i.e. compressibility, coil tendencies, and power to be at the middle of the alpha helix). In particular, K^0 seems to be a significant property with regards to invertebrate opsin evolution, with 186 nonsynonymous replacements identified from across the phylogeny. That all of the identified replacements are to the residue alanine is remarkable. Alanine is one of seven amino acids identified by Gromiha and Ponnuswamy (1993) that determine overall protein compressibility, and, with the exception of proline, has the smallest calculated K^0 ($-25.5 \times 10^{-15} \text{ m}^3 \text{ mol}^{-1} \text{ Pa}^{-1}$). Previous studies have indicated that alanine replacements can stabilize α -helices and have implicated alanine content in helical thermal stability (Argos et al. 1979; Lyu et al. 1990; O'Neil and DeGrado 1990; Padmanabhan et al. 1990; although see Pinker et al. 1993/ for a different opinion; Ptitsyn 1992; Zhang et al. 1991). The large proportion of nonsynonymous replacements in TM1 (10.2%), TMIII (17.7%), TMIV (18.3%), TMV (3.8%), and TMVI (13.4%) indicate that α -helical stability may play an important role in invertebrate opsin evolution. The exclusive use of alanine over proline at these sites is most likely due to the fact that proline residues introduce bends into alpha helices (Riek et al. 2001).

The amino acid sites corresponding to all of the four identified properties are located mainly in transmembrane domains (Figure 4). Based on site-directed mutagenesis studies in vertebrates, transmembrane domains III, VI, and VII are known to have significant interaction with the chromophore and interactions between these three helices are involved in restraining the structure of GPCRs in the inactive, non-signaling state (Chan et al. 1992; Filipek et al. 2003b; Nakayama 1991; Nathans 1990a; Nathans

1990b; Sakmar et al. 1989; Zhukovsky and Oprian 1989). Additionally, the cytoplasmic ends of helices II, VI, and VII move during GPCR activation to create a binding crevice for the G-protein (Filipek et al. 2003b). Unfortunately, the selective influences in TMVII could not be evaluated in this study because many of the available invertebrate opsin sequences are too short. However, that three of the properties (P_c , K^0 , α_m) had significant amounts of detected selection in TMIII indicates that this helix may play a crucial role in the functional diversification of invertebrate opsins.

Using a cut-off distance of 4Å identified ten sites from the TreeSAAP analyses that are potentially affecting the chromophore binding pocket. This distance is more than conservative given that residues at a distance of ~10Å have been identified as regulating chromophore wavelength absorption in other studies of spectral tuning (Briscoe 2002; Salcedo et al. 2003; Wilkie et al. 2000). These ten residues are clustered in the extracellular end of TM3 and EL2. The EL2 folds back into the cavity formed by the membrane-embedded domains, forming part of the chromophore binding pocket and acting as a ‘plug’ preventing solvent access to the Schiff base (Yan et al. 2003). Recent studies have shown this loop is important for the thermal stability of the dark state of rhodopsin (Janz et al. 2003) and in particular E181 in bovine rhodopsin affects the stability and wavelength absorption of metarhodopsin II (Yan et al. 2003), leading to the hypothesis that E181 is a counterion for metarhodopsin I (Teller et al. 2003; Yan et al. 2003). Interestingly, although the vertebrate rhodopsin counterion E113 is not conserved in invertebrates, the E181 residue is, implying that there may be more similarities in metarhodopsin dynamics between invertebrates and vertebrates than in rhodopsin photoactivation. All of the sites identified in this region (186, 187, 189) demonstrated

positive destabilizing selection for α_m , suggesting that similar forces are acting upon adaptation of residues in the EL2 ‘plug’. Future studies investigating the role of this region in metarhodopsin stabilization in invertebrates would be most interesting.

The sites clustered at the extracellular end of TM3 (113, 117, 118, 121, 122) were demonstrated to be under selection for K^0 . These sites are all on turns of the α -helix that face the chromophore binding pocket and several sites have been identified as affecting spectra tuning in vertebrates (122 Yokoyama and Tada 2000; Yokoyama et al. 1999) (Yokoyama and Tada 2003) (Nathans 1990a; Nathans 1990b) (118 Shi et al. 2001). Intriguingly, site 113 (the vertebrate counterion site) was identified to be under positive destabilizing selection in invertebrates. This site is occupied by either a tyrosine in visible- or a phenylalanine in UV-absorbing invertebrate pigments. However, previous studies have shown this site is not used as an invertebrate counterion (Nakagawa et al. 1999) and that the observed amino acid polymorphism is not responsible for the difference in absorption spectra between UV and visible pigments (Salcedo et al. 2003). However, the identification of the site in this study implies that the observed polymorphism serves an as of yet unidentified function in visible versus UV opsins.

In addition to the sites in TMIII and EL2, several residues were identified in our analyses (90, 123, 274) to be under positive destabilizing selection for more than one property that are within ~10Å of the chromophore and have been identified in other studies of spectral tuning (Briscoe 2001; Briscoe 2002; Salcedo et al. 2003; Wilkie et al. 2000). The identification of these sites confirms that amino acids affecting spectral tuning are being selected for impacts on structural aspects of the helices (in this study K^0 , α_m , and P_c). These amino acid properties in particular may affect the internal packing of

the chromophore binding site and thereby affect spectral tuning of the chromophore and signal propagation of the opsin protein.

Table 5-1. Taxonomy, GenBank accession numbers (<http://www.ncbi.nlm.nih.gov/>) for gene sequences, and wavelength of maximal absorbance (λ_{\max}) and references for opsins analyzed in this study. For λ_{\max} values, ‘m’ denotes measurements from males while ‘f’ from females and * indicates averaged values.

Taxon	Accession #	λ_{\max} (nm)	λ_{\max} Reference
Mollusca Rh			
Cephalopoda			
<i>Loligo forbesi</i>	X56788	494	(Morris et al. 1993)
<i>Loligo pealii</i>	AY450853	493	(Brown and Brown 1958; Hubbard and St. George 1958)
<i>Loligo subulata</i>	Z49108	499	(Morris et al. 1993)
<i>Sepia officinalis</i>	AF000947	492	(Brown and Brown 1958)
<i>Todarodes pacificus</i>	X70498	480	(Naito et al. 1981)
<i>Enteroctopus dofleini</i>	X07797	475	(Koutalos et al. 1989)
Arthropoda LWS			
Chelicerata			
<i>Limulus polyphemus</i> – lateral eye	L03781	520	(Hubbard and Wald 1960)
<i>Limulus polyphemus</i> – ocelli	L03782	530	(Nolte and Brown 1972)
Crustacea			
<i>Euphausia superba</i>	this study	487	(Frank and Widder 1999)
<i>Homarus gammarus</i>	this study	515	(Kent 1997)
<i>Cambarellus schufeldtii</i>	AF003544	526	(Crandall and Cronin 1997)
<i>Cambarus ludovicianus</i>	AF003543	529	(Crandall and Cronin 1997)
<i>Orconectes virilis</i>	AF003545	530	(Cronin and Goldsmith 1982; Goldsmith 1978)
<i>Procambarus milleri</i>	AF003546	522	(Crandall and Cronin 1997; Cronin and Goldsmith 1982)
<i>Procambarus clarkii</i>	S53494	533	(Zeiger and Goldsmith 1994)
<i>Archaeomysis grebnitzkii</i>	this study	496	this study
<i>Holmesimysis costata</i>	this study	512	this study
<i>Mysis relicta</i>	this study	501	this study
<i>Neomysis americana</i>	this study	520	this study
<i>Neogonodactylus oerstedi</i> Rh1	(Brown 1996)	489	(Cronin and Marshall 1989)
<i>Neogonodactylus oerstedi</i> Rh2	(Brown 1996)	528	(Cronin and Marshall 1989)
<i>Neogonodactylus oerstedi</i> Rh3	(Brown 1996)	522	(Cronin and Marshall 1989)
Insecta			
<i>Manduca sexta</i>	L78080	520	(White et al. 1983)
<i>Spodoptera exigua</i>	AF385331	515	(Langer et al. 1979)
<i>Galleria mellonella</i>	AF385330	510	(Goldman et al. 1975)
<i>Papilio xuthus</i> Rh1	AB007423	520	(Arikawa et al. 1987; Arikawa et al. 1999; Kitamoto et al. 1998)
<i>Papilio xuthus</i> Rh2	AB007424	520	(Arikawa et al. 1987; Arikawa et al. 1999; Kitamoto et al. 1998)
<i>Papilio xuthus</i> Rh3	AB007425	575	(Arikawa et al. 1987; Arikawa et al. 1999; Kitamoto et al. 1998)
<i>Pieris rapae</i>	AB177984	540	(Ichikawa and Tateda 1982)
<i>Vanessa cardui</i>	AF385333	530	(Briscoe et al. 2003)
<i>Junonia coenia</i>	AF385332	510	(Briscoe 2001)

<i>Heliconius erato</i>	AF126750	570	(Struwe 1972)
<i>Heliconius sara</i>	AF126753	550	(Struwe 1972)
<i>Bicyclus anynana</i>	AF484249	560	(Vanhoutte et al. 2002)
<i>Camponotus abdominalis</i>	U32502	510	(Popp et al. 1996)
<i>Cataglyphis bombycinus</i>	U32501	510	(Popp et al. 1996)
<i>Apis mellifera</i>	U26026	529m, 540f	(Briscoe 2001; Peitsch et al. 1992)
<i>Bombus terrestris</i>	AY485301	529	(Briscoe 2001; Peitsch et al. 1992)
<i>Osmia rufa</i>	AY572828	553	(Briscoe 2001; Peitsch et al. 1992)
<i>Schistocerca gregaria</i>	X80071	520	(Gartner and Towner 1995)
<i>Sphodromantis</i> sp.	X71665	515*	(Rossel 1979)
<hr/>			
Arthropoda MWS			
<hr/>			
Crustacea			
<i>Hemigrapsus sanguineus</i>	D50583, D50584	480	(Sakamoto et al. 1996)
<hr/>			
Insecta			
<i>Drosophila melanogaster</i> Rh6	Z86118	508	(Salcedo et al. 1999)
<i>Drosophila melanogaster</i> Rh1	AH001026	478	(Feiler et al. 1988)
<i>Calliphora erythrocephala</i> Rh1	M58334	490	(Paul et al. 1986)
<i>Drosophila melanogaster</i> Rh2	M12896	420	(Feiler et al. 1988)
<hr/>			
Insecta BLUE			
<hr/>			
<i>Schistocerca gregaria</i>	X80072	430	(Gartner and Towner 1995)
<i>Manduca sexta</i>	AD001674	450	(White et al. 1983)
<i>Papilio xuthus</i> Rh4	AB028217	460	(Arikawa et al. 1987; Eguchi et al. 1982)
<i>Apis mellifera</i>	AF004168	439	(Townson et al. 1998)
<i>Drosophila melanogaster</i> Rh5	U67905	437	(Salcedo et al. 1999)
<hr/>			
Insecta UV			
<hr/>			
<i>Apis mellifera</i>	AF004169	353	(Townson et al. 1998)
<i>Camponotus abdominalis</i>	AF042788	360	(Smith et al. 1997)
<i>Cataglyphis bombycinus</i>	AF042787	360	(Smith et al. 1997)
<i>Manduca sexta</i>	L78081	357	(White et al. 1983)
<i>Papilio xuthus</i> Rh5	AB028218	---	--
<i>Drosophila melanogaster</i> Rh4	AH001040	375	(Feiler et al. 1992)
<i>Drosophila melanogaster</i> Rh3	M17718	345	(Feiler et al. 1992)
<hr/>			
OUTGROUPS			
<hr/>			
<i>Bos taurus</i> rhodopsin	AH001149	---	---
<i>Gallus gallus</i> pineal opsin	U15762	---	---
<i>Anolis carolinensis</i> pineal opsin	AH007737	---	---
<i>Homo sapiens</i> GPR52	NM_005684	---	---
<i>Homo sapiens</i> melatonin receptor 1A	NM_005958	---	---

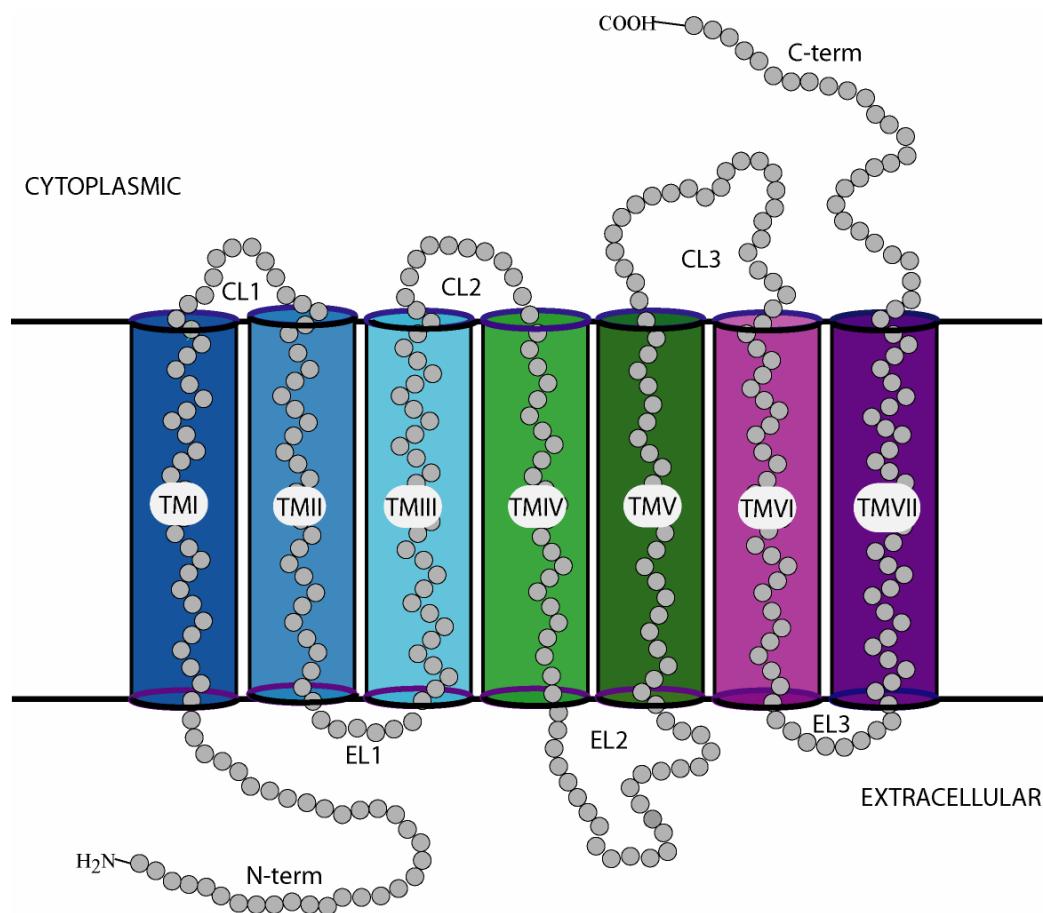


Figure 5-1. Two-dimensional schematic of an opsin protein, with domains labeled as they are referred to in the paper. The opsin gene is composed of three general domains: extracellular, transmembrane, and cytoplasmic. In common with all known G-protein coupled receptors, the opsin protein contains seven transmembrane spanning α -helices, here labeled TM I – TMVII. The extracellular domain contains the N-terminus (N-term) and three inter-helix loops (EL1, EL2, EL3). Similarly, the cytoplasmic domain includes three loops (CL1, CL2, CL3) and the C-terminus (C-term). Other structural features of the opsin protein (i.e. beta sheets and the eighth cytoplasmic alpha-helix) are not discussed in this study and therefore are not illustrated here.

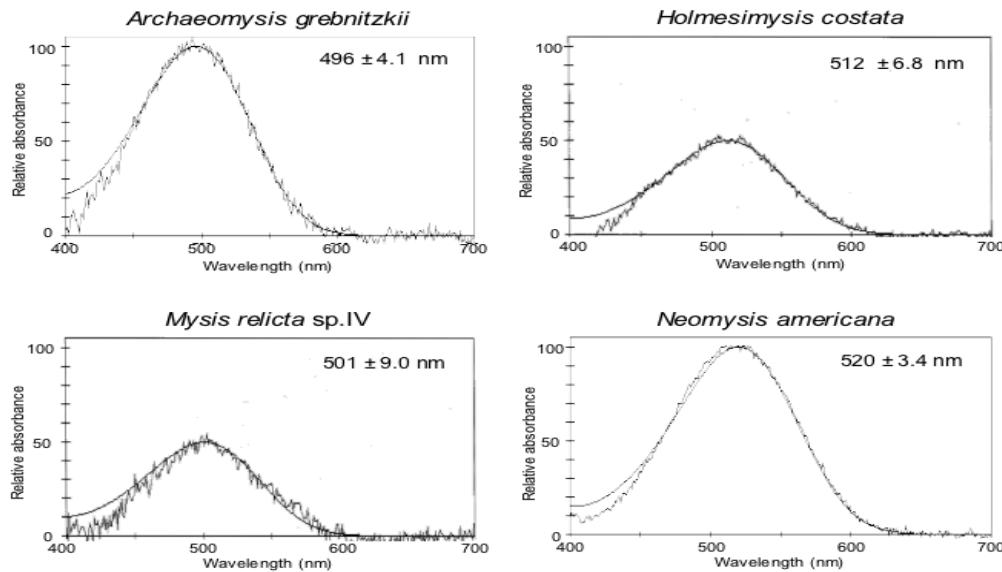


Figure 5-2. Average difference spectra for photobleaching of rhabdoms in retinal sections from four mysid species. Each panel displays the results from the indicated species together with the spectrum of the best-fit rhodopsin template (Stavenga et al. 1993). The wavelengths of the best-fit template \pm standard deviation of the measurements are indicated in each panel. The numbers of rhabdoms included in each average spectrum are as follows: *Archaeomysis grebnitzkii* - 12; *Holmesimysis costata* - 10; *Mysis relicta* sp.IV - 6; *Neomysis americana* - 13.

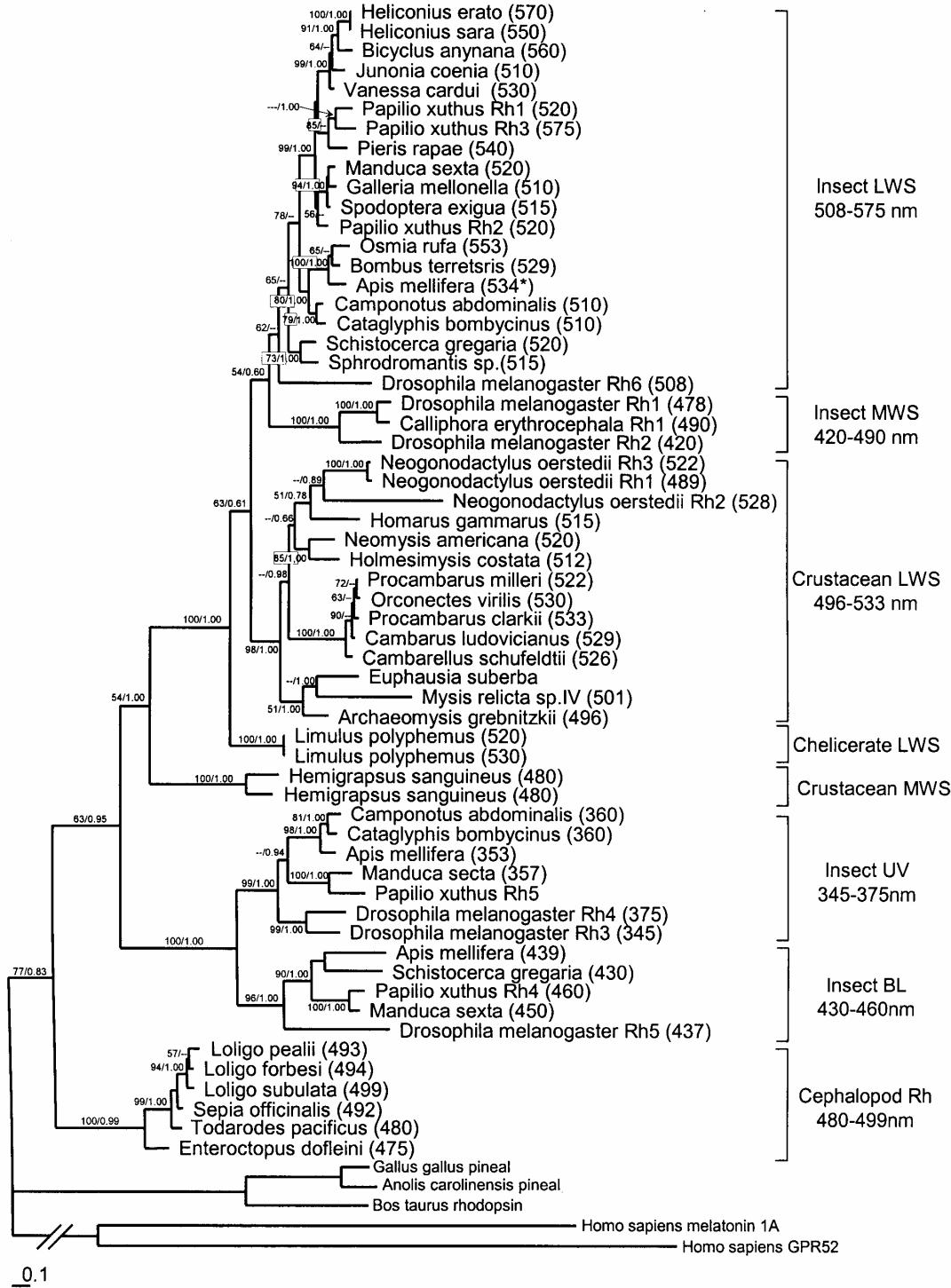


Figure 5-3. Phylogeny of invertebrate opsins based on ML amino acid analyses. Numbers above or below each branch indicate maximum likelihood bootstrap proportions (BP) / BMCMC pP. The λ_{max} for each taxon is given in parentheses after the species name. Spectral clades, with the λ_{max} variation for the represented taxa, are delineated. Accession numbers and references for λ_{max} values are given in Table 1.

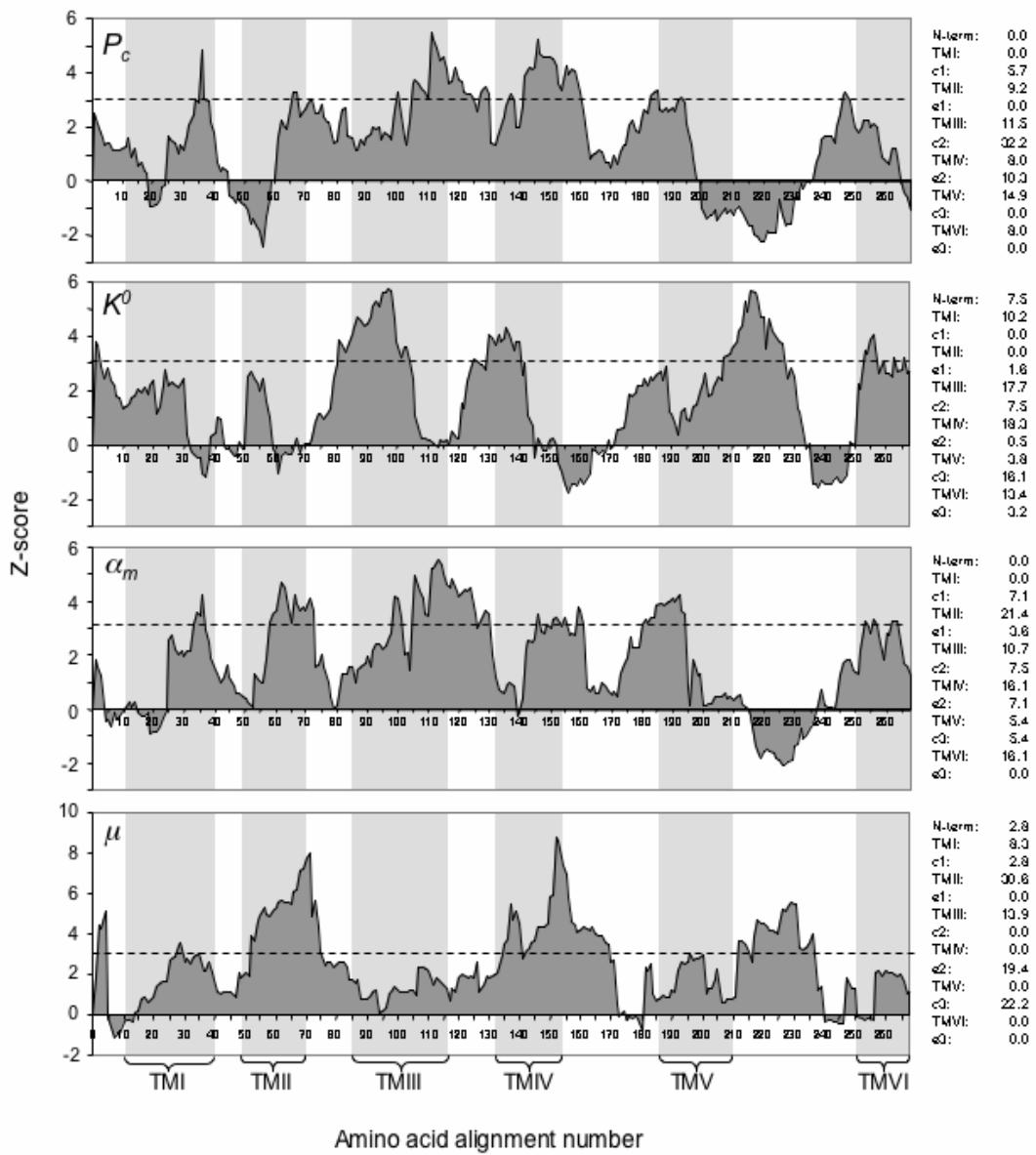


Figure 5-4. Results of TreeSAAP sliding window analyses using a window size of 20 and a step size of 1 amino acid for the four amino acid properties detected to be under positive destabilizing selection: coil tendencies (P_c), compressibility (K^0), power to be at the middle of the α -helix (α_m), and refractive index (μ). The dotted line in each panel indicates significance at the $P = 0.001$ level, with peaks above the line demonstrating regions of the protein under selection; areas under the line are not significantly different from neutral expectations. The transmembrane domains (I-VII) are indicated in each panel by shaded areas. To the right of each panel is the percentage of sites for each property found in the individual domains.

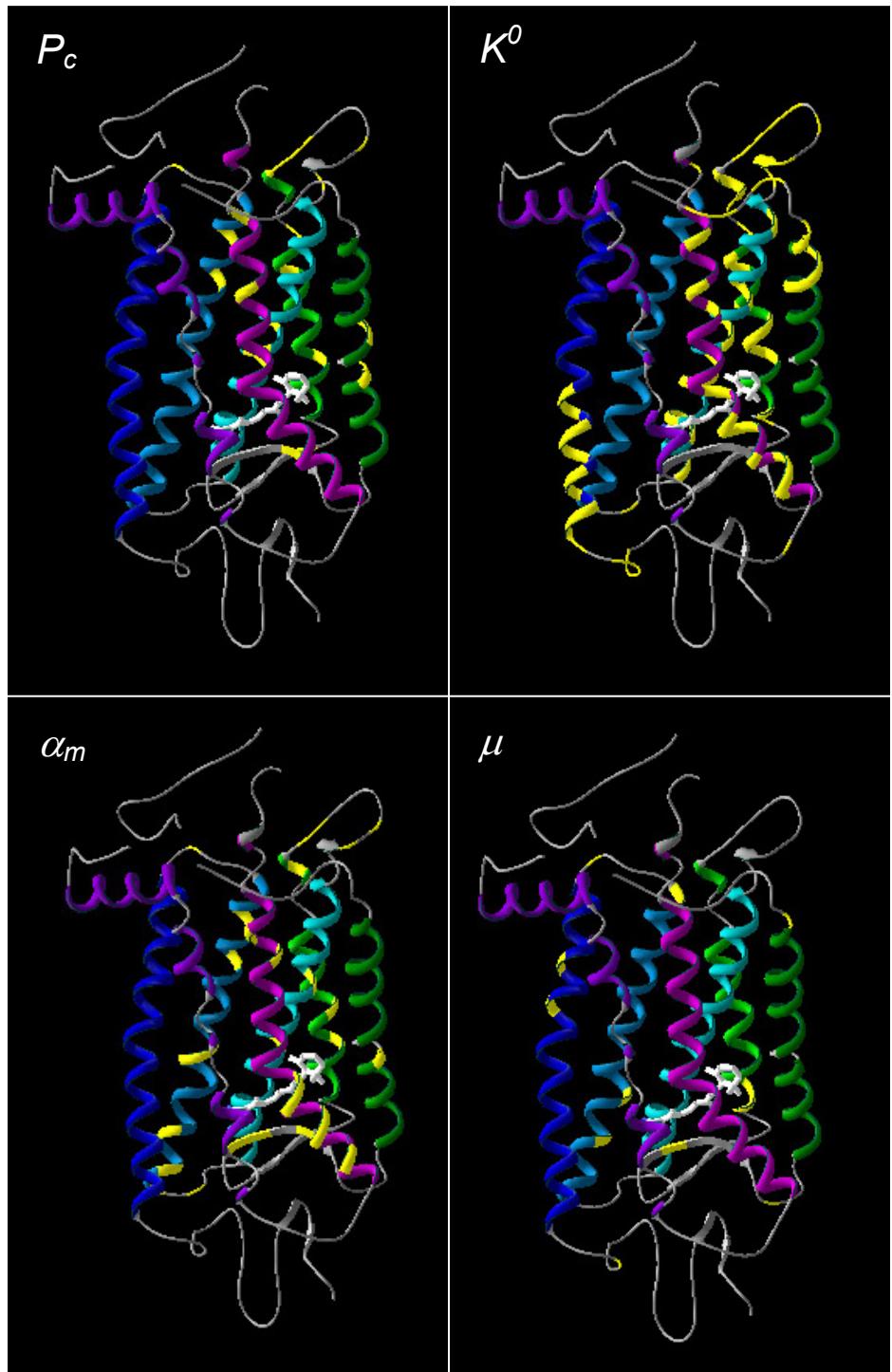


Figure 5-5. Amino acid residues identified by sliding window analyses for each property mapped to the high resolution (2.6Å) bovine rhodopsin structure. Residues with evidence for positive destabilizing selection are shaded yellow. The transmembrane domains are colored as in Figure 1: TM1 = dark blue, TM11 = mid blue, TMIII = light blue, TMIV = light green, TMV = dark green, TMVI = light purple, TMVII = dark purple. The chromophore in each panel is rendered in white. Each panel represents the sites identified for one property: coil tendencies (P_c), compressibility (K^0), power to be at the middle of the α -helix (α_m), and refractive index (μ).

CONCLUSIONS

Investigations of opsin evolution have long been focused on vertebrate systems. While much has been gained in understanding of opsin protein structure and function, few studies have investigated the applicability of this research to invertebrate systems. Those that have been accomplished have generally investigated insect taxa and have not explicitly assessed evolution across the known diversity of invertebrate opsins. Preliminary investigations of the evolutionary history of animal opsins (Chapter 1) indicate that there is much to be learned about the similarities and differences between invertebrate and vertebrate visual pigment systems. This dissertation research focused on opsin evolution in invertebrate systems, particularly from non-insect taxa. Particular conclusions resulting from the studies presented here include:

- Reverse evolution is an influential evolutionary phenomenon, particularly in the visual systems of subterranean organisms which retain functional opsin genes as the structures necessary for vision degenerate (Chapter 2)
- With the exception of a few taxonomically misplaced species, the Gastrosaccinae, Siriellinae, and Heteromysini form strongly supported clades while the Mysinae, Mysini, Leptomysini, and Erythropini are polyphyletic (Chapter 3)
- There is a potentially large number of undescribed Mysidae species (Chapter 3)
- Further molecular studies using expanded sets of genetic markers to refine phylogenetic hypotheses are needed within the Mysidae (Chapter 3)

- Incongruencies between current taxonomic classifications and the presented phylogenetic hypotheses urge taxonomic revision within the Mysidae (Chapter 3)
- Investigations of pattern heterogeneity confirm the assertion that mixed models defined *a priori* miss significant within partition variability (Chapter 3)
- There may be a limit to the ability of new Bayesian models to identify pattern heterogeneity in complex datasets (Chapter 3)
- Although the particular arrangement of most of the decapod orders are still uncertain, there are two well supported nodes corresponding to the suborder Pleoeyemata and the informal ‘Reptantia’ (Chapter 4)
- The divergence time estimates from the decapod lineages illustrate rapid diversification for particular lineages, and place the emergence of the Decapod lineage in the early Devonian (407 MYA) (Chapter 4)
- The structural amino acid properties coil tendencies, compressibility, power to at the middle of the alpha helix, and refractive index were identified to be under positive destabilizing selection in invertebrate opsins (Chapter 5)
- These properties were found at mostly transmembrane sites, and corresponded to protein regions/residues with known functional importance in opsins (Chapter 5)

The research presented here is aimed towards initiating further studies of invertebrate opsins. Hopefully, these studies will provide a sound foundation for future studies of Mysidae and Decapoda phylogenetic systematics and evolutionary studies of crustacean opsins.

REFERENCES CITED

- Abele, L. G. 1991. Comparison of morphological and molecular phylogeny of the Decapoda. Memoirs of the Queensland Museum 31:101-108.
- Abele, L. G., and B. E. Felgenhauer. 1986. Phylogenetic and phenetic relationships among the lower Decapoda. Journal of Crustacean Biology 6:385-400.
- Adachi, J., and M. Hasegawa. 1996. MOLPHY version 2.3: Programs for molecular phylogenetics based in maximum likelihood. Comput. Sci. Monogr. 28:1-150.
- Amati, L., R. M. Feldmann, and J.-P. Zonneveld. 2004. A new family of Triassic lobsters (Decapoda: Astacidea) from British Columbia and its phylogenetic context. Journal of Paleontology 78:150-168.
- Andersson, S. G. E., A. Zomorodipour, J. O. Andersson, T. Sicheritz-Ponten, U. C. M. Alsmark, R. M. Podowski, A. K. Naslund, A. S. Eriksson, H. H. Winkler, and C. G. Kurland. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 333:133-143.
- Andresen, M. C., and A. M. Brown. 1979. Photoresponses of a sensitive extraretinal photoreceptor in *Aplysia*. J. Physiol 287:267-282.
- Anisimova, M., J. P. Bielawski, and Z. Yang. 2002. Accuracy and power of Bayes prediction of amino acid sites under positive selection. Molecular Biology and Evolution 19:950-958.
- Ardendt, D., K. Tessmar-Raible, H. Snyman, A. W. Dorresteijn, and J. Wittbrodt. 2004. Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. Science 306:869-871.
- Argos, P., M. G. Rossmann, U. M. Grau, H. Zuber, G. Frank, and J. D. Tratschin. 1979. Thermal stability and protein structure. Biochemistry 18:5698-5703.
- Arikawa, K., K. Inokuma, and E. Eguchi. 1987. Pentachromatic visual system in a butterfly. Naturwissenschaften 74:297-298.
- Arikawa, K., D. G. W. Scholten, M. Kinoshita, and D. G. Stavenga. 1999. Tuning of photoreceptor spectral sensitivities by red and yellow pigments in the butterfly *Papilio xuthus*. Zool. Sci. 16:17-24.
- Aris-Brosou, S., and Z. Yang. 2002. Effects of models of rate evolution on estimation of divergence dates with special reference to the metazoan 18S ribosomal RNA phylogeny. Systematic Biology 51:703-714.
- Asenjo, A. B., J. Rim, and D. D. Oprian. 1994. Molecular determinants of red/green color discrimination. Neuron 12:1131-1138.
- Babcock, L. E., M. F. Miller, J. L. Isbell, J. W. Collinson, and S. T. Hasiotis. 1998. Paleozoic-Mesozoic crayfish from Antarctica: Earliest evidence of freshwater decapod crustaceans. Geology 26:539-542.
- Bargmann, C. 1998. Neurobiology of the *Caenorhabditis elegans* genome. Science 282:2028-2033.
- Barnes, H., and W. Klepal. 1972. Phototaxis in stage I nauplius larvae of two cirripedes. J. Exp. Mar. Ecol. 10:267-273.
- Beaudet, L., and C. W. Hawryshyn. 1999. Ecological aspects of vertebrate visual ontogeny. Pp. 413-437 in S. N. Archer, M. B. A. Djamgoz, E. R. Loew, J. C. Partridge and S. Vallerga, eds. Adaptive mechanisms in the ecology of vision. Kluwer Academic Publishers, Great Britain.

- Bellingham, J., A. Morris, and D. M. Hunt. 1998. The rhodopsin gene of the cuttlefish *Sepia officinalis*: sequence and spectral tuning. *The Journal of Experimental Biology* 201:2299-2306.
- Bellingham, J., D. J. Wells, and R. G. Foster. 2003. *In silico* characterisation and chromosomal localisation of human *RRH* (peropsin) - implications for opsin evolution. *BMC Genomics* 4, article 3, 8pp.
- Bielawski, J. P., and Z. Yang. 2004. A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. *Journal of Molecular Evolution* 59:121-132.
- Blackshaw, S., and S. Snyder. 1997. Parapinopsin, a novel catfish opsin localized in the parapineal organ, defines a new gene family. *J. Neurosci.* 17:8083-8092.
- Blackshaw, S., and S. H. Snyder. 1999. Encephalopsin: a novel mammalian extraretinal opsin discretely localized in the brain. *J. Neurosci.* 19:3681-3690.
- Boas, J. E. V. 1880. Studier over Decapodernes Slaegtskabsforhold. Danske vidensk. Selsk. Skr., (Nat.) 6:25-210.
- Boas, J. E. V. 1883. Studien über die Verwandtschaftsbeziehungen der Malakostraken. Morph. Jahrb. 8:485-579.
- Bobkova, M., P. Grève, V. B. Meyer-Rochow, and G. Martin. 2003. Description of intracerebral ocelli in two species of North American crayfish: *Orconectes limosus* (Cambaridae) and *Pacifastacus leniusculus* (Astacidae). *Invertebrate Biology* 122:158-165.
- Boden, B. P., E. M. Kampa, and B. C. Abbott. 1961. Photoreception of a planktonic crustacean in relation to light penetration in the sea. Pp. 189-196 in B. Christensen and B. Buchman, eds. *Progress in photobiology*. Elsevier, New York.
- Borowsky, R., and H. Wilkens. 2002. Mapping a cave fish genome: Polygenic systems and regressive evolution. *Journal of Heredity* 93:19-21.
- Borradaile, L. A. 1907. On the classification of the decapod crustaceans. *Annals and Magazine of Natural History* 7:457-486.
- Bowmaker, J. K. 1995. The visual pigments of fish. *Prog. Retinal Eye Res.* 15:1-31.
- Bowmaker, J. K. 1998. Evolution of colour vision in vertebrates. *Eye* 12:541-547.
- Bowmaker, J. K., S. Astell, D. M. Hunt, and J. D. Mollon. 1991. Photosensitive and photostable pigments in the retinae of old world monkeys. *Journal of Experimental Biology* 156:1-19.
- Bowmaker, J. K., V. I. Govardovskii, S. A. Shukolyukov, L. V. Zueva, D. M. Hunt, V. G. Sideleva, and O. G. Smirnova. 1994. Visual pigments and the photic environment: the cottoid fish of Lake Baikal. *Visual Research* 34:591-605.
- Bowman, T. E., and L. G. Abele. 1982. Classification of the recent Crustacea. Pp. 1-27 in L. G. Abele, ed. *The biology of the Crustacea*. Academic Press, New York.
- Bravi, S., and A. Garassino. 1998. 'Plattenkalk' of the Lower Cretaceous (Albian) of Petina, in the Alburni Mounts (Campania, S. Italy), and its decapod crustacean assemblage. *Atti Societa italiana Scienze naturali Museo civico Storia naturale Milano* 138:89-118.
- Briggs, M. H. 1961. Visual pigments of grapsiid crabs. *Nature* 190:784-786.
- Briscoe, A. D. 1998. Molecular diversity of visual pigments in the butterfly *Papilio glaucus*. *Naturwissenschaften* 85:33-35.

- Briscoe, A. D. 2000. Six opsins from the butterfly *Papilio glaucus*: Molecular phylogenetic evidence for paralogous origins of red-sensitive visual pigments in insects. *Journal of Molecular Evolution* 51:110-121.
- Briscoe, A. D. 2001. Functional diversification of Lepidopteran opsins following gene duplication. *Molecular Biology and Evolution* 18:2270-2279.
- Briscoe, A. D. 2002. Homology modeling suggests a functional role for parallel amino acid substitutions between bee and butterfly red- and green-sensitive opsins. *Molecular Biology and Evolution* 19:983-986.
- Briscoe, A. D., G. D. Bernard, A. S. Szeto, L. M. Nagy, and R. H. White. 2003. Not all butterfly eyes are created equal: rhodopsin absorption spectra, molecular identification, and localization of ultraviolet-, blue-, and green-sensitive rhodopsin-encoding mRNAs in the retina of *Vanessa cardui*. *The Journal of Comparative Neurology* 458:334-349.
- Briscoe, A. D., and L. Chitkra. 2001. The evolution of color vision in insects. *Annual Review of Entomology* 46:471-510.
- Briscoe, A. D., and R. H. White. 2005. Adult stemmata of the butterfly *Vanessa cardui* express UV and green opsin mRNAs. *Cell Tissue Res* 319:175-179.
- Britt, S. G., R. Feiler, K. Kirschfeld, and C. S. Zuker. 1993. Spectral tuning of rhodopsin and metarhodopsin in vivo. *Neuron* 11:29-39.
- Broek, J. V. 2001. Insect G protein-coupled receptors and signal transduction. *Archives of Insect Biochemistry and Physiology* 48:1-12.
- Brown, A. J. H. 1996. Isolation and characterisation of visual pigment genes from the stomatopod crustacean *Gonodactylus oerstedii*. Pp. 192. University of Sussex.
- Brown, A. M., and H. M. Brown. 1973. Light response of a giant *Aplysia* neuron. *J Gen Physiol* 62:239-254.
- Brown, P. K., and P. S. Brown. 1958. Visual pigments of the octopus and cuttlefish. *Nature* 182:1288-1290.
- Bruno, M. S., S. N. Barnes, and T. H. Goldsmith. 1977. The visual pigment and visual cycle of the lobster, *Homarus*. *Journal of Comparative Physiology A* 120:123-142.
- Bruno, M. S., and T. H. Goldsmith. 1974. Rhodopsin of the blue crab *Callinectes*: evidence for absorption differences in vitro, in vivo. *Vision Research* 14:653-658.
- Bruno, M. S., and D. Kennedy. 1962. Spectral sensitivity of photoreceptor neurons in sixth ganglion of crayfish. *Comp. Biochem. Physiol.* 6:41-&.
- Buckley, T. R. 2002. Model misspecification and probabilistic test of topology: Evidence from empirical data sets. *Systematic Biology* 51:509-523.
- Burch, C. L., and L. Chao. 1999. Evolution by small steps and rugged landscapes in the RNA virus φ6. *Genetics* 151:921-927.
- Burkenroad, M. D. 1963. The Evolution of the Eucarida (Crustacea, Eumalacostraca), in Relation to the Fossil Record. *Tulane Stud. Geol.* 2:1-17.
- Burkenroad, M. D. 1981. The higher taxonomy and evolution of decapoda (Crustacea). *Trans. San Diego Soc. Nat. Hist.* 19:251-268.
- Carleton, K. L., and T. D. Kocher. 2001. Cone opsin genes of African cichlid fishes: Tuning spectral sensitivity by differential gene expression. *Molecular Biology and Evolution* 18:1540-1550.

- Carulli, J. P., D.-M. Chen, W. S. Stark, and D. L. Hartl. 1994. Phylogeny and physiology of *Drosophila* opsins. *Journal of Molecular Evolution* 38:250-262.
- Carulli, J. P., and D. L. Hartl. 1992. Variable rates of evolution among *Drosophila* opsin genes. *Genetics* 132:193-204.
- Casanova, B., L. De Jong, and X. Moreau. 2002. Carapace and mandibles ontogeny in the Dendrobranchiata (Decapoda), Euphausiacea, and Mysidacea (Crustacea): a phylogenetic interest. *Candian Journal of Zoology* 80:296-306.
- Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* 17:540-552.
- Chan, T., M. Lee, and T. P. Sakmar. 1992. Introduction of hydroxyl bearing residues causes bathochromic spectral shifts in rhodopsin. *J. Biol. Chem.* 267:9478-9480.
- Chang, B. S. W., D. Ayers, W. C. Smith, and N. E. Pierce. 1996. Cloning of the gene encoding honeybee long-wavelength rhodopsin: A new class of insect visual pigments. *Gene* 173:215-219.
- Chang, B. S. W., K. A. Crandall, J. P. Carulli, and D. L. Hartl. 1995. Opsin phylogeny and evolution: A model for blue shift in wavelength regulation. *Molecular Phylogenetics and Evolution* 4:31-43.
- Chase, M. R., P. R. Bennett, and R. H. White. 1997. Three opsin-encoding cDNAs from the compound eye of *Manduca sexta*. *The Journal of Experimental Biology* 200:2469-2478.
- Cheroske, A. G., T. W. Cronin, and R. L. Caldwell. 2003. Adaptive color vision in *Pullosquilla litoralis* (Stomatopoda, Lysiosquilloidea) associated with spectral and intensity changes in light environment. *The Journal of Experimental Biology* 206:373-379.
- Chiao, C. C., T. W. Cronin, and N. J. Marshall. 2000. Eye design and color signaling in a stomatopod crustacean, *Gonodactylus smithii*. *Brain, Behavior and Evolution* 56:107-122.
- Chitkka, L. 1996. Does bee color vision predate the evolution of flower color? *Naturwissenschaften* 83:136-138.
- Chitkka, L. 1997. Bee color vision is optimal for coding flower color, but flower colors are not optimal for being coded - why? *Israel Journal of Plant Sciences* 45:115-127.
- Christoffersen, M. L. 1988. Phylogenetic systematics of the Eucarida (Crustacea Malacostraca). *Revista Brasileira de Zoologia* 5:325-351.
- Colgar, D. J., A. McLauchlan, G. D. F. Wilson, S. P. Livingston, G. D. Edgecombe, J. Macaranas, G. Cassis, and M. R. Gray. 1998. Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Aust. J. Zool.* 46:419-437.
- Colless, D. H. 1980. Congruence between morphometric and allozyme data for *Menidia* species: A reappraisal. *Systematic Zoology* 29:288-299.
- Cowing, J. A., S. Poopalasundaram, S. E. Wilkie, J. K. Bowmaker, and D. M. Hunt. 2002a. Spectral tuning and evolution of short wave-sensitive cone pigments in cottoid fish from Lake Baikal. *Biochemistry* 41:6019-6025.
- Cowing, J. A., S. Poopalasundaram, S. E. Wilkie, P. R. Robinson, J. K. Bowmaker, and D. M. Hunt. 2002b. The molecular mechanism for the spectral shifts between vertebrate ultraviolet- and violet- sensitive cone visual pigments. *Biochemical Journal* 367:129-135.

- Crandall, K. A. 2002. Encyclopedia of Evolution. Oxford University Press.
- Crandall, K. A., and T. W. Cronin. 1997. The molecular evolution of visual pigments of freshwater crayfishes (Decapoda: Cambaridae). *Journal of Molecular Evolution* 45:524-534.
- Crandall, K. A., and J. F. Fitzpatrick Jr. 1996. Crayfish molecular systematics: Using a combination of procedures to estimate phylogeny. *Systematic Biology* 45:1-26.
- Crandall, K. A., D. J. Harris, and J. W. J. Fetzner. 2000. The monophyletic origin of freshwater crayfish estimated from nuclear and mitochondrial DNA sequences. *Proceedings of the Royal Society of London B* 267:1679-1686.
- Crandall, K. A., and D. M. Hillis. 1997. Rhodopsin evolution in the dark. *Nature* 387:667-668.
- Crescitelli, F., M. McFall-Ngai, and J. Horwitz. 1985. The visual pigment sensitivity hypothesis: further evidence from fishes of varying habitats. *Journal of Comparative Physiology A* 157:323-333.
- Crill, W. D., H. A. Wichman, and J. J. Bull. 2000. Evolutionary reversals during viral adaptation to alternating hosts. *Genetics* 154:27-37.
- Cronin, T. W. 1985. The visual pigment of a stomatopod crustacean, *Squilla empusa*. *Journal of Comparative Physiology A* 156:679-687.
- Cronin, T. W. 1986. Optical design and evolutionary adaptation in crustacean compound eyes. *Journal of Crustacean Biology* 6:1-23.
- Cronin, T. W. 2005. Invertebrate vision in water. In: *Invertebrate Vision* (E. Warrant and D. E. Nilsson, eds.). Cambridge University Press, Cambridge UK (in press).
- Cronin, T. W., R. L. Caldwell, and M. V. Erdmann. 2002. Tuning of photoreceptor function in three mantis shrimp species that inhabit a range of depths. I. Visual pigments. *Journal of Comparative Physiology A* 188:179-186.
- Cronin, T. W., R. L. Caldwell, and N. J. Marshall. 2001. Tunable colour vision in a mantis shrimp. *Nature* 411:547.
- Cronin, T. W., and R. B. Forward. 1988. The visual pigments of crabs - I. Spectral characteristics. *Journal of Comparative Physiology A* 162:463-478.
- Cronin, T. W., and T. M. Frank. 1996. A short-wavelength photoreceptor class in deep-sea shrimp. *Proceedings of the Royal Society of London B* 263:861-865.
- Cronin, T. W., and T. H. Goldsmith. 1982. Photosensitivity spectrum of crayfish rhodopsin measured using fluorescence of metarhodopsin. *J. Gen. Physiol.* 79:313-332.
- Cronin, T. W., and R. N. Jinks. 2001. Ontogeny of vision in marine crustaceans. *Amer. Zool.* 41:1098-1107.
- Cronin, T. W., and N. J. Marshall. 1989. Multiple spectral classes of photoreceptors in the retinae of gonodactyloid stomatopod crustaceans. *Journal of Comparative Physiology A* 166:267-275.
- Cronin, T. W., and N. J. Marshall. 2004. The unique visual world of mantis shrimps. Pp. 239-268 in F. Prete, ed. *Complex Worlds from Simple Nervous Systems*. MIT Press, Cambridge MA.
- Cronin, T. W., N. J. Marshall, and R. L. Caldwell. 1993. Photoreceptor spectral diversity in the retinas of squilloid and lysiosquilloid stomatopod crustaceans. *Journal of Comparative Physiology A* 172:339-350.

- Cronin, T. W., N. J. Marshall, and R. L. Caldwell. 1994a. The retinas of mantis shrimp from low-light environments (Crustacea; Stomatopoda; Gonodactylidae). *Journal of Comparative Physiology A* 174:607-619.
- Cronin, T. W., N. J. Marshall, and R. L. Caldwell. 1996. Visual pigment diversity in two genera of mantis shrimps implies rapid evolution (Crustacea: Stomatopoda). *Journal of Comparative Physiology A* 179:371-384.
- Cronin, T. W., N. J. Marshall, and R. L. Caldwell. 2000. Spectral tuning and the visual ecology of mantis shrimps. *Philosophical Transactions of the Royal Society Series B* 355:1263-1267.
- Cronin, T. W., N. J. Marshall, R. L. Caldwell, and D. Pales. 1995. Compound eyes and ocular pigments of crustacean larvae (Stomatopoda and Decapod Brachyura). *Mar Fresh Behav Physiol* 26:219-231.
- Cronin, T. W., N. J. Marshall, R. L. Caldwell, and N. Shashar. 1994b. Specialization of retinal function in the compound eyes of mantis shrimps. *Vision Research* 34:2639-2656.
- Cronin, T. W., N. J. Marshall, C. A. Quinn, and C. A. King. 1994c. Ultraviolet photoreceptors in mantis shrimp. *Vision Research* 34:1443-1452.
- Culver, D. C., L. L. Master, M. C. Christman, and H. H. I. Hobbs. 2000. Obligate cave fauna of the 48 contiguous United States. *Conservation Biology* 14:386-401.
- Culver, D. C., and H. Wilkens. 2000. Critical review of the relevant theories of the evolution of subterranean animals. Pp. 381-398 in H. Wilkens, D. C. Culver and W. F. Humphreys, eds. *Subterranean Ecosystems*. Elsevier Science, Amsterdam, The Netherlands.
- Cummins, D. R., D.-M. Chen, and T. H. Goldsmith. 1984. Spectral sensitivity of the spiny lobster *Panulirus argus*. *Biol. Bull.* 16:269-276.
- Cummins, D. R., and T. H. Goldsmith. 1981. Cellular identification of the violet receptor in the crayfish eye. *Journal of Comparative Physiology* 42:199-202.
- Cunningham, C. W., K. E. Omland, and T. H. Oakley. 1998. Reconstructing ancestral character states: a critical reappraisal. *Trends in Ecology and Evolution* 13:361-366.
- Darwin, C. 1859. *On the Origin of Species*. John Murray.
- De Jong, L., S. Jean, and X. Moreau. 2004. The digestive system of *Euphausia krohnii* (Crustacea, Euphausiacea): functional morphology and phylogenetic relevance. *Cahiers de biologie marine* 45:71-82.
- De Jong-Moreau, L., and J.-P. Casanova. 2001. The foreguts of the primitive families of the Mysida (Crustacea, Peracarida): a transitional link between those of the Lophogastrida (Crustacea, Mysidacea) and the most evolved Mysida. *Acta Zoologica* 82:137-147.
- Denys, C. J., and P. K. Brown. 1982. Euphausiid visual pigments. The rhodopsins of *Euphausia superba* and *Meganyctiphanes norvegicus*. *J. Gen. Physiol.* 80:451-472.
- DePamphilis, C. W., N. D. Young, and A. D. Wolfe. 1997. Evolution of plastid gene *rps2* in a lineage of hemiparasitic and holoparasitic plants: Many losses of photosynthesis and complex patterns of rate variation. *Proceedings of the National Academy of Science USA* 94:7367-7372.

- Dixon, C. J., S. T. Ahyong, and F. R. Schram. 2003. A new hypothesis of decapod phylogeny. *Crustaceana* 76:935-975.
- Douglas, R. H., and N. J. Marshall. 1999. A review of vertebrate and invertebrate ocular filters. Pp. 95-162 in S. N. Archer, E. R. Loew, J. C. Partridge and S. Vallarta, eds. *Adaptive mechanisms in the ecology of vision*. Kluwer Academic Publishers, Dordrecht.
- Douglas, R. H., and J. C. Partridge. 1997. On the visual pigments of deep-sea fish. *Journal of Fish Biology* 50:68-85.
- Douglas, R. H., J. C. Partridge, and N. J. Marshall. 1998. The eyes of deep-sea fish I. Lens pigmentation, tapeta, visual pigments. *Prog. Retinal Eye Res.* 17:597-636.
- Dowling, T. E., D. P. Martasian, and W. R. Jeffery. 2002. Evidence for multiple genetic forms with similar eyeless phenotypes in the blind cavefish, *Astyanax mexicanus*. *Molecular Biology and Evolution* 19:446-455.
- Eguchi, E., K. Watanabe, T. Hariyama, and K. Yamamoto. 1982. A comparison of electrophysiologically determined spectral responses in 35 species of Lepidoptera. *Journal of Insect Physiology* 28:675-682.
- Elena, S. F., and R. E. Lenski. 2001. Epistasis between new mutations and genetic background and a test of genetic canalization. *Evolution* 55:1746-1752.
- Emlen, D. J. 2001. Costs and the diversification of exaggerated animal structures. *Science* 291:1534-1536.
- Fasick, J. I., M. L. Applebury, and D. D. Oprian. 2002. Spectral tuning in the mammalian shortwavelength sensitive cone pigments. *Biochemistry* 41:6860-6865.
- Fasick, J. I., and P. R. Robinson. 1998. Mechanism of spectral tuning in the dolphin visual pigments. *Biochemistry* 37:432-438.
- Feiler, R., R. Bjornson, K. Kirschfeld, D. Mismer, G. M. Rubin, D. P. Smith, M. Socolich, and C. S. Zuker. 1992. Ectopic expression of ultraviolet-rhodopsins in the blue photoreceptor cells of *Drosophila*: visual physiology and photochemistry of transgenic animals. *Journal of Neuroscience* 12:3862-3868.
- Feiler, R., W. A. Harris, K. Kirschfeld, C. Wehrhahn, and C. S. Zuker. 1988. Targeted misexpression of a *Drosophila* opsin gene leads to altered visual function. *Nature* 333:737-741.
- Feldmann, R. M. 2003. The Decapoda: New initiatives and novel approaches. *Journal of Paleontology* 77:1021-1038.
- Feldmann, R. M., F. J. Vega, S. P. Applegate, and G. A. Bishop. 1998. Early Cretaceous arthropods from the Tlayua formation at Tepexi de Rodriguez, Puebla, Mexico. *Journal of Paleontology* 72:79-90.
- Felgenhauer, B. E., and L. G. Abele. 1983. Phylogenetic relationships among shrimp-like decapods. *Crustacean Issues* 1:291-311.
- Felsenstein, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution* 17:368-376.
- Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791.
- Felsenstein, J. 2003. *Phylib* (Phylogenetic Inference Package). distributed by the author, Dept. of Genetics, University of Washington, Seattle.

- Fenton, G. E. 1985. Ecology and taxonomy of mysids (Mysidacea: Crustacea). Pp. 412. Dept. of Zoology. University of Tasmania.
- Fernandez, H. R. 1965. A survey of the visual pigments of decapod crustaceans of S Florida. University of Miami, Miami, Florida.
- Fernandez, H. R. 1973. Spectral sensitivity and visual pigment of the compound eye of the galatheid crab *Pleuroncodes planipes*. Marine Biology 20:148-153.
- Filipek, S., R. Stenkamp, D. C. Teller, and K. Palczewski. 2003a. G protein-coupled receptor rhodopsin: A prospectus. Ann. Rev. Physiol. 65:851-79.
- Filipek, S., D. C. Teller, K. Palczewski, and R. Stenkamp. 2003b. The crystallographic model of rhodopsin and its use in studies of other G protein-coupled receptors. Annual Review of Biophys. Biomol. Struct. 32:375-397.
- Fisher, L. R. 1967. Pigments of euphausiid eyes. J. Mar. Biol. Assoc. India:1074-1080.
- Fisher, L. R., and E. H. Goldie. 1959. The eye pigments of a euphausiid crustacean, *Meganyctiphanes norvegica* (M. Sars). Proc. XV Intern. Congr. Zool., London 533-535.
- Fisher, L. R., and E. H. Goldie. 1961. Pigments of compound eyes. Pp. 153-154 in B. Christensen and B. Buchman, eds. Progress in photobiology. Elsevier, Amsterdam.
- Flower, D. R. 1999. Modelling G-protein-coupled receptors for drug design. Biochem Biophys Acta 1422:207-234.
- Forsberg, R., and F. B. Christiansen. 2003. A codon-based model of host-specific selection in parasites, with an application to the influenza A virus. Molecular Biology and Evolution 20:1252-1259.
- Forsell, J., P. Ekström, I. Novales Flamarique, and B. Holmqvist. 2001. Expression of pineal ultraviolet- and green-like opsins in the pineal organ and retina of teleosts. Journal of Experimental Biology 204:2517-2525.
- Frank, T. M., and J. F. Case. 1988a. Visual spectral sensitivities of bioluminescent deep-sea crustaceans. Biol. Bull. 175:261-273.
- Frank, T. M., and J. F. Case. 1988b. Visual spectral sensitivity of the bioluminescent deep-sea mysid, *Gnathophausia ingens*. Biological Bulletin 175:274-283.
- Frank, T. M., and E. A. Widder. 1999. Comparative study of the spectral sensitivities of mesopelagic crustaceans. Journal of Comparative Physiology a-Neuroethology Sensory Neural and Behavioral Physiology 185:255-265.
- Fredriksson, R., M. C. Lagerström, L.-G. Lundin, and H. B. Schiöth. 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Molecular Pharmacology 63:1256-1272.
- Frélon-Raimond, M., V. B. Meyer-Rochow, A. Ugolini, and G. Martin. 2002. First cytological description of intracerebral ocelli in an amphipod: the extraretinal photoreceptors of the sandhopper *Talitrus saltator* (Crustacea, Amphipoda). Invert Neurosci 121:1-12.
- Fryxell, K. J., and E. M. Meyerowitz. 1991. The evolution of rhodopsins and neurotransmitter receptors. Journal of Molecular Evolution 33:367-378.
- Gal, G., E. R. Loew, L. G. Rudstam, and A. M. Mohammadian. 1999. Light and diel vertical migration: spectral sensitivity and light avoidance by *Mysis relicta*. Can. J. Fish. Aquat. Sci. 56:311-322.

- Garassino, A., and S. Bravi. 2003. *Palaemon antonellae* new species (Crustacea, Decapoda, Caridea) from the Lower Cretaceous 'Platydolomite' of Profeti (Caserta, Italy). *Journal of Paleontology* 77:589-592.
- Gartner, W., and P. Towner. 1995. Invertebrate visual pigments. *Photochem Photobiol* 62:1-16.
- Gatesy, J., P. O'Grady, and R. H. Baker. 1999. Corroboration among data sets in simultaneous analysis: Hidden support for phylogenetic relationships among higher level artiodactyl taxa. *Cladistics* 15:271-314.
- Gil, R., B. Sabater-Munoz, A. Latorre, F. J. Silva, and A. Moya. 2002. Extreme genome reduction in *Buchneria* spp.: Toward the minimal genome needed for symbiotic life. *Proceedings of the National Academy of Science USA* 99:4454-4458.
- Giribet, G., S. Richter, G. D. Edgecombe, and W. C. Wheeler. 2004. The position of crustaceans within the Arthropoda - Evidence from nine molecular loci and morphology. *Crustacean Issues* 16:307-352.
- Glaessner, M. F. 1969. Decapoda. Pp. R651 in R. C. Moore, ed. *Treatise on Invertebrate Paleontology*, Part R. Arthropoda 4. University of Kansas, Geological Society of America, Lawerence.
- Goldman, L. J., S. N. Barnes, and T. H. Goldsmith. 1975. Microspectrophotometry of rhodopsin and metarhodopsin in the moth *Galleria*. *Journal of General Physiology* 66:383-404.
- Goldman, N., J. P. Anderson, and A. G. Rodrigo. 2000. Likelihood-based tests of topologies in phylogenetics. *Systematic Biology* 49:652-670.
- Goldsmith, T. H. 1972. The natural history of invertebrate visual pigments. Pp. 727-742 in H. J. A. Dartnall, ed. *Handbook of sensory physiology*. Springer, Berlin.
- Goldsmith, T. H. 1978. The spectral absorption of crayfish rhabdoms: pigment, photoproduct and pH sensitivity. *Vision Research* 18:463-473.
- Goldsmith, T. H., A. E. Dizon, and H. R. Fernandez. 1968. Microspectrophotometry of photoreceptor organelles from eyes of the prawn *Palaemonetes*. *Science* 161:468-470.
- Gotow, T. 1989. Photoresponses of an extraocular photoreceptor associated with a decrease in membrane conductance in a opisthobranch mollusc. *Brain Res* 479:120-129.
- Graur, D., and W. Martin. 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in Genetics* 20:80-86.
- Gromiha, M. M., and P. K. Ponnuswamy. 1993. Relationship between amino acid properties and protein compressibility. *Journal of Theoretical Biology* 165:87-100.
- Guex, N., and M. C. Peitsch. 1997. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis* 18:2714-2723.
- Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* 52:696-704.
- Hafner, G. S., R. L. Martin, and T. R. Tokarski. 2003. Photopigment gene expression and rhabdom formation in the crayfish (*Procambarus clarkii*). *Cell Tissue Res.* 311:99-105.
- Hall, M. D., M. A. Hoon, N. J. Ryba, J. D. Pottinger, J. N. Keen, H. R. Saibil, and J. B. Findlay. 1991. Molecular cloning and primary structure of squid (*Loligo forbesi*)

- rhodopsin, a phospholipase C-directed G-protein linked receptor. Biochemical Journal 274:35-40.
- Hamacher, H., and H. Stieve. 1984. Spectral properties of the rhodopsin-system of the crayfish *Astacus leptodactylus*. Photochem Photobiol 39:379-390.
- Hamacher, K. J., and K. D. Kohl. 1981. Spectroscopical studies of the *Astacus* visual pigment. Biophysics of Structure and Mechanism 7:338.
- Hao, W., and H. K. Fong. 1996. Blue and ultraviolet light-absorbing opsin from the retinal pigment epithelium. Biochemistry 35:6251-6256.
- Hariyama, T., K. Ozaki, F. Tokunaga, and Y. Tsukahara. 1993. Primary structure of crayfish visual pigment deduced from cDNA. FEBS Letters 315:287-292.
- Hariyama, T., and Y. Tsukahara. 1993. Spectral responses, including a UV-sensitive cell type, in the eye of the isopod *Ligia exotica*. Naturwissenschaften 80:233-235.
- Hasegawa, M. 1990. Mitochondrial DNA evolution in primates: Transition rates has been extremely low in the lemur. Journal of Molecular Evolution 31:113-121.
- Hasiotis, S. T. 2002. Where is the fossil evidence for Gondwanan crayfish? Gondwana Research: International Geoscience Journal 5:872-878.
- Hasiotis, S. T., and C. E. Mitchell. 1993. A comparison of crayfish burrow morphologies: Triassic and Holocene fossil, paleo- and neo-ichnological evidence, and the identification of their burrowing signatures. Ichnos 2:291-314.
- Hattar, S., H.-W. Liao, M. Takao, D. M. Berson, and K.-W. Yau. 2002. Melanopsin-containing retinal ganglion cells: architecture, projections, and intrinsic photosensitivity. Science 295:1065-1070.
- Hays, D., and T. H. Goldsmith. 1969. Microspectrophotometry of the visual pigment of the spider crab *Libinia emarginata*. Zeitschrift fur Vergleichende Physiologie 65:218-232.
- Hertel, H. 1972. Aspekte zur photorezeption von *Artemia salina*. Verh. dtsch. Ges. Zool. Erlangen
- Higgs, P. G. 1998. Compensatory neutral mutations and the evolution of RNA. Genetica 102:91-101.
- Hill, C. A., A. N. Fox, R. J. Pitts, L. B. Kent, P. L. Tan, M. A. Chrystal, A. Cravchik, F. H. Collins, H. M. Robertson, and L. J. Zwiebel. 2002. G protein-coupled receptors in *Anopheles gambiae*. Science 298:176-178.
- Hiller-Adams, P., E. A. Widder, and J. F. Case. 1988. The visual pigments of four deep-sea crustacean species. Journal of Comparative Physiology A 163:63-72.
- Hillis, D. M., and T. M. Dixon. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. Q. Rev. Biol. 66:411-453.
- Hillman, P., F. A. Dodge, S. Hochstein, B. W. Knight, and B. Minke. 1973. Rapid dark recovery of the invertebrate early receptor potential. Journal of General Physiology 62:77-86.
- Holmquist, C. 1981. The genus *Acanthomysis* Czerniavsky, 1882 (Crustacea, Mysidacea). Zool. Jb. Syst. 108:386-415.
- Horch, K., M. Salmon, and R. Forward. 2002. Evidence for a two pigment visual system in the fiddler crab, *Uca thayeri*. Journal of Comparative Physiology A 188:493-499.
- Hubbard, R., and R. C. C. St. George. 1958. The rhodopsin system of the squid. J. Gen. Physiol. 41:510-528.

- Hubbard, R., and G. Wald. 1960. *Nature* 186:212-215.
- Hueffer, K., J. S. L. Parker, W. S. Weichert, R. E. Geisel, J. Y. Sgro, and C. R. Parrish. 2003. The natural host range shift and subsequent evolution of canine parvovirus resulted from virus-specific canine transferrin receptor. *Journal of Virology* 77:1718-1726.
- Huelsenbeck, J. P., and K. A. Crandall. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review of Ecology and Systematics* 28:437-466.
- Huelsenbeck, J. P., and N. S. Imennov. 2002. Geographic origin of human mitochondrial DNA: Accomodating phylogenetic uncertainty and model comparison. *Systematic Biology* 51:673-688.
- Huelsenbeck, J. P., B. Larget, R. E. Miller, and F. Ronquist. 2002. Potential applications and pitfalls of Bayesian inference phylogeny. *Systematic Biology* 51:673-688.
- Hunt, D. M., J. A. Cowing, S. E. Wilkie, J. W. L. Parry, S. Poopalasundaram, and J. K. Bowmaker. 2004. Divergent mechanisms for tuning of shortwave sensitive visual pigments in vertebrates. *Photochemical and Photobiological Sciences: and International Journal* 3:713-720.
- Huxley, T. H. 1880. The crayfish. C. Kegan Paul and Co., London.
- Ichikawa, T., and H. Tateda. 1982. Distribution of color receptors in the larval eyes of four species of Lepidoptera. *Journal of Comparative Physiology A* 149:317-324.
- Ikeda, Y., M. Mochizuki, R. Naito, K. Nakamura, T. Miyazawa, T. Mikami, and E. Takahashi. 2000. Predominance of canine parvovirus (CPV) in unvaccinated cat populations and emergence of new antigenic types of CPV in cats. *Virology* 278:13-19.
- Ikeda, Y., K. Nakamura, T. Miyazawa, E. Takahashi, and M. Mochizuki. 2002. Feline host range of canine parvovirus: recent emergence of new antigenic types in cats. *Emerging Infectious Diseases* 8:341-346.
- Imaiizumi, R. 1938. Fossil crayfishes from Jehol. *Science Reports of the Tokyo Imperial University, Sendal, Japan, Second Series* 19:173-179.
- Janz, J. M., J. F. Fay, and D. L. Farrens. 2003. Stability of dark state rhodopsin is mediated by a conserved ion pair in intradiscal loop E-2. *J. Biol. Chem.* 278:16982-16991.
- Jarman, S. N., S. Nicol, N. G. Elliott, and A. McMinn. 2000. 28S evolution in the Eumalacostraca and the phylogenetic position of krill. *Molecular Phylogenetics and Evolution* 17:26-36.
- Jeffery, W. R. 2001. Cavefish as a model system in evolutionary developmental biology. *Developmental Biology* 231:1-12.
- Jeffery, W. R., and D. P. Martasian. 1998. Evolution of eye regression in the cavefish *Astyanax*: Apoptosis and the Pax-6 gene. *American Zoologist* 38:685-696.
- Jinks, R. N., T. L. Markley, E. E. Taylor, G. Perovich, A. I. Dittel, C. E. Epifanio, and T. W. Cronin. 2002. Adaptive visual metamorphosis in a deep-sea hydrothermal vent crab. *Nature* 420:68-70.
- Johnson, M. L., E. Gaten, and P. M. J. Shelton. 2002. Spectral sensitivities of five marine decapod crustaceans and a review of spectral sensitivity variation in relation to habitat. *J. Mar. Biol. Ass. UK* 82:835-842.

- Jokela-Määttä, M., J. Pahlberg, M. Lindström, P. Zak, M. L. Porter, M. Ostrovskii, T. Cronin, and K. Donner. 2005. Visual pigment absorbance and spectral sensitivity of *Mysis relicta* (Crustacea, Mysidae) in different light environments. Journal of Comparative Physiology A. in press.
- Jutte, P. A., T. W. Cronin, and R. L. Caldwell. 1998a. Retinal function in the planktonic larvae of two species of *Pullosquilla* a lysiosquilloid stomatopod crustacean. Journal of Experimental Biology 201:2481-2487.
- Jutte, P. A., T. W. Cronin, and R. L. Caldwell. 1998b. Vision in two sympatric species of *Pullosquilla* (Stomatopoda, Lysiosquilloidea) living in different depth ranges. Marine Freshwater Behavior and Physiology 31:231-250.
- Kampa, E. M. 1955. Euphausiopsin, a new photosensitive pigment from the eyes of euphausiid crustaceans. Nature 175:996-998.
- Kass, R. E., and A. E. Raftery. 1995. Bayes factors. J. Am. Stat. Assoc. 90:773-795.
- Kawamura, S., and S. Yokoyama. 1996. Molecular characterization of the pigeon P-opsin gene. Gene 182:213-214.
- Kawamura, S., and S. Yokoyama. 1997. Expression of visual and nonvisual opsins in American chameleon. Vision Research 37:1867-1871.
- Kent, J. 1997. The visual pigments of deep-sea crustaceans. University of Bristol, UK.
- Kim, W., and L. G. Abele. 1990. Molecular phylogeny of selected decapod crustaceans based on 18s rRNA nucleotide sequences. Journal of Crustacean Biology 10
- King, R. C., and W. D. Stansfield. 1997. A dictionary of genetics. Oxford University Press
- Kishino, H., J. L. Thorne, and W. J. Bruno. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Molecular Biology and Evolution 18:352-361.
- Kitamoto, J. K., K. Sakamoto, K. Ozaki, Y. Mishina, and K. Arikawa. 1998. Two visual pigments in a single photoreceptor cell: identification and histological localization of three mRNAs encoding visual pigment opsins in the retina of the butterfly *Papilio xuthus*. Journal of Experimental Biology 201:1255-1261.
- Kito, Y., T. Seki, T. Suzuki, and I. Uchiyama. 1986. 3-Dehydroretinal in the eye of a bioluminescent squid, *Watasenia scintillans*. Vision Research 26:275-279.
- Kobusch, W. 1998. The foregut of the Mysida (Crustacea, Peracarida) and its phylogenetic relevance. Philosophical Transactions of the Royal Society of London B 353:559-581.
- Kojima, D., H. Mano, and Y. Fukada. 2000. Vertebrate ancient-long opsin: A green-sensitive photoreceptor molecule present in zebrafish deep brain retinal horizontal cells. J. Neurosci. 20:2845-2851.
- Koutalos, Y., T. G. Ebrey, M. Tsuda, K. Odashima, T. Lien, M. H. Park, N. Shimizu, F. Derguini, K. Nakanishi, H. R. Gilson, and B. Honig. 1989. Regeneration of bovine and octopus opsins in situ with natural and artificial retinals. Biochemistry 28:2732-2739.
- Kroeze, W. K., D. J. Sheffler, and B. L. Roth. 2003. G-protein-coupled receptors at a glance. Journal of Cell Science 116:4867-4869.
- Lall, A. B., and T. W. Cronin. 1987. Spectral sensitivity of the compound eyes in the purple land crab, *Geocarcinus lateralis*. Biological Bulletin 173:398-406.

- Land, M. F. 1981. Optics and vision in invertebrates. Pp. 471-592 in H. Autrum, ed. *Handbook of sensory physiology*. Springer-Verlag, New York.
- Langer, H., B. Haumann, and C. C. Meinecki. 1979. Tetrachromatic visual system in the moth *Spodoptera exempta* (Insecta: Noctuidae). *Journal of Comparative Physiology A* 129:235-239.
- Larimer, J. L. 1966. A functional caudal photoreceptor in blind cavernicolous crayfish. *Nature* 210:204-205.
- Lee, M. S. Y. 1999. Molecular clock calibrations and metazoan diversity. *Journal of Molecular Evolution* 49:385.
- Leggett, L. M. W. 1979. A retinal substrate for colour discrimination in crabs. *Journal of Comparative Physiology A* 133:159-166.
- Levin, B. R., V. Perrot, and N. Walker. 2000. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. *Genetics* 154:985-997.
- Lindsay, S. M., T. M. Frank, J. Kent, J. C. Partridge, and M. I. Latz. 1999. Spectral sensitivity of vision and bioluminescence in the midwater shrimp *Sergestes similis*. *Biol. Bull.* 197:348-360.
- Lindström, M. 2000. Eye function of Mysidacea (Crustacea) in the northern Baltic Sea. *Journal of Experimental and Marine Biology and Ecoloty* 246:85-101.
- Lipetz, L. E., and T. W. Cronin. 1988. Application of an invariant spectral form to the visual pigments of crustaceans: Implications regarding the binding of the chromophore. *Vision Research* 28:1083-1093.
- Lohan, A. J., and K. H. Wolfe. 1998. A subset of conserved tRNA genes in plastid DNA of nongreen plants. *Genetics* 150:425-433.
- Lowe, E. R. 1976. Light and photoreceptor degeneration in the Norway lobster, *Nephrops norvegicus*. *Proceedings of the Royal Society of London B* 193:31-44.
- Lythgoe, J. N. 1972. The adaptation of visual pigments to the photic environment. Pp. 566-603 in H. J. A. Dartnall, ed. *Handbook of sensory physiology*. Springer, Berlin.
- Lythgoe, J. N. 1980. Vision in fish: ecological adaptations. Pp. 431-445 in M. A. Ali, ed. *Environmental physiology of fishes*. Plenum Press, New York.
- Lyu, P. C., M. I. Liff, L. A. Marky, and N. R. Kallenback. 1990. Side chain contributions to the stability of alpha helical structure in peptides. *Science* 250:669-673.
- Maddison, W. P., and D. R. Maddison. 2003. *MacClade: Analysis of phylogeny and character evolution*. Sinauer Associates, Sunderland, MA.
- Maisnier-Patin, S., O. G. Berg, L. Liljas, and D. I. Andersson. 2002. Compensatory adaptation to the deleterious effect of antibiotic resistance in *Salmonella typhimurium*. *Molecular Microbiology* 46:355-366.
- Marshall, N. J., T. W. Cronin, and T. N. Frank. 2003. Visual adaptations in Crustaceans: Chromatic, developmental, and temporal aspects. Pp. 343-372 in S. Collins and N. J. Marshall, eds. *Sensory Processing in the Aquatic Environment*. Springer-Verlag, Berlin.
- Marshall, N. J., J. Kent, and T. W. Cronin. 1999. Visual adaptations in crustaceans: Spectral sensitivity in diverse habitats. Pp. 285-327 in S. N. Archer, M. B. A.

- Djamgoz, E. R. Loew, J. C. Partridge and S. Vallerga, eds. Adaptive mechanisms in the ecology of vision. Kluwer Academic, Great Britain.
- Marshall, N. J., and M. F. Land. 1993. Some optical features of the eyes of stomatopods. II. Ommatidial design, sensitivity and habitat. *Journal of Comparative Physiology A* 173:583-594.
- Martin, G. 1971. Etude préliminaire d'une structure photosensible dans la région centrale du protocérébron de *Porcellio dilatatus* (Brandt), Crustacé, Isopode, Oniscoïde. *Comp. rend. Acad. Sc. Paris* 272:269-271.
- Martin, G. 1976. Mise en évidence et étude ultrastructurale des ocelles médians chez les crustacés isopodes. *Ann. Sci. Nat. Zool. Paris* 18:405-436.
- Martin, G., and M. I. Mote. 1982. Color receptors in marine crustaceans: a second spectral class of retinular cell in the compound eyes of *Callinectes* and *Carcinus*. *Journal of Comparative Physiology* 145:549-554.
- Martin, J. W., and G. E. Davis. 2001. An updated classification of the recent Crustacea. *Natural History Museum of Los Angeles Cy. Sci. Ser.* 39:1-124.
- Martinez-Torres, D., C. Buades, A. Latorre, and A. Moya. 2001. Molecular systematics of aphids and their primary endosymbionts. *Molecular Phylogenetics and Evolution* 20:437-449.
- Massey, R. C., A. Buckling, and S. J. Peacock. 2001. Phenotypic switching of antibiotic resistance circumvents permanent costs in *Staphylococcus aureus*. *Current Biology* 11:1810-1814.
- Matsui, S., M. Seidou, I. Uchiyama, N. Sekiya, K. Hiraki, K. Yoshihara, and Y. Kito. 1988. 4-Hydroxyretinal, a new visual pigment chromophore found in bioluminescent squid, *Watasenia scintillans*. *Biochimica et Biophysica Acta* 966:370-374.
- Max, M., P. J. McKinnon, K. J. Seidenman, R. K. Barrett, M. L. Applebury, J. S. Takahashi, and R. F. Margolskee. 1995. Pineal opsin: a nonvisual opsin expressed in chick pineal. *Science* 267:1502-1506.
- McClellan, D. A. 2000. The codon-degeneracy model of molecular evolution. *Journal of Molecular Evolution* 50:131-140.
- McClellan, D. A., and K. G. McCracken. 2001. Estimating the influence of selection on the variable amino acid sites of the cytochrome *b* protein functional domains. *Molecular Biology and Evolution* 18:917-925.
- McClellan, D. A., E. J. Palfreyman, M. J. Smith, J. L. Moss, R. G. Christensen, and J. K. Sailsbury. 2004. Physicochemical evolution and molecular adaptation of the Cetacean and Artiodactyl cytochrome *b* proteins. *Molecular Biology and Evolution* 22:437-455.
- McFarland, W. N., and E. R. Loew. 1994. Ultraviolet visual pigments in marine fishes of the family pomacentridae. *Vision Research* 34:1393-1396.
- Meland, K. 2004. Species diversity and phylogeny of the deep-sea genus *Pseudomma* (Crustacea: Mysida). *Zootaxa* 649:1-30.
- Meland, K., and E. Willassen. 2004. Molecular phylogeny and biogeography of the genus *Pseudomma* (Peracarida: Mysida). *Journal of Crustacean Biology* 24:541-557.
- Merbs, S. L., and J. Nathans. 1992. Absorption spectra of hybrid pigments responsible for anomalous color vision. *Science* 258:464-466.

- Meyer-Rochow, V. B. 2001. The crustacean eye: Dark/Light adaptation, polarization, sensitivity, flicker fusion frequency, and photoreceptor damage. *Zoological Science* 18:1175-1197.
- Meyer-Rochow, V. B., and K. M. Tiang. 1984. The eye of *Jasus edwardsii* (Crustacea, Decapoda): electrophysiology, histology, and behaviour. *Zoologica* 45:1-61.
- Mezzetti, M. C., and F. Scapini. 1995. Aspects of spectral sensitivity in *Talitrus saltator* (Montagu) (Crustacea-Amphipoda). *Marine and Freshwater Behaviour and Physiology* 26:35-45.
- Minjuan, C., and L. Shujun. 1990. The spectral sensitivity of receptor system in the compound eye of *Penaeus penicillatus*. *Oceanologia et Limnologia Sinica* 21:160-165.
- Minke, B., and K. Kirschfeld. 1978. Microspectrophotometric evidence for two photo-interconvertible states of visual pigment in the barnacle lateral eye. *Journal of General Physiology* 71:37-45.
- Mirzadegan, T., G. Benko, S. Filipek, and K. Palczewski. 2003. *Biochemistry* 42:2759-2767.
- Montell, C., K. Jones, C. S. Zuker, and G. M. Rubin. 1987. A second opsin gene expressed in the ultraviolet-sensitive R7 photoreceptor cells of *Drosophila melanogaster*. *J. Neurosci.* 7:1558-1566.
- Moran, N. A., M. A. Munson, P. Baumann, and H. Ishikawa. 1993. A molecular clock in endosymbiotic bacteria is calibrated using the insect host. *Proceedings of the Royal Society of London B* 253:167-171.
- Moran, N. A., and J. J. Wernegreen. 2000. Lifestyle evolution in symbiotic bacteria: insights from genomics. *Trends in Ecology and Evolution* 15:321-326.
- Morris, A., J. K. Bowmaker, and D. M. Hunt. 1993. The molecular basis of a spectral shift in the rhodopsins of two species of squid from different photic environments. *Proceedings of the Royal Society of London B* 254:233-240.
- Moutsaki, P., J. Bellingham, B. G. Soni, Z. K. David-Gray, and R. G. Foster. 2000. Sequence, genomic structure and tissue expression of carp (*Cyprinus carpio* L.) vertebrate ancient (VA) opsin. *FEBS Letters* 473:316-322.
- Muachline, J. 1980. The biology of mysids and euphausiids. Academic Press, London.
- Muller, G. B. 2002. *Encyclopedia of Evolution*. Oxford University Press
- Müller, O. F. 1776. *Zoologiae Danicae prodromus, seu animalium Daniae et Norvegiae indigenarum: characteres, nomina, et synonyma imprimis popularium.* (Typis Hallageriis, Havniae [Copenhagen]) i-xxxii:1-282.
- Naito, T., K. Nashima-Hayama, K. Ohtsu, and Y. Kito. 1981. Photoreactions of cephalopod rhodopsin. *Vision Research* 21:935-941.
- Nakagawa, M., T. Iwasa, S. Kikkawa, M. Tsuda, and T. G. Ebrey. 1999. How vertebrate and invertebrate visual pigments differ in their mechanism of photoactivation. *Proceedings of the National Academy of Science USA* 96:6189-6192.
- Nakayama, T. A. 1991. Mapping of the amino acids in membrane-embedded helices that interact with the retinal chromophore in bovine rhodopsin. *J. Biol. Chem.* 266:4269-4275.
- Nathans, J. 1987. Molecular biology of visual pigments. *Ann. Rev Neurosci* 10:163-94.

- Nathans, J. 1990a. Determinants of visual pigment absorbance: identification of the retinylidene Schiff's base counterion in bovine rhodopsin. *Biochemistry* 29:9746-9752.
- Nathans, J. 1990b. Determinants of visual pigment absorbance: role of charged amino acids in putative transmembrane segments. *Biochemistry* 29:937-942.
- Nathans, J., C. M. Davenport, I. H. Maumenee, R. A. Lewis, J. F. Hejtmancik, M. Litt, E. Lovrien, R. Weleber, B. Bachynski, F. Zwas, R. Klingaman, and G. Fishman. 1989. Molecular genetics of human blue cone monochromacy. *Science* 245:831-838.
- Nathans, J., and D. S. Hogness. 1983. Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin. *Cell* 34:807-814.
- Neitz, M., J. Neitz, and G. H. Jacobs. 1991. Spectral tuning of pigments underlying red-green colour vision. *Science* 252:971-974.
- Nevo, E. 1999. Mosaic evolution of subterranean mammals: Regression, progression, and global convergence. Oxford University Press, Oxford.
- Nielsen, H. S., M. B. Oleksiewicz, R. Forsberg, T. Stadejek, A. Botner, and T. Storgaard. 2001. Reversion of a live porcine reproductive and respiratory syndrome virus vaccine investigated by parallel mutations. *Journal of General Virology* 82:1263-1272.
- Nielsen, R., and Z. Yang. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 134:1271-1276.
- Nilsson, D.-E. 1989. Optics and evolution of the compound eye. Pp. 30-73 in D. G. Stavenga and R. C. Hardie, eds. *Facets of vision*. Springer-Verlag, Berlin.
- Nolte, J., and J. E. Brown. 1972. *J. Gen. Physiol.* 59:186-200.
- Norell, M. 1992. Taxic origin and temporal diversity: The effect of phylogeny. Pp. 89-118 in N. J. Novacek and Q. D. Wheeler, eds. *Extinction and phylogeny*. Columbia University Press, New York.
- Nylander, J. A. A., F. Ronquist, J. P. Huelsenbeck, and J. L. Nieves-Aldrey. 2004. Bayesian phylogenetic analysis of combined data. *Systematic Biology* 53:47-67.
- Oakley, T. H., and D. R. Huber. 2004. Differential expression of duplicated opsin genes in two eye types of ostracod crustaceans. *Journal of Molecular Evolution* 59:239-249.
- Ochman, H., and N. A. Moran. 2001. Genes lost and genes found: Evolution of bacterial pathogenesis and symbiosis. *Science* 292:1096-1098.
- Okano, T., T. Yoshizawa, and Y. Fukada. 1994. Pinopsin is a chicken pineal photoreceptive molecule. *Nature* 372:94-97.
- Oliveira, L., T. Hulsen, D. L. Hulsik, A. C. M. Paiva, and G. Vriend. 2004. Heavier-than-air flying machines are impossible. *FEBS Letters* 564:269-273.
- Omland, K. E. 1997. Examining two standard assumptions of ancestral reconstructions: repeated loss of dichromatism in dabbling ducks (Anatini). *Evolution* 51:1636-1646.
- Omland, K. E., and S. M. Lanyon. 2000. Reconstructing plumage evolution in orioles (*Icterus*): Repeated convergence and reversal in patterns. *Evolution* 54:2119-2133.

- O'Neil, K. T., and W. F. DeGrado. 1990. A thermodynamic scale for the helix forming tendencies of the commonly occurring amino acids. *Science* 250:646-650.
- O'Tousa, J. E., W. Baehr, R. L. Martin, J. Hirsh, W. L. Pak, and M. L. Applebury. 1985. The *Drosophila ninaE* gene encodes an opsin. *Cell* 40:839-850.
- Padmanabhan, S., S. Marqusee, T. Ridgeway, T. M. Laue, and R. L. Baldwin. 1990. Relative helix forming tendencies of nonpolar amino acids. *Nature* 344:268-270.
- Pagel, M., and A. Meade. 2004. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence of character-state data. *Systematic Biology* 53:571-581.
- Palczewski, K., T. Kumashita, T. Hori, and e. a. co-authors). 2000. Crystal structure of rhodopsin: a G protein coupled receptor. *Science* 289:739-745.
- Pandey, S., J. C. Blanks, C. Spee, M. Jiang, and H. K. Fong. 1994. Cytoplasmic retinal localization of an evolutionary homolog of the visual pigments. *Exp Eye Res* 58:605-613.
- Partridge, J. C. 1989. Visual pigments of deep-sea fishes: ecophysiology and molecular biology. *Prog. Underwater Sci.* 14:17-31.
- Partridge, J. C., S. N. Archer, and J. N. Lythgoe. 1988. Visual pigments in the individual rods of deep-sea fishes. *Journal of Comparative Physiology A* 162:543-550.
- Partridge, J. C., S. N. Archer, and J. Van Oostrum. 1992. Single and multiple visual pigments in deep-sea fishes. *J Mar Biol Ass UK* 72:113-130.
- Partridge, J. C., J. Shand, S. N. Archer, J. N. Lythgoe, and W. A. H. M. van Groningen-Luyben. 1989. Interspecific variation in the visual pigments of deep-sea fishes. *Journal of Comparative Physiology A* 164:513-529.
- Paul, R., A. Steiner, and R. Gemperlein. 1986. Spectral sensitivity of *Calliphora erythrocephala* and other insect species studien with Fourier interferometric stimulation (FIS). *Journal of Comparative Physiology A* 158:669-680.
- Peitsch, D., A. Feitz, H. Hertel, J. de Souza, D. F. Ventura, and R. Menzel. 1992. The spectral input systems of hymenopteran insects and their receptor-based colour vision. *Journal of Comparative Physiology A* 170:23-40.
- Pérez-Losada, M., J. T. Hoeg, and K. A. Crandall. 2004. Unraveling the evolutionary radiation of the Thoracican barnacles using molecular and morphological evidence: A comparison of several divergence time estimation approaches. *Systematic Biology* 53:244-264.
- Philp, A. R., J. M. Garcia-Fernandez, B. G. Soni, R. J. Lucas, J. Bellingham, and R. G. Foster. 2000. Vertebrate ancient (VA) opsin and extraretinal photoreception in the Atlantic salmon (*Salmo salar*). *Journal of Experimental Biology* 203:1925-1936.
- Pichaud, F., A. Briscoe, and C. Desplan. 1999. Evolution of color vision. *Current Opinion in Neurobiology* 9:622-627.
- Pinker, R. J., L. Lin, G. D. Rose, and N. R. Kallenbach. 1993. Effects of alanine substitution in alpha-helices of sperm whale myoglobin on protein stability. *Protein Science* 2:1099-1105.
- Pol, D. 2004. Empirical problems of the hierarchical likelihood ratio test for model selection. *Systematic Biology* 53:949-962.

- Popp, M. P., R. Grisshammer, P. A. Hargrave, and W. C. Smith. 1996. Ant opsins: sequences from the Saharan silver ant and the carpenter ant. *Invert Neurosci* 1:323-329.
- Posada, D., and T. R. Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of Akaike information criterion and Bayesian approaches over Likelihood ratio tests. *Systematic Biology* 53:793-808.
- Posada, D., and K. A. Crandall. 1998. Modeltest: Testing the model of DNA substitution. *Bioinformatics* 14:817-818.
- Posada, D. P., and K. A. Crandall. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution* 16:37-44.
- Ptitsyn, O. B. 1992. Secondary structure formation and stability. *Current Opinion in Structural Biology* 2:13-20.
- Rambaut, A., and L. Broham. 1998. Estimating divergence data from molecular sequences. *Molecular Biology and Evolution* 15:442-448.
- Rambaut, A., and A. J. Drummond. 2003. Tracer v1.2, available from <http://evolve.zoo.ox.ac.uk/>.
- Remerie, T., B. Bulckaen, J. Calderon, T. Deprez, J. Mees, J. Vanfleteren, A. Vanreusel, A. Vierstraete, M. Vincx, K. J. Wittmann, and T. Wooldridge. 2004. Phylogenetic relationships within the Mysidae (Crustacea, Peracarida, Mysida) based on nuclear 18S ribosomal RNA sequences. *Molecular Phylogenetics and Evolution* 32:770-777.
- Richter, S., and G. Scholtz. 2001. Phylogenetic analysis of the Malacostraca (Crustacea). *J. Zool. Syst. Evol. Research* 39:113-136.
- Riek, R. P., I. Rigoutsos, J. Novotny, and R. M. Graham. 2001. Non-alpha-helical elements modulate polytopic membrane protein architecture. *Journal of Molecular Evolution* 306:349-362.
- Robas, N., M. O'Reilly, S. Katugampola, and M. Fidock. 2003. Maximizing serendipity: strategies for identifying ligands for orphan G-protein-coupled receptors. *Current Opinion in Pharmacology* 3:121-126.
- Rode, A. L., and L. E. Babcock. 2003. Phylogeny of fossil and extant freshwater crayfish and some closely related Nephropid lobsters. *Journal of Crustacean Biology* 23:418-435.
- Rokyta, D., M. R. Badgett, I. J. Molneux, and J. J. Bull. 2002. Experimental genomic evolution: Extensive compensation for loss of DNA ligase activity in a virus. *Molecular Biology and Evolution* 19:230-238.
- Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.
- Rossel, S. 1979. Regional differences in photoreceptor performance in the eye of the praying mantis (*Tenodera australasiae*). *Journal of Comparative Physiology* 131:95-112.
- Saiki, R., D. H. Gelfand, S. Stofell, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487-491.
- Sakamoto, K., O. Hisatomi, F. Tokunaga, and E. Eguchi. 1996. Two opsins from the compound eye of the crab *Hemigrapsus sanguineus*. *The Journal of Experimental Biology* 199:441-450.

- Sakmar, T. P., R. R. Franke, and H. G. Khorana. 1989. Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. *Proceedings of the National Academy of Science USA* 86:8309-8313.
- Salcedo, E., A. Huber, S. Henrich, L. V. Chadwell, W.-H. Chou, R. Paulsen, and S. G. Britt. 1999. Blue- and green-absorbing visual pigments of *Drosophila*: ectopic expression and physiological characterization of the R8 photoreceptor cell-specific Rh5 and Rh6 rhodopsins. *Journal of Neuroscience* 19:10716-10726.
- Salcedo, E., L. Zheng, M. Phistry, E. E. Bagg., and S. G. Britt. 2003. Molecular basis for ultraviolet vision in invertebrates. *The Journal of Neuroscience* 23:10873-10878.
- Sandeman, D. C., R. E. Sandeman, and H. G. DeCoutet. 1990. Extraretinal photoreceptors in the brain of the crayfish *Cherax destructor*. *Journal of Neurobiology* 21:619-629.
- Sanderson, M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* 19:101-109.
- Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Molecular Biology and Evolution* 19:1218-1231.
- Savill, N. J., D. C. Hoyle, and D. G. Higgins. 2001. RNA sequence evolution with secondary structure constraints: Comparison of substitution rate models using maximum likelihood methods. *Genetics* 157:399-411.
- Schiff, H., and M. E. Hendrickx. 1997. An introductory survey of ecology and sensory receptors of tropical eastern Pacific crustaceans. *Italian Journal of Zoology* 64:13-30.
- Scholtz, G., and S. Richter. 1995. Phylogenetic systematics of the reptantian Decapoda (Crustacea, Malacostraca). *Zool. Jour. Linn. Soc.* 113:289-328.
- Schöniger, M., and A. von Haeseler. 1994. A stochastic model for the evolution of autocorrelated DNA sequences. *Molecular Phylogenetics and Evolution* 3:240-247.
- Schrag, S., V. Perrot, and B. R. Levin. 1997. Adaptation to the fitness costs of antibiotic resistance in *Escherichia coli*. *Proceedings of the Royal Society of London B* 264:1287-1291.
- Schram, F. R. 1984. Relationships within eumalacostracan Crustacea. *Trans. San Diego Soc. Nat. Hist.* 20:301-312.
- Schram, F. R. 1986. Crustacea. Oxford University Press, New York.
- Schram, F. R. 2000. Phylogeny of decapods. *Studi e Ricerche, Assoc. Amici Mus. civ. 'G. Zannato', Montecchio Maggiore*, 61.
- Schram, F. R. 2001. Phylogeny of decapods: moving towards a consensus. *Hydrobiologia* 449:1-20.
- Schram, F. R., and C. J. Dixon. 2003. Fossils and decapod phylogeny. *Contributions to Zoology* 72:169-172.
- Schram, F. R., R. M. Feldmann, and M. J. Copeland. 1978. The late Devonian Palaeopalaemonidae and the earliest decapod crustaceans. *Jour. Paleo* 52:1375-1387.

- Schram, F. R., and C. J. Hof. 1998. Fossils and the interrelationships of major Crustacean groups. Pp. 233-302 in G. D. Edgecombe, ed. Arthropod Fossils and Phylogeny. Columbia University Press, New York.
- Schweitzer, C. E., R. M. Feldmann, G. Gonzales-Barba, and F. J. Vega. 2002. New crabs from the Eocene and Oligocene of Baja California Sur, Mexico and an assessment of the evolutionary and paleobiogeographic implications of Mexican fossil decapods. The Paleontological Society Memoir 59:1-43.
- Scott, S., and M. I. Mote. 1974. Spectral sensitivity in some marine crustacea. Vision Research 14:659-663.
- Seki, T., and K. Vogt. 1998. Evolutionary aspects of the diversity of visual pigment chromophores in the class Insecta. Comp. Biochem. Physiol. B 119:53-64.
- Shand, J. 1993. Changes in the spectral absorption of cone visual pigments during the settlement of the goatfish *Upeneus tragula*: the loss of red sensitivity as a benthic existence begins. Journal of Comparative Physiology A 173:115-121.
- Shand, J. 1994. Changes in retinal structure during development and settlement of the goatfish *Upeneus tragula*. Brain Behav Evol 43:51-60.
- Shand, J., J. C. Partridge, S. N. Archer, G. W. Potts, and J. N. Lythgoe. 1988. Spectral absorbance changes in the violet/blue sensitive cones of the juvenile pollack, *Pollachius pollachius*. Journal of Comparative Physiology A 163:699-703.
- Shen, D., M. Jiang, W. Hao, L. Tao, M. Salazar, and H. K. Fong. 1994. A human opsin-related gene that encodes a retinaldehyde-binding protein. Biochemistry 33:13117-13125.
- Shi, Y., F. B. Radlwimmer, and S. Yokoyama. 2001. Molecular genetics and the evolution of ultraviolet vision in vertebrates. Proceedings of the National Academy of Science USA 98:11731-11736.
- Shi, Y., and S. Yokoyama. 2003. Molecular analysis of the evolutionary significance of ultraviolet vision in vertebrates. Proceedings of the National Academy of Science USA 100:8308-8313.
- Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki, and H. Ishikawa. 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. Nature 407:81-86.
- Shimizu, I., Y. Yamakawa, Y. Shimazaki, and T. Iwasa. 2001. Molecular cloning of *Bombyx* cerebral opsin (Boceropsin) and cellular localization of its expression in the silkworm brain. Biochem. Biophys. Res. Commun. 287:27-34.
- Shimodaira, H. 2001. Multiple comparisons of log-likelihoods and combining nonnested models with applications to phylogenetic tree selection. Commun. Stat. Theory Methods 30:1751-1772.
- Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Systematic Biology 51:492-508.
- Shimodaira, H., and M. Hasegawa. 1999a. An approximately.:.
- Shimodaira, H., and M. Hasegawa. 1999b. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution 16:1114-1234.
- Shukolyukov, S. A., O. Kalishevich, A. D. Polyanovsky, and F. G. Gribakin. 1984. Vision in the crab *Hemigrapsus sanguineus*: spectral sensitivity and ommatidium fine structure. Marine Biology 2:53-59.

- Simpson, A. A., V. Chandrasekar, B. Hebert, G. M. Sullivan, M. G. Rossmann, and C. R. Parrish. 2000. Host range and variability of calcium binding by surface loops in the capsids of canine and feline parvoviruses. *Journal of Molecular Biology* 300:597-610.
- Smith, A. B. 1994. Systematics and the fossil record. Blackwell, Oxford, UK.
- Smith, K. C., and E. R. Macagno. 1990. UV photoreceptors in the compound eye of *Daphnia magna* (Crustacea, Branchiopoda) - a 4th spectral class in single ommatidia. *Journal of Comparative Physiology a-Sensory Neural and Behavioral Physiology* 166:597-606.
- Smith, W. C., D. M. Ayers, M. P. Popp, and P. A. Hargrave. 1997. Short wavelength-sensitive opsins from the Saharan silver and carpenter ants. *Invert Neurosci* 3:49-56.
- Smith, W. C., D. A. Price, R. M. Greenberg, and B.-A. Battelle. 1993. Opsins from the lateral eyes and ocelli of the horseshoe crab, *Limulus polyphemus*. *Proceedings of the National Academy of Science USA* 90:6150-6154.
- Soltis, P. S., D. E. Soltis, V. Savolainen, P. R. Crane, and T. G. Barraclough. 2002. Rate heterogeneity among lineages of tracheophytes: Integration of molecular and fossil data and evidence for molecular living fossils. *Proceedings of the National Academy of Science USA* 99:4430-4435.
- Soni, B. G., and R. G. Foster. 1997. A novel and ancient vertebrate opsin. *FEBS Letters* 406:279-283.
- Soni, B. G., A. R. Philp, R. G. Foster, and B. E. Knox. 1998. Novel retinal photoreceptors. *Nature* 394:27-28.
- Spaethe, J., and A. D. Briscoe. 2004. Early duplication and functional diversification of the opsin gene family in insects. *Molecular Biology and Evolution*???????
- Stavenga, D. G., R. P. Smits, and B. J. Hoenders. 1993. Simple exponential functions describing the absorbance bands of visual pigment spectra. *Vision Research* 33:1011-1017.
- Stearns, D. E., and R. B. J. Forward. 1984. Photosensitivity of the calanoid copepod, *Acartia tonsa*. *Marine Biology* 82:85-90.
- Stowe, S. 1980. Spectral sensitivity, retinal pigment movement in the crab *Leptograpsus variegatus* (Fabricius). *Journal of Experimental Biology* 87:73-98.
- Strickler, A. G., K. Famuditimi, and W. R. Jeffery. 2002. Retinal homeobox genes and the role of cell proliferation in cavefish eye degeneration. *International Journal of Developmental Biology* 46:285-294.
- Strickler, A. G., Y. Yamamoto, and W. R. Jeffery. 2001. Early and late changes in Pax6 expression accompany eye degeneration during cavefish development. *Development Genes and Evolution* 211:138-144.
- Strimmer, K., and A. Rambaut. 2002. Inferring confidence sets of possible misspecified gene trees. *Proceedings of the Royal Society of London B* 269:137-142.
- Struwe, G. 1972. Spectral sensitivity of the compound eye in butterflies (*Heliconius*). *Journal of Comparative Physiology A* 79:191-196.
- Sun, H., D. J. Gilbert, N. G. Copeland, N. A. Jenkins, and J. Nathans. 1997. Peropsin, a novel visual pigment-like protein located in the apical microvilli of the retinal pigment epithelium. *Proceedings of the National Academy of Science USA* 94:9893-9898.

- Suzuki, T., K. Arikawa, and E. Eguchi. 1985. The effects of light and temperature on the rhodopsin-porphyrin visual system of the crayfish, *Procambarus clarkii*. *Zoological Science* 2:455-461.
- Suzuki, T., and E. Eguchi. 1987. Survey of 3-dehydroretinal as a visual pigment chromophore in various species of crayfish and other freshwater crustaceans. *Experientia* 43:1111-1113.
- Suzuki, T., M. Makino-Tasaka, and E. Eguchi. 1984. 3-Dehydroretinal (vitamin A₂ aldehyde) in crayfish eye. *Vision Research* 24:783-787.
- Suzuki, T., K. Uji, and Y. Kito. 1976. Studies on cephalopod rhodopsin: photoisomerization of the chromophore. *Biochem Biophys Acta* 428:321-338.
- Swofford, D. L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, Massachusetts.
- Tamas, I., L. Klasson, B. Canback, A. K. Naslund, A. S. Eriksson, J. J. Wernegreen, J. P. Sandstrom, N. A. Moran, and S. G. E. Andersson. 2002. 50 Million years of genomic stasis in endosymbiotic bacteria. *Science* 296:2376-2379.
- Taylor, R. S., S. Yan-bin, and F. R. Schram. 1998. New pygocephalomorph crustaceans from the Permian of China and their phylogenetic relationships. *Paleontology* 41:815-834.
- Taylor, S. D., K. Dittmar de la Cruz, M. L. Porter, and M. F. Whiting. 2005. Characterization of the long-wavelength opsin from Mecoptera and Siphonaptera: Does a Flea see? *Molecular Biology and Evolution* 22:1165-1174.
- Teller, D. C., R. E. Stenkamp, and K. Palczewski. 2003. Evolutionary analysis of rhodopsin and cone pigments: connecting the three-dimensional structure with spectral tuning and signal transfer. *FEBS Letters* 555:151-159.
- Teotónio, H., and M. R. Rose. 2000. Variation in the reversibility of evolution. *Nature* 408:463-466.
- Teotónio, H., and M. R. Rose. 2001. Perspective: Reverse Evolution. *Evolution* 55:653-660.
- Teotónio, H., and M. R. Rose. 2002. Reverse evolution of fitness in *Drosophila melanogaster*. *Journal of Evolutionary Biology* 15:608-617.
- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 24:4876-4882.
- Thorley, J. L., and R. D. M. Page. 2000. RadCon: phylogenetic tree comparison and consensus. *Bioinformatics* 16:486-487.
- Thorne, J. L., and H. Kishino. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* 51:689-702.
- Thorne, J. L., H. Kishino, and I. S. Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15:1647-1657.
- Towner, P., and W. Gartner. 1994. The primary structure of mantid opsin. *Genetics* 143:227-231.
- Towner, P., P. Harris, A. J. Wolstenholme, C. Hill, K. Worm, and W. Gartner. 1997. Primary structure of locust opsins: a speculative model which may account for ultraviolet wavelength light detection. *Vision Research* 37:495-503.

- Townson, S. M., B. S. W. Chang, E. Salcedo, L. V. Chadwell, N. E. Pierce, and S. G. Britt. 1998. Honeybee blue- and ultraviolet-sensitive opsins: Cloning, heterologous expression in *Drosophila*, and physiological characterization. The Journal of Neuroscience 18:2412-2422.
- Tshudy, D. 2003. Clawed lobster (Nephropidae) diversity through time. Journal of Crustacean Biology 23:178-186.
- Tudge, C. C., and D. M. Scheltinga. 2002. Spermatozoal morphology of the freshwater anomuran *Aegla longirostris* Bond-Buckup and Buckup, 1994 (Crustacea: Decapoda: Aeglidae) from South America. Proceedings of the the Biological Society of Washington 115:118-128.
- Väinölä, R. 1986. Sibling species and phylogenetic relationships of *Mysis relicta* (Crustacea: Mysidacea). Annls. Zool. Fenn. 23:207-221.
- Väinölä, R. 1990. Molecular time scales for evolution in *Mysis* and *Pontoporeia*. Ann. Zool. Fennici 27:211-214.
- Väinölä, R. 1992. Evolutionary genetics of marine *Mysis* spp. (Crustacea: Mysidacea). Marine Biology 114:539-550.
- Väinölä, R., B. J. Riddoch, R. D. Ward, and R. I. Jones. 1994. Genetic zoogeography of the *Mysis relicta* species group in northwestern Europe and North America. Can. J. Fish. Aquat. Sci. 51:1490-1505.
- Van Dover, C. L., E. Z. Szuts, S. C. Chamberlain, and J. R. Cann. 1989. A novel eye in 'eyeless' shrimp from hydrothermal vents of the Mid-Atlantic Ridge. Nature 337:458-460.
- van Ham, R. C. H. J., J. Kamerbeek, C. Palacios, C. Rausell, F. Abascal, U. Bastolla, J. M. Fernandez, L. Jimenez, M. Postigo, F. J. Silva, J. Tamames, E. Viguera, A. Latorre, A. Valencia, F. Moran, and A. Moya. 2003. Reductive genome evolution in *Buchnera aphidicola*. Proceedings of the National Academy of Science USA 100:581-586.
- Van Straelen, V. 1928. On a fossil freshwater crayfish from eastern Mongolia. Bulletin of the Geological Society of China 7:173-178.
- Vanhoutte, K. J. A., B. J. L. Eggen, J. J. M. Janssen, and D. G. Stavenga. 2002. Opsin cDNA sequences of a UV and green rhodopsin of the satyrine butterfly *Bicyclus anynana*. Insect Biochemistry and Molecular Biology 32:1383-1390.
- Vassilatis, D. K., J. G. Hohmann, H. Zeng, F. Li, J. E. Ranchalis, M. T. Mortrud, A. Brown, S. S. Rodriguez, J. R. Weller, A. C. Wright, J. E. Bergmann, and G. A. Gaitanaris. 2003. The G protein-coupled receptor repertoires of human and mouse. Proceedings of the National Academy of Science USA 100:4903-4908.
- Verslycke, T. A., N. Fockedey, C. L. McKenney, S. D. Roast, M. B. Jones, J. Mees, and C. R. Janssen. 2004. Mysid crustaceans as potential test organisms for the evaluation of environmental endocrine disruption: A review. Environmental Toxicology and Chemistry 23:1219-1234.
- Wake, M. H. 1992. "Regressive" evolution of special sensory organs in Caecilians (Amphibia: Gymnophiona): Opportunity for morphological innovations. Zool. Jb. Anat. 122:325-329.
- Wald, G., and E. B. Seldin. 1968. Spectral sensitivity of the common prawn, *Palaemonetes vulgaris*. J. Gen. Physiol. 51:694-700.

- Watling, L. 1981. An alternative phylogeny of peracarid crustaceans. *Journal of Crustacean Biology* 1:201-210.
- Watling, L. 1983. Peracaridan disunity and its bearing on eumalacostracan phylogeny with a redefinition of eumalacostracan superorders. *Crustacean Issues* 1:213-228.
- Watling, L. 1999. Toward understanding the relationships of the Peracarida orders: the necessity of determining exact homologies. Pp. 73-89 in F. R. Schram and J. C. von Vaupel Klein, eds. *Crustaceans and the Biodiversity Crisis*. Proceedings of the Fourth International Crustacean Congress. Leiden, The Netherlands: Brill NV, Amersterdam.
- Weiyun, A., and C. Minjuan. 1990. Electrophysiological studies on the visual characteristics of *Portunus trituberculatus* (Miers). *Oceanologia et Limnologia Sinica* 21:490-494.
- Wernegreen, J. J., H. Ochman, I. B. Jones, and N. A. Moran. 2000. Decoupling of genome size and sequence divergence in a symbiotic bacterium. *Journal of Bacteriology* 182:3867-3869.
- Wheeler, W. C., G. Giribet, and G. D. Edgecombe. 2004. Arthropod systematics: The comparative study of genomic, anatomical, and paleontological information. Pp. 281-295 in J. Cracraft and M. J. Donoghue, eds. *Assembling the Tree of Life*. Oxford University Press, New York.
- Whelan, S., and N. Goldman. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular Biology and Evolution* 18:691-699.
- White, R. H., P. K. Brown, A. K. Hurley, and P. R. Bennett. 1983. Rhodopsins, retinula cell ultrastructure and receptor potenials in the developing pupal eye of the moth *Manduca sexta*. *Journal of Comparative Physiology* 150:153-163.
- Whitfield, R. P. 1880. Notice of new forms of fossil crustaceans from the Upper Devonian rocks of Ohio, with decriptions of new genera and species. *Am. J. Sci.* 19:33-42.
- Whiting, M. F. 2002. Mecoptera is paraphyletic: Multiple genes and phylogeny of Mecoptera and Siphonaptera. *Zoologica Scripta* 31:93-104.
- Whiting, M. F., S. Bradler, and T. Maxwell. 2003. Loss and recovery of wings in stick insects. *Nature* 421:264-267.
- Whiting, M. F., J. C. Carpenter, Q. D. Wheeler, and W. C. Wheeler. 1997. The Strepsiptera problem: Phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Systematic Biology* 46:1-68.
- Wiens, J. J. 1999. Phylogenetic evidence for multiple losses of a sexually selected character in phrynosomatid lizards. *Proceedings of the Royal Society of London B* 266:1529-1535.
- Wiens, J. J. 2001. Widespread loss of sexually selected traits: how the peacock lost its spots. *Trends in Ecology and Evolution* 16:517-523.
- Wilkens, L. A., and J. L. Larimer. 1976. Photosensitivity in the sixth abdominal ganglion of decapod crustaceans: A comparative study. *Journal of Comparative Physiology A* 106:69-75.

- Wilkie, S. E., P. R. Robinson, T. W. Cronin, S. Poopalasundaram, J. K. Bowmaker, and D. M. Hunt. 2000. Spectral tuning of avian violet- and ultraviolet-sensitive visual pigments. *Biochemistry* 39:7895-7901.
- Williams, A. J., D. M. Hunt, J. K. Bowmaker, and J. D. Mollon. 1992. The polymorphic photopigments of the marmoset: spectral tuning and genetic basis. *EMBO J* 11:2039-2045.
- Wills, M. A. 1998. A phylogeny of recent and fossil Crustacea derived from morphological characters. Pp. 189-209 in R. A. Fortey and R. H. Thomas, eds. *Arthropod Relationships*. Chapman & Hall, London.
- Wittmann, K. J. 1999. Global biodiversity in Mysidacea, with notes on the effects of human impact. Pp. 511-525 in F. R. Schram and J. C. von Vaupel Klein, eds. *Crustaceans and the Biodiversity Crisis*. Proceedings of the Fourth International Crustacean Congress. Leiden, The Netherlands: Brill NV, Amersterdam.
- Wooley, S., J. Johnson, M. J. Smith, K. A. Crandall, and D. A. McClellan. 2003. TreeSAAP: Selection on amino acid properties using phylogenetic trees. *Bioinformatics* 19:671-672.
- Yamamoto, Y., and W. R. Jeffery. 2000. Central role for the lens in cave fish eye degeneration. *Science* 289:631-633.
- Yan, E. C., M. A. Kazmi, S. De, B. S. W. Chang, C. Seibert, E.P. Marin, R.A. Mathies, and T. P. Sakmar. 2003. Function of extracellular loop 2 in rhodopsin: Glutamic acid 181 modulates stability and absorption wavelength of metarhodopsin II. *Biochemistry* 41:3620-3627.
- Yang, Z. 1996. Among-site rate variation and its impact on phylogenetic analysis. *Trends in Ecology and Evolution* 11:367-372.
- Yang, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* 13:555-556.
- Yang, Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Molecular Biology and Evolution* 15:568-573.
- Yang, Z. 2004. A heuristic rate smoothing procedure for maximum likelihood estimation of species divergence times. *Acta Zoologica Sinica* 50:645-656.
- Yang, Z., and J. P. Bielawski. 2000. Statistical methods for detecting molecular adaptation. *Trends in Ecology and Evolution* 15:496-503.
- Yang, Z., and R. Nielsen. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *Journal of Molecular Evolution* 46:409-418.
- Yang, Z., R. Nielsen, N. Goldman, and A.-M. K. Pedersen. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155:431-449.
- Yang, Z., W. S. W. Wong, and R. Nielsen. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution* 22:1107-1118.
- Yang, Z., and A. D. Yoder. 2003. Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cute-looking mouse Lemur species. *Systematic Biology* 52:705-716.
- Yoder, A. D., and Z. Yang. 2000. Estimation of primate speciation dates using local molecular clocks. *Molecular Biology and Evolution* 17:1081-1090.

- Yoder, A. D., and Z. Yang. 2004. Divergence dates for Malagasy lemurs estimated from multiple gene loci: geologic and evolutionary context. *Molecular Ecology* 13:757-773.
- Yokoyama, R., and S. Yokoyama. 1990. Convergent evolution of the red- and green-like visual pigmen genes in fish, *Astyanax fasciatus*, and human. *Proceedings of the National Academy of Science USA* 87:9315-9318.
- Yokoyama, S. 2000. Evolution of vertebrate visual pigments. *Progress in Retinal and Eye Research* 19:385-419.
- Yokoyama, S. 2002. Molecular evolution of color vision in vertebrates. *Gene* 300:69-78.
- Yokoyama, S., and F. B. Radlwimmer. 1999. The molecular genetics of red and green color vision in mammals. *Genetics* 153:919-932.
- Yokoyama, S., and F. B. Radlwimmer. 2001. The molecular genetics and evolution of red and green color vision in vertebrates. *Genetics* 158:1697-1710.
- Yokoyama, S., F. B. Radlwimmer, and N. S. Blow. 2000. Ultraviolet pigments in birds evolved from violet pigments by a single amino acid chage. *Proceedings of the National Academy of Science USA* 97:7366-7371.
- Yokoyama, S., and Y. Shi. 2000. Genetics and evolution of ultraviolet vision in vertebrates. *FEBS Letters* 486:167-172.
- Yokoyama, S., and T. Tada. 2000. Adaptation of the African and Indonesian coelacanths to deep-sea environments. *Gene* 261:35-42.
- Yokoyama, S., and T. Tada. 2003. The spectral tuning in the short wavelength-sensitive type 2 pigments. *Gene* 306:91-98.
- Yokoyama, S., and H. Zhang. 1997. Cloning and characterisation of the pineal gland-specific opsin gene of marine lamprey (*Petromyzon marinus*). *Gene* 202:89-93.
- Yokoyama, S., H. Zhang, F. B. Radlwimmer, and N. S. Blow. 1999. Adaptive evolution of color vision of the Comoran coelacanth (*Latimeria chalumnae*). *Proceedings of the National Academy of Science USA* 96:6279-6284.
- Zeiger, J., and T. H. Goldsmith. 1989. Spectral properties of porphyropsin from an invertebrate. *Vision Research* 29:519-527.
- Zeiger, J., and T. H. Goldsmith. 1993. Packaging of rhodopsin and porphyropsin in the compound eye of the crayfish. *Vis. Neurosci* 10:193-202.
- Zeiger, J., and T. H. Goldsmith. 1994. Behavior of crayfish rhodopsin and metarhodopsin in digitonin: the 510 and 562 nm 'visual pigments' are artifacts. *Vision Research* 34:2679-2688.
- Zhang, X.-J., W. A. Baase, and B. W. Matthews. 1991. Towards a simplification of the protein folding problem: A stabilizing polyalanine alpha-helix engineered in T4 lysozyme. *Biochemistry* 30:2012-2017.
- Zhukovsky, E. A., and D. D. Oprian. 1989. Effect of carboxylic acid side chains on the absorption maximum of visual pigments. *Science* 246:928-930.
- Zhukovsky, E. A., P. R. Robinson, and D. D. Oprian. 1992. Changing the location of the Schiff base counterion in rhodopsin. *Biochemistry* 31:10400-10405.
- Ziedins, I. P., and V. B. Meyer-Rochow. 1990. ERG-determined spectral and absolute sensitivities in relation to age and size in the half crab *Petrolisthes elongatus*. *Experimental Biology* 48:319-328.

- Zuckerkandl, E., and L. Pauling. 1965. Evolutionary divergence and convergence in proteins. Pp. 97-166 in V. Bryson and H. J. Vogel, eds. *Evolving genes and proteins*. Academic Press, New York.
- Zuker, C. S. 1996. The biology of vision in *Drosophila*. *Proceedings of the National Academy of Science USA* 93:571-576.
- Zuker, C. S., C. Montell, K. Jones, T. Laverty, and G. M. Rubin. 1987. A rhodopsin gene expressed in photoreceptor cell R7 of the *Drosophila* eye: homologies with other signal-transducing molecules. *J. Neurosci.* 7:1550-1557.