



Molecular evolution of a long wavelength-sensitive opsin in mimetic *Heliconius* butterflies (Lepidoptera: Nymphalidae)

RUBY HSU¹, ADRIANA D. BRISCOE², BELINDA S. W. CHANG³ and NAOMI E. PIERCE

Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA

Received 18 May 2000; accepted for publication 3 November 2000

This study examines the pattern of opsin nucleotide and amino acid substitution among mimetic species 'rings' of *Heliconius* butterflies that are characterized by divergent wing colour patterns. A long wavelength opsin gene, *OPS1*, was sequenced from each of seven species of *Heliconius* and one species of *Dryas* (Lepidoptera: Nymphalidae). A parsimony analysis of *OPS1* nucleotide and amino acid sequences resulted in a phylogeny that was consistent with that presented by Brower & Egan in 1997, which was based on mitochondrial *cytochrome oxidase I* and *II* as well as nuclear *wingless* genes. Nodes in the *OPS1* phylogeny were well supported by bootstrap analysis and decay indices. An analysis of specific sites within the gene indicates that the accumulation of amino acid substitutions has occurred independently of the morphological diversification of *Heliconius* wing colour patterns. Amino acid substitutions were examined with respect to their location within the opsin protein and their possible interactions with the chromophore and the G-protein. Of the 15 amino acid substitutions identified among the eight species, one nonconservative replacement (A226Q) was identified in a position that may be involved in binding with the G-protein.

© 2001 The Linnean Society of London

ADDITIONAL KEYWORDS: visual pigment – mimicry – phylogeny – invertebrate vision – butterfly – lepidopteran.

INTRODUCTION

The butterfly genus *Heliconius* (Lepidoptera: Nymphalidae) is well known for its classic examples of mimicry, in which distantly-related species display wing patterns that are nearly indistinguishable to the casual observer. The bright coloration of these butterflies advertises the presence of toxins to potential predators (Brower, Brower & Collins, 1963; Nahrstedt & Davis, 1981), and the convergence of wing colour and pattern between unpalatable species is often cited as an archetypal example of Müllerian mimicry. The enormous diversity, rapid morphological evolution, and patterns

of convergence within *Heliconius* have made it a model system for testing hypotheses of micro- and macro-evolution (Turner, 1971; Benson, 1972; Turner, 1981; Gilbert, 1983; Brower, 1994b; McMillan, Jiggins & Mallet, 1997).

Relatively little is known about vision in *Heliconius* despite the longstanding importance of the genus in butterfly behavioral ecological and evolutionary studies. Electrophysiology studies on *H. erato*, *H. numata*, and *H. sara* have suggested that these species have three spectral types of photoreceptor cells with λ_{\max} in the ranges of 370–390 nm, 450–470 nm, and 550–570 nm (Struwe, 1972a,b). In addition, a number of behavioral and neurophysiological experiments have been performed to determine the relationship between wing colour and colour preference. Some studies indicate that the strongest colour preference or peak wavelength sensitivity typically occurs at wavelengths matching the species' own wing colour (Eltringham, 1933). For instance Crane (1955) showed that, when courting, *H. erato* hydara is most strongly attracted to the orange-red colour found on its own forewings. Swihart (1971) found that *H. charithonia*, a black

¹ Current address: Howard Hughes Medical Institute and Center for Neurobiology and Behavior, College of Physicians and Surgeons of Columbia University, 722 West 168th Street, New York, NY 10032, USA

² Corresponding author. Current address: University of Colorado Health Sciences Center, Department of Cellular and Structural Biology, 4200 East 9th Ave., Denver, CO 80262. E-mail: adriana.briscoe@uchsc.edu

³ Current address: Rockefeller University, 1230 York Ave., Box 284, New York, NY 10021, USA

butterfly with yellow bands on its wings, exhibited the strongest preference for yellow hues. However, these studies focus on only two species, and a general correlation between peak spectral sensitivity, at the level of the individual photoreceptor cell or at higher levels of visual processing, and wing colour has not been established conclusively.

Because vision is crucial for survival and, in particular, the detection of conspecifics and mates, we decided to examine the patterns of amino acid substitution in an opsin gene in eight species that have undergone rapid shifts in wing colour usage. Most evolutionary change in *Heliconius* wing colour has occurred in the yellow to red portions of the spectrum (475–700 nm) (Swihart, 1967), which is in the long wavelength (LW) region. In addition, the colour variation in the wing patterns of the *Heliconius* species used in this study is primarily in the LW region. We therefore focused on cloning a LW-sensitive opsin to examine whether or not there was any correlation between changes in wing colour and the pattern of opsin amino acid evolution. If there were spectral tuning in relation to wing colour, we might expect convergent evolution of opsins in butterflies with mimetic wing patterns.

Because of their central role in vision, opsins provide a unique model for examining the interplay between an organism's genotype and its phenotype. The spectral sensitivity of a photoreceptor cell is in many cases largely determined by the absorption spectrum maximum (λ_{\max}) of its visual pigments. Visual pigments are composed of a opsin protein that is covalently bound to a light-sensitive retinal chromophore (Wald, 1968). Opsins are members of the large family of G-protein-coupled receptors. This family is characterized by seven α -helical transmembrane domains, which in opsins form a barrel around a central binding pocket for the chromophore. Interactions between amino acid side chains and the chromophore within this pocket are primarily responsible for spectral tuning (for review, see Sakmar, 1998). In the first step of the biochemical visual cascade that occurs in a photoreceptor, a photon of light is absorbed by the chromophore, which then isomerizes and causes a conformational change in the associated opsin. This change leads to the activation of the heterotrimeric G protein transducin, which triggers a biochemical cascade that eventually results in release of neurotransmitter (Nathans, 1987).

Comparative sequence analysis makes use of a phylogenetic hypothesis upon which amino acid substitutions potentially associated with shifts in maximum sensitivity (λ_{\max}) can be mapped (Yokoyama & Yokoyama, 1990; Chang *et al.*, 1995; Yokoyama, 1995). This method identifies naturally occurring substitutions whose phenotypic expression can be tested using site-directed mutagenesis. For example, Chang

et al. (1995) proposed a model for blue spectral shifts in vertebrate visual pigments using comparative sequence analysis, which was later confirmed experimentally through site-directed mutagenesis and resonance Raman spectroscopy (Lin *et al.*, 1998). Amino acid residues potentially involved in wavelength regulation have also been identified in freshwater crayfishes (Crandall & Cronin, 1997).

Four putative LW opsin genes have been characterized to date in the butterfly *Papilio glaucus*, identified as *PglRh1*, *PglRh2*, *PglRh3*, and *PglRh4* (Briscoe, 1998). In this study, we have cloned a LW opsin in *Heliconius*, which we have named *OPS1*. After generating a phylogenetic hypothesis based on *OPS1* molecular characters and evaluating the data for congruence with other *Heliconius* cladograms generated from *wingless*, *COI* and *COII* (Fig. 1) (Brower, 1994a; Brower & Egan, 1997) data, comparative sequence analysis was used to identify residues that are potentially involved in spectral tuning and/or other regulatory aspects of the visual system. Additional comparisons were made with sites that may be involved in LW spectral tuning in other organisms, such as crayfish (Crandall & Cronin, 1997), humans (Merbs & Nathans, 1993; Asenjo, Rim & Oprian, 1994), and New World monkeys (Shyue *et al.*, 1998).

MATERIAL AND METHODS

DNA AND RNA ISOLATION

Butterflies were obtained as live specimens from Dr Larry Gilbert, University of Texas, Austin, or as genomic DNA extracts from Dr. Andrew Brower, Oregon State University, Corvallis (Table 1).

After being sexed with a dissecting scope, butterflies were frozen in liquid nitrogen and stored at -80°C . Genomic DNA (gDNA) was extracted from whole abdomens of all individuals. Approximately 100 μg of abdominal tissue was homogenized in 400 μl of 50 mM Tris-Cl (pH 8.0/2% SDS) and digested with 2 μl of Proteinase K. After incubating overnight at 37°C , 100 μl of saturated NaCl was added and the solution was placed on ice for 35 minutes. The cellular debris was eliminated in a centrifuge spun for 15 minutes at 13 000 rpm and 4°C , and DNA was ethanol precipitated.

In some instances, complementary DNA (cDNA) was synthesized in addition to gDNA in order to verify that the opsin locus was being transcribed and to supplement the data set. Head tissue RNA from two available species, *H. melpomene* and *H. erato*, was extracted with Trizol, and cDNA was then synthesized with either SuperScript RNase H-Reverse Transcriptase (GibcoBRL) with a polyT adapter primer or the Marathon cDNA Amplification kit (Clontech).

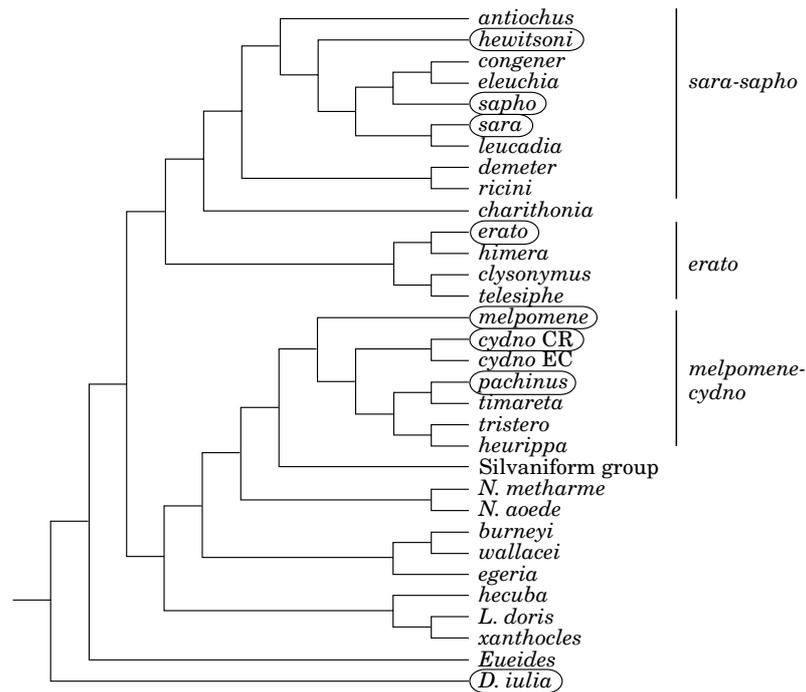


Figure 1. Phylogeny of *Heliconius* inferred from mitochondrial *COI* and *COII* and nuclear *wingless* genes (redrawn from Figure 3 in Brower & Egan, 1997). Species included in this study are circled. Select clades have been condensed and some terminal branches eliminated for simplicity. Clades named by Brower & Egan are indicated to the right.

Table 1. Butterfly specimens used in this study. u.i. = unavailable information

Species	Specimen name	GenBank #	Sex	Wing colour	Location
<i>Heliconius cydnno</i>	isolate E-1-2	AF126755	u.i.	black, white, blue	Pichincha Prov., Ecuador
<i>H. erato</i>	HER-1	AF126750	F	black, red, yellow	Sirena Station, Costa Rica
<i>H. hewitsoni</i>	HHW-1	AF126752	M	black, yellow	Pacific Slope, Costa Rica
<i>H. melpomene rosina</i>	HMP-1 & 13	AF126751	M	black, red, yellow	Pacific Slope, Costa Rica
<i>H. pachinus</i>	HPA-1	AF126756	F	black, yellow	Sirena Station, Costa Rica
<i>H. sapho</i>	isolate P-16-5	AF126754	u.i.	black, white, blue	Darien Prov., Panama
<i>H. sara fulgidus</i>	HSA-1	AF126753	M	black, yellow	Sirena Station, Costa Rica
<i>Dryas iulia</i>	DRJ-1	AF126757	M	orange	Texas

PCR, CLONING, AND SEQUENCING

A short fragment of the opsin gene spanning transmembrane domains VI and VII was first amplified using a polymerase chain reaction (PCR) from each species with *Taq* DNA polymerase (Qiagen, Inc.) and the degenerate primers 80 and OPSRD (Fig. 2, Table 2). In order to obtain sequence for the five remaining transmembrane domains, this initial fragment was then used to design a gene-specific reverse primer, 5'RACE Rh1 (Fig. 2, Table 2), which was paired with a degenerate forward primer, LWFD (Fig. 2, Table 2), designed against other known insect LW-sensitive opsins and located in the N-terminus. Once this larger genomic fragment was cloned, further gene-specific

primers were designed by 'walking' down the cloned gene product.

A PCR thermocycling protocol of 1' at 94°C, 30 cycles of [1' at 94°C, 30'' at 55°C, and 1' at 68°C], and 10' at 68°C was the most successful at amplifying opsins. The PCR products were purified using a QIAquick PCR purification kit (Qiagen, Inc.). Most PCR products were not easy to sequence directly, and these were cloned into a pCR2.1-TOPO plasmid (TOPO TA cloning or Original TA Cloning, Invitrogen). After *EcoRI* digestion and screening for inserts, the targeted plasmids were prepared using a QIAprep Spin Miniprep Kit and then cycle sequenced (Dye Primer and Dye Terminator Cycle Sequencing Ready Reaction kits, Perkin Elmer)

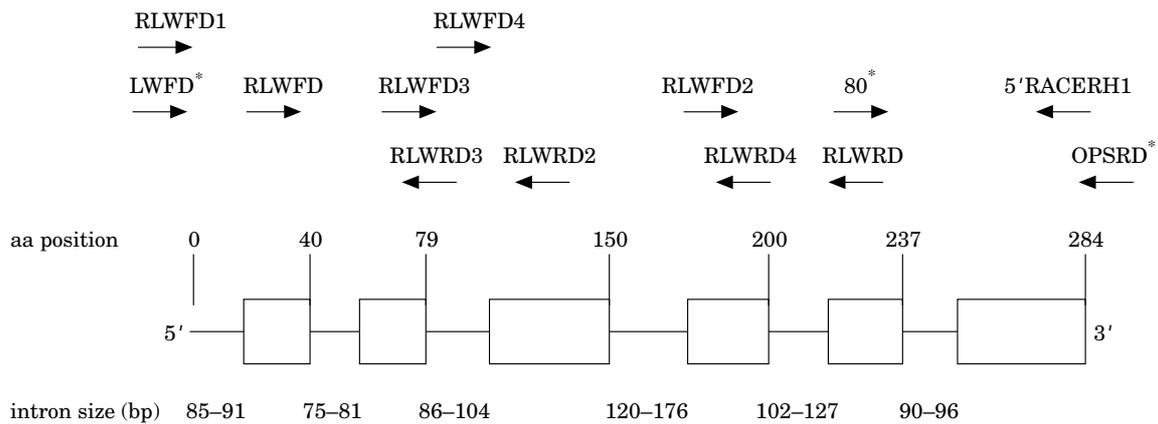


Figure 2. Primers used for both PCR and sequencing mapped onto the opsin genomic sequence. Open boxes indicate exons. Amino acid positions are assigned relative to the *H. cydno* sequence (GenBank accession #AF126755). Degenerate PCR primers (not *OPS1*-specific) are noted with an asterisk.

Table 2. Sequences of primers used for both PCR and sequencing in this study. Amino acid positions are numbered relative to the *H. cydno* sequence (GenBank accession #AF126755). Degenerate PCR primers (not *OPS1*-specific) are noted with an asterisk

Primer	Sequence	Source
LWFD*	5'-CAY HTN RTN GAY CCN CAY TGG-3'	Briscoe, 1998
RLWFD	5'-TNG GNT TYA THT CNG TNA CNG GN-3'	newly designed
RLWRD	5'-NAC RTT CAT YTT YTT NGC-3'	newly designed
RLWFD1	5'-CAY TTN ATH GAY CCN CAY TGG-3'	newly designed
RLWFD2	5'-CCN GAR GGN AAY ATG ACN G-3'	newly designed
RLWRD2	5'-DAT NAC RTT RTA NCG RTC G-3'	newly designed
RLWFD3	5'-ATG GGT ATT CCG TAG GTT K-3'	newly designed
RLWRD3	5'-AAC SKW CCG AAT ACC CAT-3'	newly designed
RLWFD4	5'-ACA AAA ATC AAT GGG AGA GTA-3'	newly designed
RLWRD4	5'-TGA ACA ATG AAG AAG TAG GAG-3'	newly designed
80*	5'-GAA CAR GCW AAR AAR ATG A-3'	Chang <i>et al.</i> , 1996
5'RACERH1	5'-TGG TTA CAA TAG GGC TGA-3'	Briscoe, 1998
OPSRD*	5'-CCR TAN ACR ATN GGR TTR TA-3'	Chang <i>et al.</i> , 1996

on an ABI 370A DNA automated fluorescent cycle sequencing machine (Applied Biosystems). In almost all cases, two or more independent PCR products were cloned and sequenced in both directions. Primers that were also used for amplification were also used for sequencing. In addition, multiple PCR products were directly sequenced in all cases in order to verify the sequences from cloned products and to check for *Taq* DNA polymerase error. Electropherograms generated by automated sequencing were imported into Sequencher 3.0 (Gene Codes Corp., 1995) for manual editing. Ambiguous sites were resolved by comparison across electropherograms within a species. Contigs were then assembled automatically in Sequencher 3.0.

DATA MATRIX AND PHYLOGENETIC ANALYSIS

Amino acid and nucleotide sequences were viewed in SeqEd 1.0.3 (Applied Biosystems, 1992). Coding regions were strictly conserved and were aligned by eye, but introns showed greater variation and were aligned using ClustalW (Thompson, Higgins & Gibson, 1994). Exhaustive searches using maximum parsimony were performed on *OPS1* data sets with a test version of PAUP*4.0.0d62 (Swofford, 1998). A step matrix of amino acid substitutions constructed from a larger data set of G protein-coupled receptors (Rice, 1994) was employed as a weighting scheme in order to take into account the unequal probabilities of state changes among amino acids of different classes. Par-

simony analysis with the nucleotide data was unweighted. To establish the robustness of the tree topologies, bootstrap searches were performed with 10 000 replicates and 10 heuristic searches per replicate. The heuristic search parameters included tree bisection-reconnection (TBR) with MULPARS and steepest descent invoked.

A phylogenetic analysis was also performed on the *OPS1* amino acid sequences combined with opsin sequences from 28 other invertebrates. A heuristic search of 100 replicates using TBR was employed, again with MULPARS and steepest descent invoked. The same profile was used for the bootstrap analysis, with 100 replicates and 10 additions per replicate. Due to the large size of the data matrix, uninformative characters were excluded from the bootstrap calculations. The amino acid step matrix (Rice, 1994) was employed in both the parsimony and the bootstrap analyses. Phylogenetic trees and amino acid changes were viewed in MacClade 3.0 (Maddison & Maddison, 1993), and pairwise divergences between species were calculated using MEGA 1.01 (Kumar, Tamura & Nei, 1993).

RESULTS

Molecular sequence was collected for a 1612 basepair region of the LW opsin *OPS1*, including coding and non-coding regions as well as gaps. The data matrix used in the coding region analysis consisted of 284 amino acid characters and 852 basepairs across all eight species (Appendix 1). Full-length (1612 basepairs) genomic sequences were collected for all species except for *H. erato* and *H. melpomene*, in which nucleotides 1–153, 1107–1612 (*H. erato*) and 1–173, 1107–1612 (*H. melpomene*) could not be amplified from our sample of genomic DNA. Information in these regions was obtained from the cDNA sequences, which showed complete identity and were therefore combined with the gDNA. The 5' and 3' primer sites were excluded from all matrices.

Six introns were found within the gene region characterized, all of which began with the nucleotides GT and ended with AG. The splice sites of the introns were conserved in all eight species except for *D. iulia*, which was missing the first intron, and *H. erato* and *H. melpomene*, in which genomic information was not available for the regions containing the first, fifth, and sixth introns. The six introns ranged in size from 75 to 176 basepairs, occurring after amino acid positions 0, 40, 200, and 237 and interrupting amino acids 79 and 150 (Fig. 2, and Appendix 1, numbered according to the *H. cydno* coding region).

In addition, the coding region of *OPS1* was found to be highly conserved within the eight species, with only 83 of 852 (9.7%) variable nucleotides and 15 of 284

(5.3%) variable amino acid residues (Table 3). Third codon positions showed the highest variability, with 61 informative sites compared to 16 in the first position and 6 in the second. The data also showed a higher proportion of synonymous than nonsynonymous substitutions, but both proportions were relatively low throughout the data set, never exceeding 0.025 (nonsynonymous) or 0.239 (synonymous) between any two species.

PHYLOGENETIC ANALYSIS

Phylogenetic analysis was performed on the nucleotide sequences to infer the relatedness of the species. A heuristic maximum parsimony search with 10 000 replicates (unweighted) of the coding region only (852 basepairs) yielded three most parsimonious trees, which included all three possible resolutions of the bootstrap consensus tree (Fig. 3). These trees were each 96 steps in length. The consensus tree separated *Heliconius* into an *erato-hewitsoni-sara-sapho* clade and a *melpomene-pachinus-cydno* clade that were well supported (bootstrap values = 98% and 100%, decay indices = 7 and >10). The data fit the trees very well, as indicated by a consistency index of 0.906 and a retention index of 0.907.

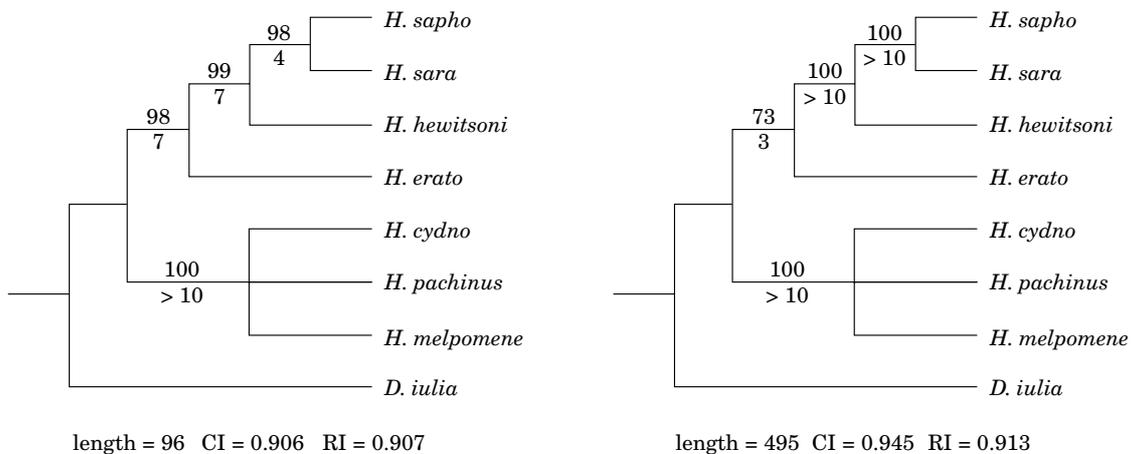
An exhaustive parsimony search was also performed on the nucleotide data matrix with the inclusion of introns (1612 basepairs), with gaps in *H. erato* and *H. melpomene* treated as missing information. This search yielded three most parsimonious trees of 495 steps each, and the consensus retained the same topology as the tree generated from coding regions only (Fig. 3). Overall, the genomic tree was more robust than the coding region tree, with 100% bootstrap values (10 000 heuristic search replications, 10 random taxon additions per replicate) and >10 decay indices. The genomic data also fit the tree better, with a consistency index of 0.945 and a retention index of 0.913. The *erato-hewitsoni-sara-sapho* clade was not as well supported, however, with a bootstrap value of 73% and a decay index of 3.

In order to ensure that the missing introns in *H. erato* and *H. melpomene* did not affect the inferred phylogeny, the analysis was repeated on a reduced data set (933 basepairs) in which the missing regions had been trimmed in all species. This analysis recovered the identical consensus topology as the full genomic matrix, with bootstrap values within 3%.

In addition, an exhaustive maximum parsimony search was performed on the coding region using amino acid sequences (284 amino acids), which resulted in one most parsimonious tree that was consistent with the nucleotide analyses (Fig. 4). The identical tree topology was recovered both with and without a step matrix (Rice, 1994). The resolution of the *melpomene-*

Table 3. Character of amino acid and nucleotide sequences in *OPS1* coding regions (284 amino acids, 852 nucleotides sequenced)

Character status	Nucleotides				Amino acids
	1st position	2nd position	3rd position	All positions	
Constant	268	278	223	769	269
Variable and uninformative	8	1	28	37	10
Informative	8	5	33	46	5
Total	284	284	284	852	284

**Figure 3.** 50% consensus trees produced by maximum parsimony analysis on nucleotide data without (left) and with (right) introns using an unweighted exhaustive search. The tree shown to the right was generated using the full genomic nucleotide matrix with missing introns treated as missing information. Bootstrap values shown above branches (10 000 heuristic search replicates, 10 additions per replicate), decay indices given below. Trees rooted using *D. iulia*.

pachinus-cydno and the *erato-hewitsoni-sara-sapho* clades was again well supported with bootstrap values (10 000 replications) of 100% and 95% and decay indices of >10 and 6, respectively. All nodes within the two clades were unresolved, but the consistency and retention indices were both 1.000.

A final phylogenetic analysis was performed on the *OPS1* amino acid data set combined with known UV, blue, blue-green, and LW opsins from 28 other invertebrates obtained from the GenBank database (see Appendix 2 for accession numbers). A maximum parsimony bootstrap heuristic search (100 replications, 10 additions per replicate) and weighted with a step matrix (Rice, 1994) produced a single most parsimonious tree with a length of 1894 steps (Fig. 4). With uninformative characters excluded, the consistency index was 0.641 and the retention index was 0.739, indicating that the data fit the tree reasonably well. Within the tree, the eight *OPS1* sequences clustered in a monophyletic group with a bootstrap value of 100%.

AMINO ACID SUBSTITUTIONS

Fifteen amino acid residues were variable within the coding region sequenced in *OPS1*, six of which were located in one of the transmembrane domains (Table 4, Fig. 5). Three of these six, A62F (A→F at amino acid 62), G133F, and L189F, were substitutions from small amino acids to the larger phenylalanine. All three involve nonpolar residues and are not at sites likely to face the chromophore binding pocket (Baldwin, 1993). When mapped onto the cladogram of the species, none of the fifteen substitutions within the gene exhibited convergence between species (either according to mimetic patterns or otherwise); all informative replacements followed phylogenetic lines (Table 4).

DISCUSSION

ORTHOLOGY AND HOMOLOGY OF OPS1

Because opsins constitute a multi-gene family, in sequencing opsins from multiple species several steps

Table 4. Amino acid substitution positions and identities in coding regions sequenced in *OPS1* (asterisks indicate phylogenetically informative sites). Amino acid positions are numbered relative to the *H. cydno* sequence (GenBank accession #AF126755). TM=transmembrane. Hypothetical examples of convergent substitution patterns according to mimicry rings are given using two character states, which present any nonconservative substitution (0=ancestral state, 1=derived state).

Convergence in: red ring	0	1	0	0	0	0	0	1
yellow ring	0	0	1	0	0	1	1	0
black-and-white ring	0	0	0	1	1	0	0	0
No convergence:	0	0	0	0	1	1	1	1

aa Position	Domain	DRJ	HMP	HPA	HCY	HSP	HSA	HHW	HER
1	N-terminus	Y	H	H	H	H	H	H	H
41	loop I-II	S	S	S	S	S	S	S	T
62	TM II	A	F	F	F	F	F	F	F
133*	TM IV	G	F	F	F	G	G	G	G
143*	TM IV	V	V	V	V	I	I	I	I
164	loop IV-V	D	D	D	D	D	D	E	D
168	loop IV-V	K	Q	Q	Q	Q	Q	Q	Q
170*	loop IV-V	M	I	I	I	F	F	F	F
173*	loop IV-V	I	M	M	M	R	R	R	R
174*	loop IV-V	S	T	T	T	S	S	S	S
177	TM V	V	L	L	L	L	L	L	L
181	TM V	V	I	I	I	I	I	I	I
189	TM V	L	F	F	F	F	F	F	F
226*	loop V-VI	A	Q	Q	Q	A	A	A	A
263	loop VI-VII	E	K	K	K	K	K	K	K

were taken to ensure that the cloned opsin genes were orthologous, that is descended from the same ancestral gene, and not paralogous, or descended from duplicated ancestral genes. First, gene-specific primers were carefully designed from preliminary PCR products that had been cloned and sequenced. Second, the sequences generated from *OPS1* were combined with other invertebrate opsins in a phylogenetic analysis. The eight *OPS1* sequences formed a well-supported monophyletic group within the tree (bootstrap value=100%), which provided strong evidence that the *Heliconius* and *Dryas* sequences are from the same gene. In addition, the sequences were highly conserved within the species, and the relatively low levels of substitution suggested orthology (Tables 3, 4). Finally,

an examination of the intron splice sites from the available genomic data indicated that at least three of the introns sites are strictly conserved in *OPS1*.

Comparisons were also made with other opsins in order to support the classification of *OPS1* as a LW opsin. In examining intron sites in relation to those found in invertebrate opsins, none of the six *OPS1* sites corresponded to the intron positions in UV- and blue-sensitive visual pigments (Briscoe, 1999). However, the data set showed that the number and position of the *OPS1* introns were identical to those in *PglRh3*, the only putative LW opsin fully characterized genomically from butterflies (Briscoe, 1999). *PglRh1*, *PglRh2*, and *PglRh4*, the three other putative LW opsins characterized in *Papilio glaucus* (Briscoe, 2000),

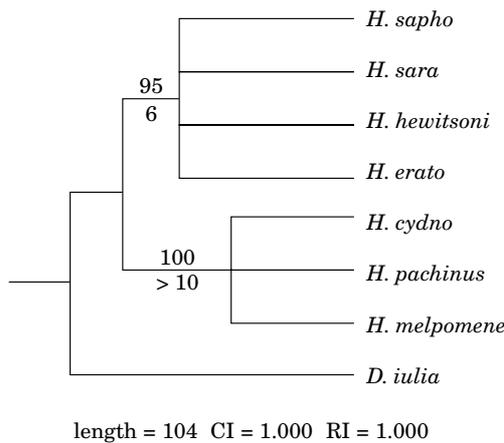


Figure 4. Most parsimonious tree produced by maximum parsimony analysis on amino acid data (coding regions only) using an exhaustive search. Weighted using an amino acid step matrix (Rice, 1994). Bootstrap values shown above branches (10 000 heuristic search replicates, 10 additions per replicate), decay indices given below. Tree rooted using *D. iulia*.

have only one intron identified to date, which corresponds to the sixth intron in *OPS1*.

An examination of amino acid residues that are highly conserved among opsins was also useful in classifying *OPS1*. The Glu-113 residue known to act as the stabilizing counterion to the protonated Schiff's base in bovine rhodopsin (Sakmar, Franke & Khorana, 1989; Zhukovsky & Oprian, 1989; Nathans, 1990; Zhukovsky, Robinson & Oprian, 1992) was present in *OPS1* as Tyr-87 (Fig. 5). The glutamine at this site is conserved in all vertebrate opsins, but all invertebrate opsins except the UV share a tyrosine at that site (Chang *et al.*, 1995). The *OPS1* sequences also exhibited the highly conserved Trp-161 residue (numbered Trp-135 in *OPS1*) that is present in all visual pigments except *Drosophila Rh3* UV opsins, and may be involved in the shifting of *Rh3* toward such short wavelengths (Chang, 1995).

Within the invertebrate opsins that were analysed, the *OPS1* sequences were nestled in the LW clade and most closely related to *Rh3* and *Rh1* of the *Papilio* butterfly (Kitamoto *et al.*, 1998; Briscoe, 1998, 2000) (Fig. 6). The conserved introns between the genes and

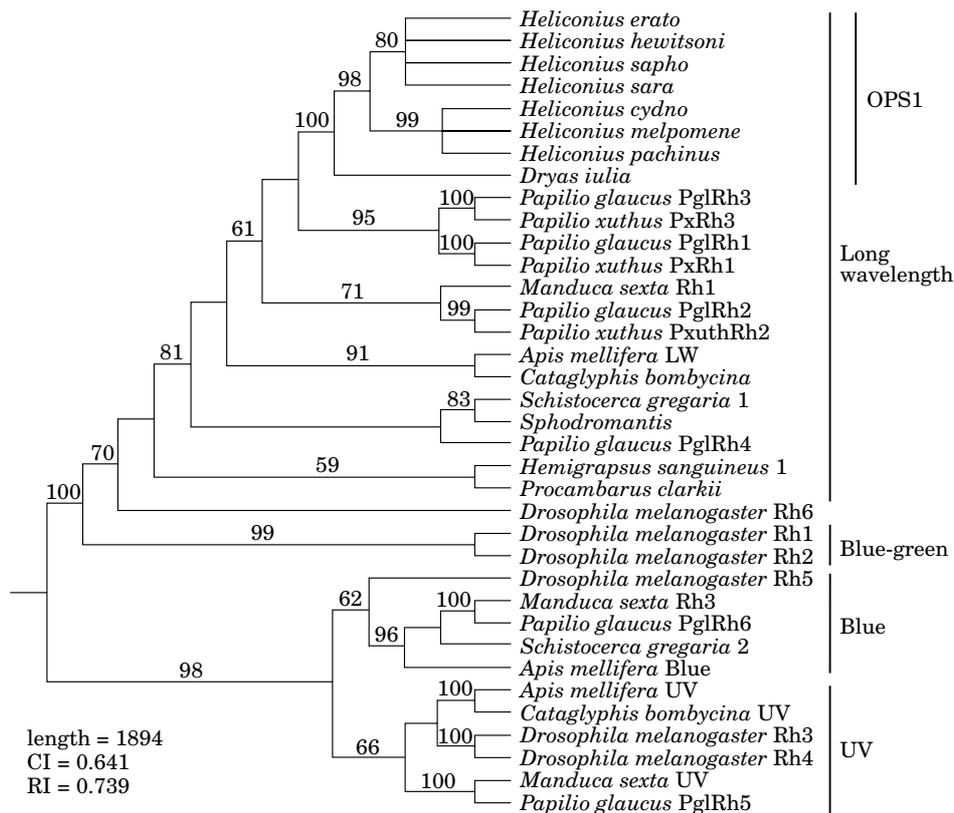


Figure 5. Single most parsimonious tree of *OPS1* and other invertebrate opsins. Obtained from a heuristic search of amino acids using 100 replicates. Bootstrap values obtained from an analysis using 100 heuristic search replicates with 10 random taxon additions per replicate (uninformative sites excluded). All analyses were weighted using an amino acid step matrix (Rice, 1994), and the tree was rooted using the UV- and blue-sensitive opsins (Chang *et al.*, 1995).

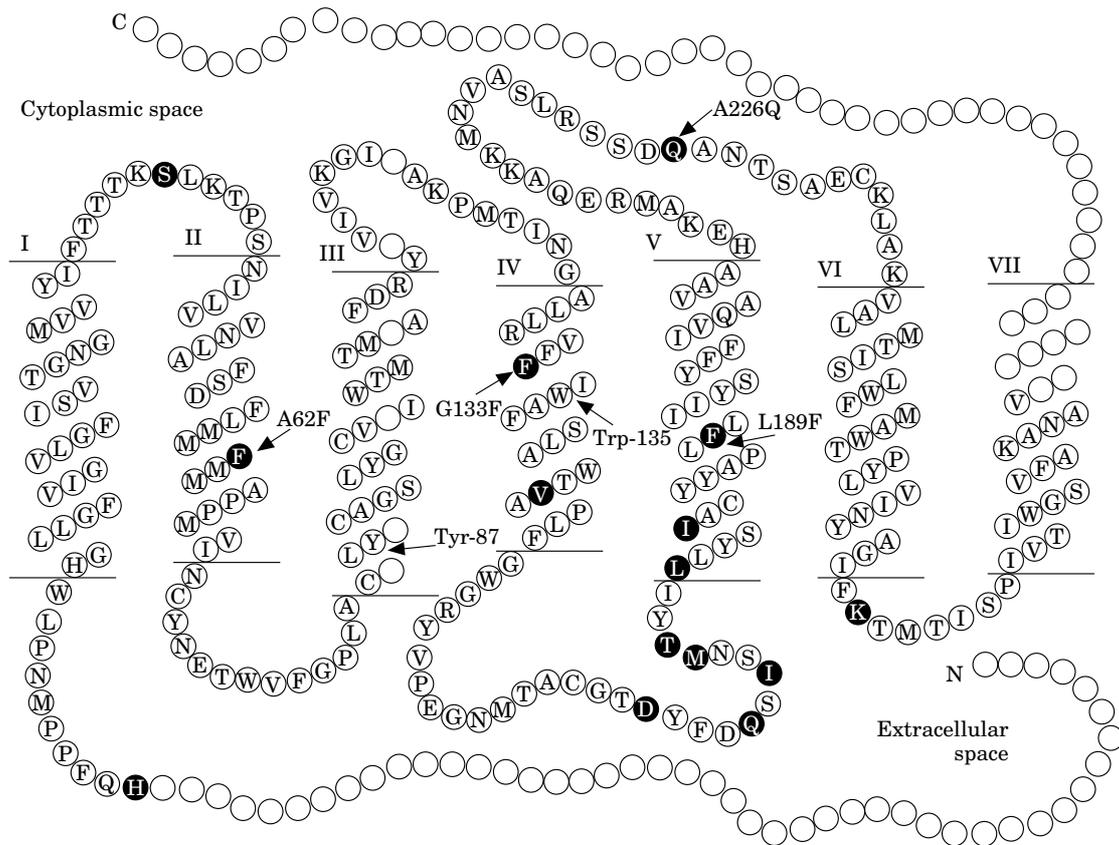


Figure 6. Topographical model of the opsin protein encoded by *OPS1*. Bars indicate the location of the microvillous membrane of the photoreceptor cell, and roman numerals refer to the seven α -helical transmembrane domains of the opsin. The protein sequence of *H. cydno* is given (GenBank accession #AF126755). Shaded circles are substitution sites and arrows indicate residues of interest (see text).

the close grouping within the opsin family suggest that *OPS1* is indeed a LW opsin that could be homologous with the opsins *Papilio Rh3* or *Rh1* (not to be confused with the blue-green- or UV-sensitive opsins found in *Drosophila*, labeled *Dmel Rh1* and *Dmel Rh3*). However, the bootstrap support for the grouping of *OPS1* with *Papilio Rh3* and *Rh1* is less than 50%, so without additional information, any proposed homologies are preliminary. If *OPS1* is indeed homologous to *Papilio Rh3* or *Rh1*, it is possible that the *Papilio* gene duplication was more recent than that of *OPS1*.

SYSTEMATICS OF *HELICONIUS* AND PHYLOGENETIC UTILITY OF *OPS1*

Heliconius as a group ranges from the United States to Argentina, but most species are confined to the moist tropics (DeVries, 1987). Of the eight species used in this study, *H. hewitsoni* and *H. pachinus* are Central American endemics. Like all *Heliconius* species, their hostplants are members of the diverse tropical genus *Passiflora* (Passifloraceae). The taxa are sample spe-

cies selected from three different mimicry 'rings', or groups of sympatric species that have converged on the same warning colour pattern (Brower *et al.*, 1963; Turner, 1984; Mallet & Gilbert, 1995). *H. melpomene* and *H. erato* belong to a 'red' ring; *H. pachinus*, *H. hewitsoni*, and *H. sara* to a 'yellow' ring; and *H. cydno* and *H. sapho* to a third 'black-and-white' ring (Mallet & Gilbert, 1995). *D. iulia* belongs to a separate 'orange' ring (DeVries, 1987). The seven species were chosen because they exhibit well-established patterns of morphological convergence across three mimicry rings, and they show a diversity of wing colours. It has also been demonstrated in Costa Rican *Heliconius* that species within a mimicry ring are sympatric and therefore subjected to the same selective forces (e.g. predation) which drive mimetic evolution (Mallet & Gilbert, 1995).

Until recently, the phylogenetic classification of *Heliconius* has been based on morphological, behavioral, and cytological characters (Michener, 1942; Emsley, 1973; summarized in Brown, 1981). Because of the high degree of polymorphism within species and the convergence among distantly related species, resolving

the systematics of these taxa has been challenging. The first phylogenetic study to combine morphological and molecular data involved a restriction analysis of ribosomal RNA (rRNA) genes; the molecular data was largely uninformative and their cladogram was mostly supported by morphological characters (Lee *et al.*, 1992). More recent and comprehensive analyses of *Heliconius* have used molecular characters of mitochondrial (*COI* and *COII*) (Brower, 1994a) and nuclear (*wingless*) genes (Brower & Egan, 1997) (Fig. 1). These were the first extensive systematic studies of *Heliconius* species based on molecular data. Brower & Egan's hypothesis generally corroborates morphological studies (see Brown, 1981; Penz, 1999) with respect to the two major clades of *Heliconius* considered in the present study. Collectively, these studies strongly support the independent evolution of wing patterns in the co-mimics under investigation, thus revealing instances of striking convergence in wing-pattern morphology.

OPS1 produced a cladogram for *Heliconius* that was consistent with Brower & Egan's (1997) analysis using mitochondrial *COI* and *II* and nuclear *wingless* genes. This indicates that the genealogical lineage of *OPS1* matches the species lineage in *Heliconius*. The nucleotide data yielded a cladogram that strongly supported the monophyly of the two clades, which Brower & Egan refer to as the *erato-sara-sapho* (combined) clade and the *melpomene-cydno* clade (Fig. 1). The division into two clades separates the mimetic pairs and corroborates the independent evolution of the wing patterns (Brown, 1981; Lee *et al.*, 1992; Brower, 1994a). Within the *erato-sara-sapho* clade, *H. sara* and *H. sapho* were grouped as most closely related species, in concordance with Brower & Egan (1997). The topology is robust, as indicated by high bootstrap values and decay indices (Fig. 3). The phylogeny produced using amino acid data also strongly supported the division of the two clades.

The inclusion of the intron data in the nucleotide analysis generally increased the robustness of the tree (raising both bootstrap values and decay indices), but the *melpomene-cydno* clade was still unresolved. Similarly, Brower & Egan (1997) also found this clade to be problematic. Brower & Egan's analyses used a data matrix that included two *H. cydno* samples (one from Ecuador and one from Costa Rica). In the separate mitochondrial and nuclear phylogenies (Brower, 1994a; Brower & Egan, 1997), these two samples of *H. cydno* were polyphyletic. Only the tree from the combined data set recovered them as a monophyletic group.

The lack of complete resolution in the *OPS1* phylogeny may also be due to a lower rate of substitution in nuclear genes (5.3% amino acids and 9.7% nucleotides variable in *OPS1*). A comparison with the *COI*, *COII*,

and *wingless* data shows that slight discrepancies may exist between cladograms generated from mitochondrial and nuclear genes, most likely due to higher rates of evolution in mitochondrial DNA (Brown, George & Wilson, 1979; Harrison, 1989). As a result, unless the substitutions are saturated, mitochondrial genes may yield higher resolution than nuclear genes (Brower & DeSalle, 1998). The collapsed *melpomene-pachinus-cydno* node may also be due to the small number of species investigated. Autapomorphic substitutions could potentially become phylogenetically informative with the inclusion of more species, thus giving more resolution in the tree.

PATTERNS OF AMINO ACID SUBSTITUTION

The physical structure of the rhodopsin protein has been recently visualized using crystallography (Palczewski *et al.*, 2000), and the inferred structure has been extensively studied in vertebrates (reviewed in Nathans, 1990, and Yokoyama & Yokoyama, 1996). It is known that substitutions within the transmembrane domains which face the binding pocket may affect the opsin-chromophore interaction. Based on models of bovine rhodopsin (Hargrave *et al.*, 1983; Baldwin, 1993), the approximate location of *OPS1* substitutions within the protein could be determined through homology (see Chang *et al.*, 1995, for an alignment of vertebrate and invertebrate opsin sequences). Six amino acid replacements were located in one of the seven transmembrane domains (Fig. 5), but all six were in areas of the helices that face away from the chromophore binding pocket, as determined from the bovine rhodopsin model.

None of the amino acid residues in the transmembrane domains involved a change in amino acid side chain polarity, which can affect the absorption spectra of the visual pigment (Merbs & Nathans, 1993; Sun, Macke & Nathan, 1997). In addition, no substitutions were homologous with any of the sites thought to be responsible for LW spectral shifts in crayfish, humans, or New World monkeys (data not shown).

However, comparisons with vertebrate opsins can be problematic, as invertebrate opsins are known to differ somewhat both structurally and functionally (Autrum, 1979; Britt *et al.*, 1993; Townson *et al.*, 1998). In this light, it may be worth noting that the three replacements within the transmembrane domains (residues 62, 133, and 189) were from alanine, glycine, and leucine, all of which have very small side chains, to phenylalanine, which has a large side chain with a benzene ring. If located within the binding pocket, this increase in size and the acquisition of an aromatic ring could potentially affect interactions with the chromophore, either by constricting the size or shape of

the binding pocket or through the increased polarizability of the side chain. Constricting the size of the binding pocket may involve an increased sensitivity to light activation or a more narrow range of sensitivity, as may be the case in rod versus cone opsins (Chang & Donoghue, in prep.). Increased polarizability of the side chain is thought to be responsible for blue spectral shifts in both vertebrate and insect blue opsins (Chang, 1995; Lin *et al.*, 1998). Alternatively, these substitutions to larger amino acids could alter spectral sensitivities through indirect steric effects that result in tertiary structural changes. Small spectral shifts based on such effects have been postulated in crayfish (Crandall & Cronin, 1997).

One substitution, A226Q, is located in the cytoplasmic loop between transmembrane domains V and VI and could be functionally significant (Fig. 6). There is evidence which indicates that residues in this loop are involved in the coupling with the G-protein in vertebrates. The A226Q substitution was also one of the five phylogenetically informative substitutions (Table 6), and it is nonconservative, showing a change from a hydrophobic, very small amino acid, alanine, to a polar, larger amino acid, glutamine.

CONCLUSION

The study of *Heliconius* colour and wing patterns has historically been a rewarding field for evolutionary biologists, inspiring theories on natural and sexual selection by Darwin (1874) and Müller (1879). Because the recognition of conspecific butterflies acts as a strong selective pressure on the visual system, this data set was used to determine whether the morphological divergence observed in the wings of mimetic *Heliconius* species correlates with molecular divergence in the visual system. Although the results showed relatively little variation in the *OPS1* gene, and no wing colour-correlated pattern of amino acid substitution, future electrophysiological studies could be useful in determining the precise λ_{\max} of each of these mimetic species. This could establish whether or not the visual systems are conserved at the functional level, as well as provide valuable information regarding photoreceptor classes in *Heliconius*. Other avenues for research on the *Heliconius* visual system could involve analysis of the mechanism of selective enhancement, the arrangement and relative numbers of photoreceptor cells, or sequences from additional opsin genes.

In addition, amino acid sites that were identified in this study could be involved in spectral tuning or other aspects of the visual cascade. Site-directed mutagenesis could be applied to these residues, particularly to determine potential effects of the A226Q substitution on the G-protein interaction. Finally, this

study demonstrates the potential utility of the *OPS1* gene for systematic purposes.

ACKNOWLEDGEMENTS

We thank R. C. Lewontin in whose lab this work was conducted, and R. Collins for his hospitality while this project was being finished. L. E. Gilbert and A. V. Z. Brower provided specimens and discussions for which we are grateful. A. Mignault and D. Haig provided helpful comments on the manuscript. This work was supported by grants from the National Science Foundation (IBN-9700941 to R.H., N.E.P. and A.D.B.), Sigma Xi (A.D.B.), and the Putnam Expeditionary Fund of the Museum of Comparative Zoology (A.D.B.). A.D.B. was supported by a Predoctoral Fellowship from the Howard Hughes Medical Institute and a Research Grant from the Canadian Institute for Advanced Research. B.S.W.C. is an NSF/Alfred P. Sloan Postdoctoral Fellow in Molecular Evolution. R.H. was supported by grants from the Harvard College Research Program and the Radcliffe Research Partnership program.

REFERENCES

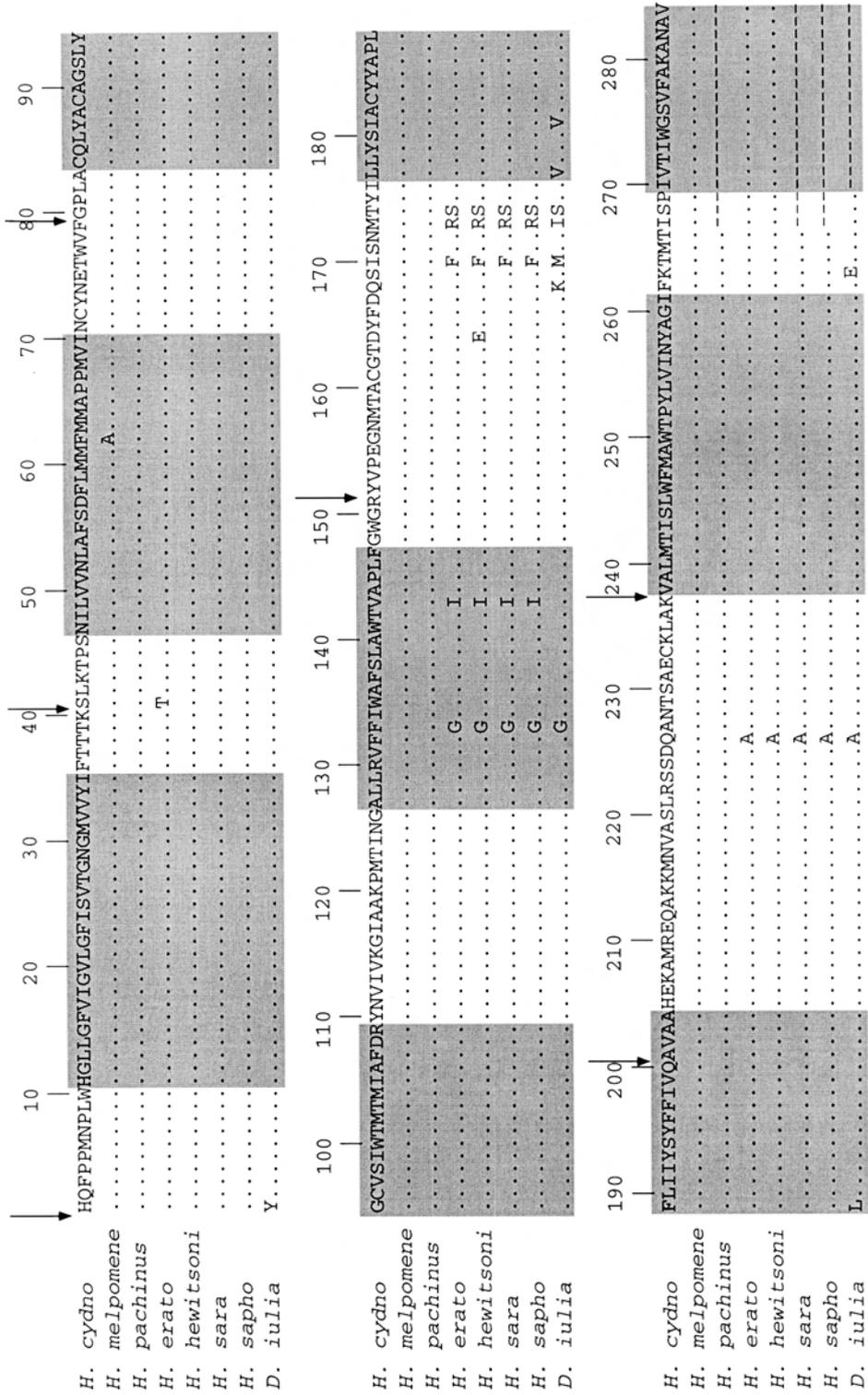
- Asenjo AB, Rim J, Oprian DD. 1994. Molecular determinants of human red/green color discrimination. *Neuron* **12**: 1131–1138.
- Autrum H, ed. 1979. *Comparative Physiology and Evolution of Vision in Invertebrates. A: Invertebrate Photoreceptors. Handbook of Sensory Physiology*. Berlin: Springer-Verlag.
- Baldwin JM. 1993. The probable arrangement of the helices in G protein-coupled receptors. *European Molecular Biology Organization Journal* **12**: 1693–1703.
- Benson WW. 1972. Natural selection for Müllerian mimicry in *Heliconius erato* in Costa Rica. *Science* **176**: 936–939.
- Briscoe AD. 1998. Molecular diversity of visual pigments in the butterfly *Papilio glaucus*. *Naturwissenschaften* **85**: 33–35.
- Briscoe AD. 1999. Intron splice sites of *Papilio glaucus PglRh3* corroborate insect opsin phylogeny. *Gene* **230**: 101–109.
- Briscoe AD. 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.
- Britt SG, Feiler R, Kirschfeld K, Zuker CS. 1993. Spectral tuning of rhodopsin and metarhodopsin *in vivo*. *Neuron* **11**: 29–39.
- Brower AVZ. 1994a. Phylogeny of *Heliconius* butterflies inferred from mitochondrial DNA sequences (Lepidoptera: Nymphalidae). *Molecular Phylogenetics and Evolution* **3**: 159–174.
- Brower AVZ. 1994b. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato*

- inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA* **91**: 6491–6495.
- Brower AVZ, DeSalle R. 1998.** Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of *wingless* as a source of characters for phylogenetic inference. *Insect Molecular Biology* **7**(1): 73–82.
- Brower AVZ, Egan MG. 1997.** Cladistic analysis of *Heliconius* butterflies and relatives (Nymphalidae: Heliconiini): a revised phylogenetic position for *Eueides* based on sequences from mtDNA and a nuclear gene. *Proceedings of the Royal Society of London B* **264**: 969–977.
- Brower LP, Brower JVZ, Collins CT. 1963.** Experimental studies of mimicry. 7. Relative palatability and Müllerian mimicry among Neotropical butterflies of the subfamily Heliconiinae. *Zoologica, NY* **48**: 65–84.
- Brown KS, Jr. 1981.** The biology of *Heliconius* and related genera. *Annual Review of Entomology* **26**: 427–456.
- Brown WM, George M, Wilson AC. 1979.** Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences, USA* **76**: 1967–1971.
- Chang BSW. 1995.** Opsin phylogeny and evolution. Ph.D. Thesis. Harvard University.
- Chang BSW, Ayers D, Smith WC, Pierce NE. 1996.** Cloning of the gene encoding honeybee long-wavelength rhodopsin: a new class of insect visual pigments. *Gene* **173**: 215–219.
- Chang BSW, Crandall KA, Carulli JP, Hartl DL. 1995.** Opsin phylogeny and evolution: a model for blue shifts in wavelength regulation. *Molecular Phylogenetics and Evolution* **4**(1): 31–43.
- Crandall KA, Cronin TW. 1997.** The molecular evolution of visual pigments of freshwater crayfishes (Decapoda: Cambaridae). *Journal of Molecular Evolution* **45**: 524–534.
- Crane J. 1955.** Imaginal behavior of a Trinidad butterfly, *Heliconius erato hydara* Hewitson, with special reference to the social use of color. *Zoologica, NY* **40**(16): 167–196.
- Darwin C. 1874.** *The Descent of Man and Selection in Relation to Sex*. 2nd revised ed. New York: Merrill.
- DeVries PJ. 1987.** *The Butterflies of Costa Rica*. Princeton, NJ: Princeton University Press.
- Eltringham H. 1933.** *The Senses of Insects*. London: Methuen.
- Emsley M. 1973.** A morphological study of Imagine Heliconiinae (Lep: Nymphalidae) with a consideration of the evolutionary relationships within the group. *Zoologica, NY* **48**: 85–130.
- Gilbert LE. 1983.** Coevolution and mimicry. In: Futuyma D, Slatkin M, eds. *Coevolution*. Sunderland, MA: Sinauer Associates, 263–281.
- Hargrave PA, McDowell JH, Curtis DR, Wang JK, Juszczak E, Fong SL, Rao JKM, Argos P. 1983.** The structure of bovine rhodopsin. *Biophysics of Structure and Mechanism* **9**: 235–244.
- Harrison RG. 1989.** Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends in Ecology and Evolution* **6**: 6–11.
- Kitamoto J, Sakamoto K, Ozaki K, Mishina Y, Arikawa K. 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.
- Kumar S, Tamura K, Nei M. 1993.** MEGA. Ver. 1.01. The Pennsylvania State University, University Park, PA.
- Lee CS, McCool BA, Moore JL, Hillis DM, Gilbert LE. 1992.** Phylogenetic study of heliconiine butterflies based on morphology and restriction analysis of ribosomal RNA genes. *Zoological Journal of the Linnean Society* **106**: 17–31.
- Lin SW, Kochendoerfer GG, Carroll KS, Wang D, Mathies RA, Sakmar TP. 1998.** Mechanisms of spectral tuning in blue cone visual pigments. Visible and raman spectroscopy of blue-shifted rhodopsin mutants. *Journal of Biological Chemistry* **273**: 24583–24591.
- Maddison WP, Maddison DR. 1993.** *MacClade*. Ver. 3.0. Sinauer Associates, Sunderland, MA.
- Mallet J, Gilbert LE. 1995.** Why are there so many mimicry rings? Correlations between habitat, behaviour and mimicry in *Heliconius* butterflies. *Biological Journal of the Linnean Society* **55**: 159–180.
- McMillan WO, Jiggins CD, Mallet J. 1997.** What initiates speciation in passion-vine butterflies? *Proceedings of the National Academy of Science USA* **94**: 8628–8633.
- Merbs SL, Nathans J. 1993.** Role of hydroxyl-bearing amino acids in differentially tuning the absorption spectra of the human red and green cone pigments. *Photochemistry and Photobiology* **58**: 706–710.
- Michener CD. 1942.** A generic revision of the Heliconiinae (Lepidoptera: Nymphalidae). *American Museum Novitates* **1197**: 1–8.
- Müller F. 1879.** *Ituna and Thridia*: a remarkable case of mimicry in butterflies. *Proceedings of the Entomological Society of London* *xx–xxix*.
- Nahrstedt A, Davis RH. 1981.** The occurrence of cyanoglucosides, linamarin and lotaustralin, in *Acraea* and *Heliconius* butterflies. *Comparative Biochemistry and Physiology* **68B**: 575–577.
- Nathans J. 1987.** Molecular biology of visual pigments. *Annual Review of Neuroscience* **10**: 163–194.
- Nathans J. 1990.** Determinants of visual pigment absorbance: identification of the retinylidene Schiff's base counterion in bovine rhodopsin. *Biochemistry* **29**: 9746–9752.
- Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Trong IL, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M. 2000.** Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**: 739–745.
- Penz CM. 1999.** Higher level phylogeny for the passion-vine butterflies (Nymphalidae, Heliconiinae) based on early stage and adult morphology. *Zoological Journal of the Linnean Society* **127**: 277–344.
- Rice K. 1994.** The origin, evolution, and classification of G-protein-coupled receptors. PhD thesis. Harvard University.
- Sakmar TP. 1998.** Rhodopsin: a prototypical G protein-coupled receptor. *Progress in Nucleic Acid Research and Molecular Biology* **59**: 1–34.

- Sakmar TP, Franke RR, Khorana HG. 1989.** Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. *Proceedings of the National Academy of Sciences, USA* **86**: 8309–8313.
- Sequencher. 1995.** Ver. 3.0 for Macintosh. Gene Codes Corporation, Ann Arbor, MI.
- Shyue S, Boissinot S, Schneider H, Sampaio I, Schneider MP, Abee CR, Williams L, Hewett-Emmett D, Sperling HG, Cowing JA, Dulai KS, Hunt DM, Li WH. 1998.** Molecular genetics of spectral tuning in New World monkey color vision. *Journal of Molecular Evolution* **46**: 697–702.
- Struwe G. 1972a.** Spectral sensitivity of the compound eye in butterflies (*Heliconius*). *Journal of Comparative Physiology* **79**: 191–196.
- Struwe G. 1972b.** Spectral sensitivity of single photoreceptors in the compound eye of a tropical butterfly (*Heliconius numata*). *Journal of Comparative Physiology* **79**: 197–201.
- Sun H, Macke JP, Nathans J. 1997.** Mechanisms of spectral tuning in the mouse green cone pigment. *Proceedings of the National Academy of Sciences, USA* **94**: 8860–8865.
- Swihart SL. 1967.** Neural adaptations in the visual pathway of certain Heliconiine butterflies, and related forms, to variations in wing coloration. *Zoologica NY* **52**: 1–14.
- Swihart CA. 1971.** Colour discrimination by the butterfly, *Heliconius charitonius* Linn. *Animal Behavior* **19**: 156–164.
- Swofford DL. 1998.** PAUP*: *Phylogenetic Analysis Using Parsimony*. Ver. 4.0.Od62 for Macintosh. Laboratory of Molecular Systematics, Smithsonian Institution.
- Thompson JD, Higgins DG, Gibson TJ. 1994.** ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673–4680.
- Townson SM, Chang BSW, Salcedo E, Chadwell LV, Pierce NE, Britt SG. 1998.** Honeybee blue- and ultraviolet-sensitive opsins: cloning, heterologous expression in *Drosophila*, and physiological characterization. *The Journal of Neuroscience* **18**: 2412–2422.
- Turner JRG. 1971.** Studies on Müllerian mimicry and its evolution in burnet moths and heliconid butterflies. In: Creed ER, ed. *Ecological Genetics and Evolution*. Oxford: Blackwell, 224–260.
- Turner JRG. 1981.** Adaptation and evolution in *Heliconius*. A defense of neo-Darwinism. *Annual Review of Ecology and Systematics* **12**: 99–121.
- Turner JRG. 1984.** Mimicry: the palatability spectrum and its consequences. In: Vane-Wright RI, Ackery PR, eds. *The Biology of Butterflies*. London: Academic Press, 141–161.
- Wald G. 1968.** Molecular basis of visual excitation. *Science* **162**: 230–239.
- Yokoyama R, Yokoyama S. 1990.** Convergent evolution of the red- and green-like visual pigment genes in fish, *Astyanax fasciatus*, and human. *Proceedings of the National Academy of Sciences, USA* **87**: 9315–9318.
- Yokoyama S. 1995.** Amino acid replacements and wavelength absorption of visual pigments in vertebrates. *Annual Review of Ecology and Systematics* **27**: 543–567.
- Yokoyama S, Yokoyama R. 1996.** Adaptive evolution of photoreceptors and visual pigments in vertebrates. *Annual Review of Ecology and Systematics* **27**: 543–567.
- Zhukovsky EA, Oprian DD. 1989.** Effect of carboxylic acid side chains on the absorption maximum of visual pigments. *Science* **246**: 928–930.
- Zhukovsky EA, Robinson PR, Oprian DD. 1992.** Changing the location of the Schiff base counterion in rhodopsin. *Biochemistry* **31**: 10400–10405.

APPENDIX 1

Alignment of seven *Heliconius* and one *Dryas* amino acid sequences (coding regions only). Numbering of amino acids according to the *H. cydno* sequence (GenBank accession #AF126755). Arrows indicate location of introns, shaded positions indicate location of transmembrane domains within the opsin protein.



APPENDIX 2

TABLE OF INVERTEBRATE OPSINS USED IN A PHYLOGENETIC ANALYSIS WITH *OPS1*.

Opsin name	Species	GenBank Accession #
BeeUV	<i>Apis mellifera</i>	AF004169
BeeBlue	<i>Apis mellifera</i>	AF004168
BeeRh1	<i>Apis mellifera</i>	U26026
Cataglyph	<i>Cataglyphis bombycinus</i>	U32501
CatagUV	<i>Cataglyphis bombycinus</i>	AF042787
Crab	<i>Limulus polyphemus</i>	L03782
Crayfish	<i>Procambarus clarkii</i>	S53494
DmelRh1	<i>Drosophila melanogaster</i>	K02315
DmelRh2	<i>Drosophila melanogaster</i>	M12896
DmelRh3	<i>Drosophila melanogaster</i>	M17718
DmelRh4	<i>Drosophila melanogaster</i>	M17730
DmelRh5	<i>Drosophila melanogaster</i>	U67905
DmelRh6	<i>Drosophila melanogaster</i>	Z86118
Lo1	<i>Schistocerca gregaria</i>	X80071
Lo2	<i>Schistocerca gregaria</i>	X80072
Mantid	<i>Sphodromantis</i>	X71665
ManRh1	<i>Manduca sexta</i>	76082
ManRh3	<i>Manduca sexta</i>	1249561
MandUV	<i>Manduca sexta</i>	109852
PglRh1	<i>Papilio glaucus</i>	AF077189
PglRh2	<i>Papilio glaucus</i>	AF077190
PglRh3	<i>Papilio glaucus</i>	AF067080
PglRh4	<i>Papilio glaucus</i>	AF077193
PglRh5	<i>Papilio glaucus</i>	AF077191
PglRh6	<i>Papilio glaucus</i>	AF077192
PxRh1	<i>Papilio xuthus</i>	AB007423
PxRh2	<i>Papilio xuthus</i>	AB007424
PxRh3	<i>Papilio xuthus</i>	AB007425