Adaptive evolution of color vision as seen through the eyes of butterflies

Francesca D. Frentiu*, Gary D. Bernard†, Cristina I. Cuevas*, Marilou P. Sison-Mangus*, Kathleen L. Prudic‡, and Adriana D. Briscoe*§

*Comparative and Evolutionary Physiology Group, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697; †Department of Electrical Engineering, University of Washington, Seattle, WA 98195; and ‡Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721

Butterflies and primates are interesting for comparative color vision studies, because both have evolved middle- (M) and long-wavelength- (L) sensitive photopigments with overlapping absorbance spectrum maxima (λ_{max} values). Although positive selection is important for the maintenance of spectral variation within the primate pigments, it remains an open question whether it contributes similarly to the diversification of butterfly pigments. To examine this issue, we performed epimicrospectrophotometry on the eyes of five *Limenitis* butterfly species and found a 31-nm range of variation in the λ_{max} values of the L-sensitive photopigments (514–545 nm). We cloned partial *Limenitis* L opsin gene sequences and found a significant excess of replacement substitutions relative to polymorphisms among species. Mapping of these L photopigment λ_{max} values onto a phylogeny revealed two instances within Lepidoptera of convergently evolved L photopigment lineages whose λ_{max} values were blue-shifted. A codon-based maximum-likelihood analysis indicated that, associated with the two blue spectral shifts, four amino acid sites (Ile17Met, Ala64Ser, Asn70Ser, and Ser137Ala) have evolved substitutions in parallel and exhibit significant d_N/d_S >1. Homology modeling of the full-length *Limenitis arthemis astyanax* L opsin placed all four substitutions within the chromophore-binding pocket. Strikingly, the Ser137Ala substitution is in the same position as a site that in primates is responsible for a 5- to 7-nm blue spectral shift. Our data show that some of the same amino acid sites are under positive selection in the photopigments of both butterflies and primates, spanning an evolutionary distance >500 million years.

Limenitis | viceroy | visual pigment

Ever since Darwin, there has been intense interest in the evolution of complex traits, with one of the most notable examples being the evolution of the eyes. Only three phyla have evolved image-resolving eyes: arthropods, mollusks, and chordates (1). Although there are vast morphological and neurobiological differences among the eyes of these lineages, there are also many similarities that may provide new insight into the molecular mechanisms governing the evolution of this complex trait.

Image-resolving eyes are composed of photoreceptor cells, which contain photopigments, and they also have accessory pigment cells, which shield the photopigments from stray light. Photopigments are made up of a light-sensitive chromophore (e.g., 11-cis-retinal) and an opsin protein. Although the isolated chromophore has an absorbance spectrum maximum (λ_{max} value) at 380 nm (2), the λ_{max} values of photopigments can vary from 360 to 600 nm through the spectral tuning of the chromophore by specific interactions with amino acids in the chromophore-binding pocket of the opsin protein. Because photopigment sensitivities represent clear adaptations to an animal's light environment (3), the amino acid sites of opsins involved in spectral tuning may be under positive selection, which has been most extensively studied in fish, mammals, and primates (4–6). Comparing these photopigment results with those of a more distant lineage such as butterflies could provide a better understanding of the molecular mechanisms of the eye and how these molecules change under selection. The visual systems of butterflies and primates may be under similar selective pressures. Butterflies and primates share a similar light environment and contain species that are nocturnally or diurnally active and species that are fruit-feeders. Like primates that are leaf eaters, butterfly females need to discriminate among foliage types for oviposition, which may have a strong impact on their color vision (see refs. 7 and 8). As well, both use vision in the detection of conspecifics and mates (9). (Detection of predators is probably visual for butterflies but does not depend on color vision; it probably depends on motion vision instead.) Thus, various aspects of the color vision systems of butterflies and primates may have undergone convergent evolution (see below).

Phylogenetic analyses indicate that the opsin gene family duplicated several times before the radiation of metazoans, giving rise to as many as seven protein subfamilies (10), including the ciliary and rhabdomeric opsins, each associated with a distinct photoreceptor cell type (11). All photoreceptor cells expand their membranes to express opsins, but ciliary photoreceptor cells expand their ciliary side, the side closest to the cell body, and express ciliary opsins, whereas rhabdomeric photoreceptor cells expand their apical side and express rhabdomeric opsins (11). In general, vertebrates have ciliary opsins that are expressed in the photoreceptor cells of the retina, whereas insects have rhabdomeric opsins that are expressed in the photoreceptor cells of the ommatidia of the compound eye.

Color vision adds to the complexity of the eye. With a single spectral class of photoreceptor, only achromatic (brightness-contrast) vision is possible. Both mammalian long-wavelength-sensitive (L) cones and butterfly L photoreceptors provide outputs for brightness processing (12, 13). Color vision, on the other hand, is the ability to discriminate between different wavelengths of light, regardless of relative intensity and depends on the presence of at least two spectrally distinct classes of photoreceptors, as well as appropriate neuronal connections in the brain. Natural selection has recruited both the vertebrate ciliary opsins and the insect rhabdomeric opsins for use in achromatic and color vision. Moreover, mammals use all their cone photopigments for color vision, and although not yet fully investigated, butterflies likely use all their major spectral receptor types for color vision (8, 14, 15).

There is variation in both the photopigment sensitivity and the range of color vision in mammals. Color vision among placental mammals is typically dichromatic based on outputs from short-wavelength-sensitive (S) and L cone pigments encoded by distinct opsin genes expressed in the cone photoreceptor cells of the vertebrate retina (12, 16). Some primates have even evolved trichromatic color vision based on an additional middle-wavelength-sensitive (M) cone pigment, which allows them to discriminate colors in the green-to-red part of the visible light spectrum. The most parsimonious explanation for the variation in primate color vision systems is that the common ancestor of primates contained a single S opsin variant, but the ancestor was polymorphic at another locus for alleles encoding M and L cone pigments (17). In some lineages, such as in nocturnal prosimians, either the M or the L cone pigment has been lost (18). In other lineages, such as in diurnal New World monkeys, as many as three opsin alleles have been maintained at high frequencies by balancing selection (19). In addition, the functionally distinct M and L allelic variants have been fixed independently at least twice (or evolved convergently) as duplicate genes in the New World monkey genus *Alouatta* (20, 21) and in the common ancestor of humans, apes, and Old World monkeys.

Among insects, some butterflies have evolved red–green color vision (7, 22), convergently with primates by gene duplication (23–25), and also through the use of heterogeneously expressed filtering pigments, in combination with a single L opsin (26). In both mammals and butterflies, natural selection has led to the evolution of photopigments with similarly diverse sets of spectral sensitivities in the long wavelength part of the visible light spectrum. For mammals, the L/M pigment $\lambda_{max} = 495–565$ nm (27), whereas in butterflies, L pigment $\lambda_{max} = 500–600$ nm (28). It is fascinating that the ciliary-type L and M pigments of mammals have evolved similar spectral phenotypes as the rhabdomeric-type L pigments of butterflies, and that some primates and butterflies have evolved red–green color vision in parallel. Because the genetic mechanisms underlying this expanded sensitivity to longer wavelengths of light have been elucidated largely in mammals and primates, we investigated whether similar genetic mechanisms accounted for the parallel phenotypes found in butterflies.

To address this question, we determined the peak sensitivity of the L photopigment from a large number ($n > 20$) of butterflies in the family Nymphalidae. In the course of this survey, we found surprisingly large variation in the λ_{max} values of the L photopigment in eyes of butterflies of the genus *Limenitis*. We therefore decided to concentrate our efforts on evaluating whether positive selection has driven the spectral diversification of these closely related photopigments. Using molecular, population-genetic, and molecular evolutionary approaches, we identified several candidate spectral tuning sites in the chromophore-binding pocket of the butterfly L opsin protein. Our combined approaches suggest that similar amino acid sites are indeed involved in the evolution of color vision in both primates and butterflies.

Results and Discussion

Nymphalid Butterflies Comprise a Useful System for Studying Color Vision. Of the five families of butterflies, we have focused our studies on the diverse family Nymphalidae, because it contains a number of species that are model systems in butterfly ecology and evolutionary biology, such as the monarch butterfly *Danaus plexippus* (mimicry and migration) (29, 30), *Heliconius erato* (speciation) (31, 32), and *Bicyclus anynana* (evolution and development) (33). Likewise, there are ample anatomical, molecular, and behavioral data for color vision in nymphalids.

Anatomically, the basic unit of the butterfly compound eye is the ommatidium, composed of eight elongate photoreceptor cells (R1–R8) and a small basal ninth cell (R9) (Fig. 1 *A* and *B*). The photopigment-containing microvillar membranes produced by all photoreceptor cells of an ommatidium are fused into a single
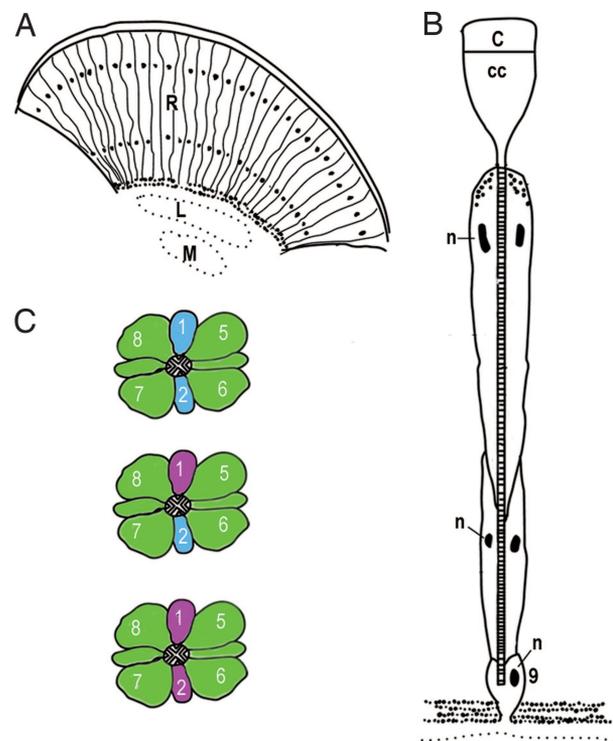

Fig. 1. Butterfly compound eye and opsin expression patterns. (*A*) Diagram of a longitudinal section through the compound eye showing the ommatidial units. Black dots indicate location of photoreceptor nuclei. R, retina; L, lamina; M, medulla. (*B*) Schematic of an ommatidium. C, cornea; CC, crystalline cone; n, nuclei; 9, the ninth photoreceptor cell that sits just above the basement membrane. (*C*) Opsin mRNA expression patterns. The cross-sections of three ommatidia are shown. The cross-hatched area in the middle of each depicts the fused microvillous membranes of the rhabdomeres that contain the visual pigment proteins. Numbers refer to the photoreceptor cells (R1–R8), and the colors refer to the opsin expression patterns: violet, UV opsin mRNA; blue, blue opsin; green, long-wavelength opsin. Modified from Sauman *et al.* (35).

optical-sensing structure known as a rhabdom. Molecular studies in *D. plexippus*, *H. erato*, and *Vanessa cardui* have shown that the photopigments are encoded by two S opsin (UV, ultraviolet; B, blue) and one L opsin gene present in the ancestor of all butterflies and moths. In the main retina of these species, the R1 and R2 photoreceptor cells express either the UV or B opsin mRNAs, and the R3–R9 photoreceptors express the L opsin mRNA (Fig. 1*C*) (26, 34, 35). Some nymphalid butterflies (e.g., *H. erato* and monarchs) also have colored "lateral" filtering pigments that coat the rhabdom and modify the wavelengths of light available to photoisomerize the photopigments, whereas others do not (e.g., *Vanessa atalanta*).

Behavioral studies have shown that *H. erato*, although it expresses UV, B, and only one L opsin, discriminates colors in the blue to L range tested (440–640 nm), whereas *V. atalanta*, despite having color vision in the 440- to 590-nm range, is unable to discriminate between longer-wavelength (590–640 nm) colors (26). Thus, these two nymphalid butterflies display color vision, with the difference in color vision range between them being due to the presence or absence of heterogeneously expressed filtering pigments.

Absorbance Spectra and Opsin Sequences of *Limenitis* L Photopigments. To test for evidence of positive selection of spectral tuning sites in nymphalid L photopigments, we first searched for L photopigments whose λ_{max} values varied among closely related species. The λ_{max} values of L photopigments from a number of nymphalid species were evaluated by using epimicrospectrophotometry, a method (36) that takes advantage of a mirrored tapetum
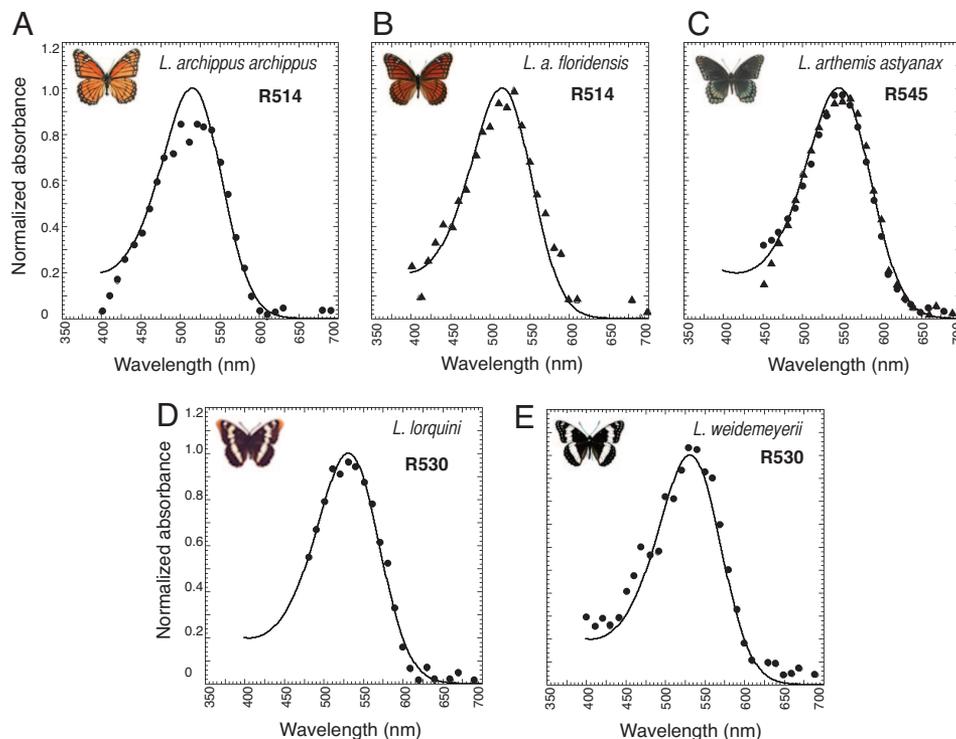
Fig. 2. Normalized absorbance spectra of L photopigments in the *Limenitis* genus measured by epimicrospectrophotometry. Idealized spectra (solid curves) based on the 1987 Bernard template (56). λ_{max} values are shown in upper right corner. R545 from *L. arthemis astyanax*: dots represent dorsal retina, and triangles represent ventral retina.

underlying each ommatidium of the nymphalid eye to resolve peak absorbances of the expressed photopigments [see *Materials and Methods* and supporting information (SI) *Methods*). Using epimicrospectrophotometry, we found that the λ_{max} values of the L photopigments of the five *Limenitis* species ranged from 514 nm in *Limenitis archippus archippus* (36) and *Limenitis archippus floridensis*, to 530 nm in *Limenitis lorquini* and *Limenitis weidmeyerii*, to 545 nm in *Limenitis arthemis astyanax* (Fig. 2). A spectral range of 31 nm within a single butterfly genus is remarkable compared with the other nymphalid genera we surveyed and is similar to the 33-nm range observed for the human green and red cone pigments. This unusual spectral range among the five L pigments within the genus *Limenitis* provided us with a unique collection of photopigments for further analysis.

To evaluate candidate spectral tuning sites, we next determined the opsin gene sequences that correspond to the encoded L photopigment spectral phenotypes. We thus cloned the full-length L opsin cDNA from *L. arthemis astyanax* (GenBank accession no. AY918903) and found no evidence of gene duplication in this species, as has been found in other butterflies (25, 37). We also cloned partial L opsin gene sequences from each of the other four North American *Limenitis* species. The gene sequences ranged from 1366 to 1391 bp in length and were trimmed of introns. For all five L opsin sequences, only transmembrane domains I–VI were used in subsequent analyses, because this region includes the chromophore binding pocket (SI Fig. 5).

There was a robust correlation between the spectral phenotypes and the number of amino acid substitutions of the L opsin proteins of *Limenitis*. Accordingly, the largest number of amino acid substitutions was observed between the two pigments with the largest spectral difference (31 nm, *L. archippus archippus*–*L. arthemis astyanax*, 15/263 substitutions), whereas the smallest number of substitutions was between the pigments with the smallest spectral difference (0 nm, *L. archippus archippus*–*L. archippus floridensis*,

3/263 substitutions; 0 nm, *L. lorquini*–*L. weidemeyerii*, 0/263 substitutions).

Population Genetic and Molecular Evolutionary Analyses Suggest L Opsins Are Under Positive Selection. The large range of variation in the λ_{max} values among the *Limenitis* L photopigments (514, 530, and 545 nm) is strikingly similar to that observed in the three-allele system of New World monkeys (≈530, 545, and 560 nm) (6) and suggests that the *Limenitis* L photopigments may also be maintained by positive selection. We therefore undertook four population genetic and molecular evolutionary approaches, the McDonald–Kreitman (MK) test, character mapping, parallel/convergent change analysis, and branch-site test of selection, to define specific spectral tuning sites that may be under positive selection.

MK Test for Selection. Differences in the amino acid sequences of proteins among species may be due to the accumulation of neutral mutations by drift, the fixation of adaptive mutations by positive selection, or a combination of the two. We therefore used the MK test (38) to examine whether the amino acid differences between the *L. archippus archippus* and *L. arthemis astyanax* pigments had evolved by neutral evolution. Although some demographic scenarios may result in the MK test erroneously indicating adaptive evolution (39), generally this test is free of demographic concerns, because neutral and selected sites interspersed throughout a gene share the same phylogeny and have the same effective population size (38, 40).

For the MK test, we genotyped 24 *L. arthemis astyanax* individuals for a region (1,056–1,069 bp) on both chromosomes and identified five alleles and 11 polymorphic sites, compared with the *L. archippus archippus* sequence (Table 1). We found that the ratio of nonsynonymous to synonymous fixed differences between species (i.e., the number of substitutions that produce an amino acid change compared with the number of substitutions that do not) deviated strongly from the ratio of nonsynonymous to synonymous

Table 1. Allelic variation in 597 bp of coding region of L opsin gene in 24 *L. arthemis astyanax* individuals (48 chromosomes) and one *L. archippus archippus*

Allele*	N	16	51	99	120	184	186	190	192	205	209	216	245	265	285	291	345	366	384	409	415	438	509	529	554	582	588	592
1	28	A	C	C	T	T	C	G	G	G	A	T	C	G	T	T	A	C	C	T	T	C	C	T	T	C	C	G
2	4	-	-	-	-	G	A	T	-	-	-	-	-	-	-	C	G	T	A	T	-	-	-	-	-	-	-	-
3	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-
4	5	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	T	-	-	-
L. archippus archippus	2	G	G	-	A	G	A	T	T	A	G	C	G	A	C	G	T	-	-	G	G	-	T	G	C	-	T	A
Amino acid		N	I			V	V	A		V	N		A	A						S	S		A	F	M			V
replacements		↓	↓			↓	↓	↓		↓	↓		↓	↓						↓	↓		↓	↓	↓			↓
		D	M			F	F†	S		I	S		G	T						A	A		V	V	T			I
Position		16	17			62	62	64		69	70		82	89						137	139		169	176	184			197

A dash (-) indicates identity with the most common allele. Amino acid replacements are given in single-letter notation, and position refers to the numbering of amino acid residues in the alignment shown in SI Fig. 5.
*GenBank accession nos. EF156437–EF156441.
†A substitution at nucleotide position 186 may lead to a replacement (V→L) but was never observed.

polymorphisms within species (Fisher's exact test, two-tailed $P = 0.006$) (Table 2). The higher proportion of fixed nonsynonymous substitutions (12) compared with fixed synonymous substitutions (4) rejected the hypothesis of neutral evolution and suggested that most of the observed replacement substitutions are probably due to positive selection-driven fixation of advantageous mutations. The results of this analysis indicate that the *L. archippus* and *L. arthemis astyanax* L opsin genes have evolved under positive selection for divergent functions. This allowed us to focus on specific sites within the encoded protein that are under positive selection and involved in spectral tuning.

Character Mapping of L Opsin λ_{max} Values. Methodological advances have been made in recent years in detecting positive selection at particular amino acids or codon sites. These statistical approaches depend on ancestral state reconstructions of either amino acid or nucleotide sequences at different branching points (nodes) along a gene tree and are greatly strengthened by the availability of functional data mapped onto the tree, which permits the *a priori* selection of branches for investigation. These tests are particularly compelling if convergent phenotypes have evolved along one or more independent lineages of the tree (41).

To test this possibility, we first reconstructed an opsin gene tree by using L opsin gene data from 12 nymphalid, one papilionid, and one pierid butterfly species for which λ_{max} values are also available (SI Table 4), by using maximum-likelihood and Bayesian algorithms. Both methods recovered identical trees with good bootstrap support (>50%) in all except the basal node (Fig. 3). Our phylogeny is in general agreement with the most recent phylogeny of butterflies based on molecular and morphological data (42).

A character map was constructed by mapping the λ_{max} values onto the gene tree. Visual inspection of the distribution of λ_{max}

values on the phylogeny revealed two instances of spectral convergence of L pigment λ_{max} values. Two pigment lineages, one leading from *L. arthemis astyanax* to other *Limenitis* species (Fig. 3, Nodes A to B) and the lineage ancestral to the *Siproeta stelenes* and *Junonia coenia* clade (Fig. 3, Nodes C to D), have evolved peak spectral sensitivities that are blue-shifted (i.e., shifted to shorter-wavelength light) compared with ancestral nodes.

Parallel/Convergent Change Analysis. Using maximum-likelihood and maximum-parsimony ancestral state reconstructions of amino acid sequences, we next asked whether statistically significant parallel and/or convergent amino acid changes occurred along the two blue-shifted opsin lineages (i.e., along Nodes A to B and C to D) (43, 44). Whether a substitution is a parallel or convergent
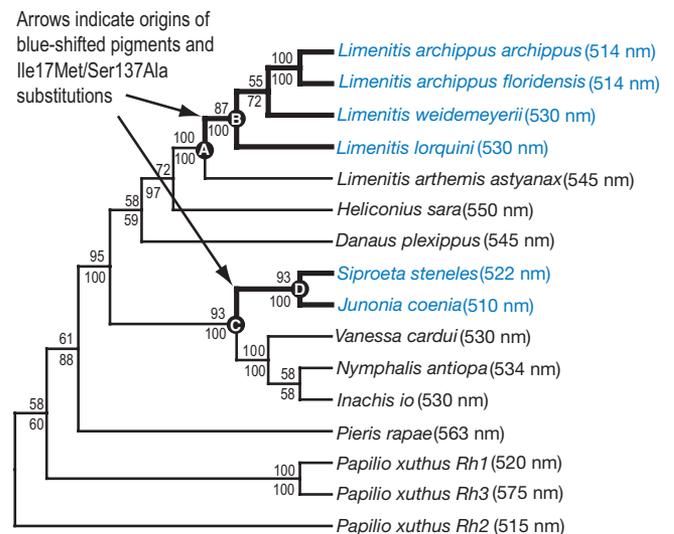

Fig. 3. Character mapping of L photopigment λ_{max} values onto an L opsin gene tree. Numbers shown above branch indicate maximum-likelihood bootstrap support values and numbers below branch show Bayesian clade credibility values. Thick black lines indicate nymphalid branches along which blue spectral shifts occurred that were investigated in the parallel/convergent change and branch-site tests of selection. Nodes in which ancestral state reconstructions shown in Table 3 were performed are indicated by letters.

Table 2. MK test results for the L opsins

	Synonymous	Nonsynonymous	P value
Polymorphic*	9	2	
Fixed divergent†	4	12	0.0063‡

*Number of polymorphic sites among 24 individuals of *L. arthemis astyanax* genotyped for 597 bp of L opsin coding region.
†Divergent with respect to *L. archippus archippus*.
‡Fisher's exact test, two-tailed *P* value.

Table 3. Ancestral state reconstructions along blue-shifted butterfly L opsin lineages shown in the opsin gene tree (Fig. 3)

Amino acid site	545 nm Node A	514–530 nm Node B	Amino acid site	530 nm Node C	510–522 nm Node D	Amino acid site	510–522 nm Node D	510 nm J. coenia
6	N 0.998	D 1.000	16	V 0.704	F 0.541	8	L 0.985	M
17	**I 0.998**	**M 1.000**	17	**I 0.997**	**M 1.000**	20	L 0.998	I
64	A 0.996	S 1.000	18	G 0.945	A 0.747	26	T 0.981	A
70	N 0.998	S 1.000	131	L 0.994	F 0.981	40	T 0.882	S
137	**S 1.000**	**A 1.000**	137	**S 1.000**	**A 1.000**	61	F 0.541	T
			258	A 0.640	S 0.717	**64**	**A 0.522**	**S**
						70	**N 0.912**	**S**
						93	F 0.997	C
						180	V 0.758	L

Parallel amino acid changes are indicated in bold. Numbers above nodes indicate inferred λ_{max} values. Numbers after amino acid residues indicate Bayes Empirical Bayes posterior probabilities.

change depends on the reconstructed amino acid in the ancestral sequences. For example, a parallel change would occur if both ancestral sequences contained the same amino acid residue at a site, whereas a convergent change would occur if the ancestral sequences contained different amino acid residues at a site. Convergent changes are statistically less likely and stronger signatures of selection.

We observed two instances of parallel change associated with blue spectral shifts by using both reconstructions. At amino acid residue 17, an isoleucine-to-methionine substitution was observed, and at residue 137, a serine-to-alanine substitution was observed. The number of observed substitutions was significantly greater than expected by chance ($P < 0.01$) (Table 3). When the ancestral state reconstruction was expanded to include substitutions that occurred along the terminal branch connecting node D to the even more blue-shifted *J. coenia* opsin, two further parallel amino acid changes were observed associated with parallel blue spectral shifts: a alanine-to-serine substitution at amino acid residue 64 and an asparagine-to-serine substitution at amino acid residue 70.

Branch-Site Models of Selection. Although parallel/convergent change analysis depends on reconstruction of amino acid sequences, a much stronger positive selection inference can be made by using reconstruction of nucleotide sequences. One of the signatures of positive selection on a particular codon site is a much higher rate of substitution, causing amino acid replacements, d_N, than the rate of silent substitutions, d_S, leading to the ratio (ω) of $d_N/d_S > 1$ (45). Using branch-site models of selection, we tested whether specific amino acid sites along the two identified blue-shifted lineages are evolving in a manner consistent with positive selection. For the two blue-shifted lineages tested, the alternative model (specifically testing for positive selection) was a significantly better fit ($2\Delta\lambda = 3.951$, df $= 1$, $P < 0.05$) than the null model. The analysis indicated that seven sites were likely to be under positive selection (sites 6, 17, 18, 64, 70, 131, and 137). However, only three (17, 64, and 137; Fig. 4 and SI Fig. 5) had Bayes Empirical Bayes posterior probabilities >0.85, with two sites, 17 and 137, having posterior probabilities that were >0.95 (data not shown).

This result is provocative, because previous attempts to identify amino acid sites under positive selection in both butterfly and vertebrate photopigments by using similar methods have largely failed despite robust physiological data (i.e., absorbance spectra), suggesting functional divergence and the observation of parallel evolution (refs. 5 and 24, but see ref. 4). Remarkably, four of the identified sites (17, 64, 70, and 137) were the same as those evolving in a manner consistent with positive selection in the parallel/convergent change analysis.

Homology Modeling Places L Opsin Sites Under Selection Close to the Chromophore. A homology model based on the *L. arthemis astyanax* sequence (Fig. 4) indicated that four of the sites under selection in

the two blue-shifted lineages, sites 17, 64, 70, and 137, directly face the chromophore (Fig. 4). In particular, sites 17, 64, and 70 are situated close to the Schiff-base end of the chromophore. The amino acid sites Ile17Met, Ala64Ser, Asn70Ser, and Ser137Ala in the *Limenitis* sequences correspond to Met[44], Phe[91], Thr[97], and Ala[164] on bovine rhodopsin, respectively. Met[44] is part of the region surrounding the Schiff base of the 11-cis-retinal chromophore (46, 47), and a Met44Thr substitution in bovine rhodopsin causes a 3-nm blue shift (48). Phe[91] produces a small spectral tuning effect when the equivalent site (Asp[71]) is mutagenized in squid retinochrome, which binds the retinal chromophore and functions as a retinal photoisomerase in squid photoreceptor cells (49). Site 137 corresponds to Ala[164] in bovine rhodopsin or Ser[180]/Ala[180] in human L/M
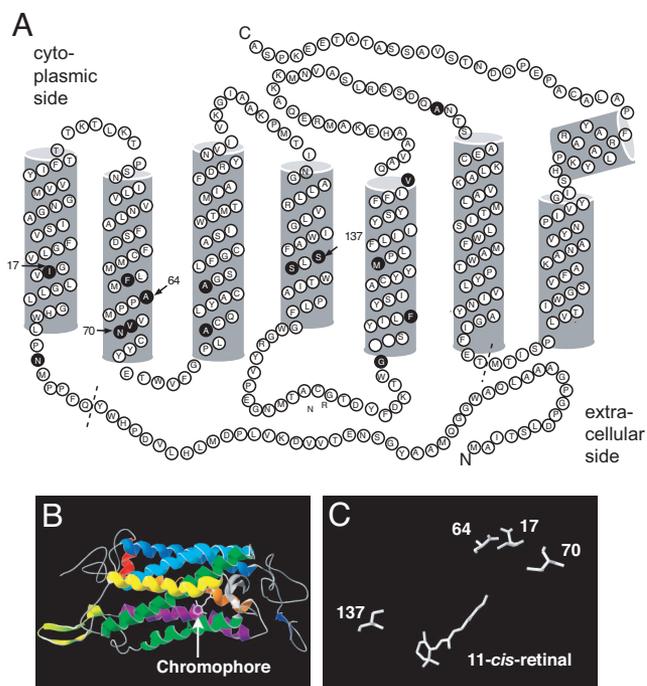

Fig. 4. Topographical and homology models of *L. arthemis astyanax* L opsin. (*A*) Topographical opsin map showing amino acid differences (filled circles) between *L. arthemis astyanax* and *L. archippus archippus* opsins. The diagram is based on a model of bovine rhodopsin (46). Dashed line indicates boundaries of alignment used in phylogenetic analyses. The amino acid residues are numbered relative to the alignment shown in SI Fig. 5. Black arrows, amino acid sites under positive selection. (*B*) Homology model of *L. arthemis astyanax* L opsin, with chromophore shown (arrow). (*C*) Candidate spectral tuning sites that directly face the chromophore (11-cis-retinal).

pigments. Intriguingly, in human cone pigments, the substitution Ser164Ala causes a 5- to 7-nm blue shift (50, 51) and is one of the sites polymorphic among the New World monkey alleles and under balancing selection. Therefore, the correspondence between sites implicated previously in blue spectral shifts in vertebrate opsins and those found in butterflies in this study strongly suggests that some of the same amino acid sites are involved in the evolution of L color vision in both primates and butterflies.

Conclusions

Expanded, or trichromatic, color vision in some primate lineages evolved through gene duplication and diversification of an ancestral polymorphic M/L pigment gene. Expanded color vision is advantageous over a more restricted color vision range in individuals foraging for ripe fruits and young leaves, especially under dim light (52–54). We have shown that, like the polymorphism observed in New World monkeys, the photopigment genes of *Limenitis* have evolved under positive selection. Remarkably, we have further shown that two closely related butterfly species, *L. archippus* and *L. arthemis astyanax* (which can hybridize in nature), show a commonly observed primate polymorphism. Like the primate cone pigments that contain this substitution, the butterfly photopigments are blue-shifted in sensitivity. This suggests that genetic mechanisms of spectral tuning may have evolved in parallel across 500 million years of evolution.

Materials and Methods

Epimicrospectrophotometry. The technique was used as described (36, 55) (a more extensive description is provided in *SI Methods*). Briefly, the visual pigments of a butterfly ommatidium are contained within a rhabdom waveguide. Light propagates down the waveguide, being absorbed by rhodopsins as it goes. Light reaching the end of the rhabdom is reflected by the tapetum, a specialized tracheolar layer, back out of the eye as eyeshine. The reflectance spectrum of eyeshine was measured in dark-adapted live insects, with a series of dim monochromatic flashes set as a reference spectrum. Reflectance spectra were then measured after partial bleaching of the rhabdom of its L visual pigment after a series of photoisomerizing flashes. The rhodopsin photoproduct-free difference spectrum yields a two-way absorbance spectrum. An estimate of λ_{max} is then obtained by least-squares fitting to template absorbance spectra (56).

Sample Collecting. All *Limenitis* specimens were collected as adults in the field (*L. archippus archippus*, Goose Lake, Dane County, Wisconsin; *L. archippus floridensis*, Yankeetown, Florida, and Archbold Biological Station, Lake Placid, Florida; *L. lorquini*, Ash Canyon, Carson City County, Nevada; *L. weidemeyerii*, Willow Creek, Catron County, New Mexico, and Kelber Pass, Colorado; and *L. arthemis axtyanax*, Due West, South Carolina, $n = 2$; Madison, WI, $n = 1$; Amherst, MA, $n = 1$; and Lowndes County, Georgia, $n = 1$). The remaining 19 *L. arthemis astyanax* specimens used for the population sample were gifts from Austin Platt and were collected in Green Ridge State Forest, Maryland.

PCR, Cloning, and Sequencing. *L. arthemis astyanax* cDNA was synthesized from total RNA from one frozen head by using the Marathon cDNA Amplification Kit (BD Biosciences Clontech, Mountain View, CA). The full-length L opsin cDNA, including 3′ and 5′ UTRs, was obtained by amplification of 3′ RACE products by using a degenerate primer (5′-GAA CAR GCW AAR AAR ATG A-3′) followed by amplification of 5′ RACE products by using a gene-specific reverse primer (5′-CAG AGC CCC AAA TGG TCA CTA A-3′). The resulting products were cloned into the pGEM T-Easy Vector System II (Promega, Madison, WI) and sequenced at the University of California, Irvine, DNA core sequencing facilities.

Genomic DNA from the other *Limenitis* species was extracted from individual adult butterflies (one per species) by using phenol-chloroform. First, a highly conserved 300-bp part of the gene was amplified by the PCR by using degenerate primers 5′-GAA CAR GCW AAR AAR ATG A-3′ and 5′-CCR TAN ACR ATN GGR TTR TA-3′, which was cloned and sequenced as above. These sequences allowed the design of species-specific reverse primers that were used in combination with a degenerate forward primer (5′-CAY YTN ATH GAY CCN CAY TGG-3′) to amplify a 1,366- to 1,391-bp fragment. For the *L. arthemis astyanax* population sample, PCR products were then directly sequenced. Heterozygous individuals were identified by visual inspection. Haplotypes were confirmed by cloning into the pGEM T-easy vector as described above.

For the MK test (38), introns were manually removed from the 1,056- to 1,069-bp region of the 24 *L. arthemis astyanax* and one *L. archippus archippus* individuals genotyped, leaving 597 bp of coding region that was aligned. The number of sites within the *L. arthemis astyanax* population sample that was polymorphic for synonymous and nonsynonymous substitutions was counted and compared with the number of sites that was fixed between species for synonymous and nonsynonymous substitutions by using a Fisher's exact test.

Phylogenetic Reconstruction. Sequences for *Limenitis* species were added to other L opsin sequences obtained from GenBank for which physiological data were also available (SI Table 4). The resulting alignment was edited to retain only coding sequence, 263 aa in length, spanning transmembrane domains I–VI (Fig. 4 and SI Fig. 5). Phylogenetic relationships were reconstructed from nucleotide data by using maximum-likelihood and Bayesian methods, with the moth *Manduca sexta* L opsin sequence as an outgroup by using all three nucleotide positions. The optimal DNA substitution model for the maximum-likelihood phylogenetic analysis was determined by using a hierarchical likelihood ratio test in Modeltest (57). A maximum-likelihood analysis was conducted in PHYML with TrN93 + I + G (invariant sites and gamma-distributed rates for sites) substitution model, and the reliability of the tree obtained was tested by bootstrapping with 500 replicates. Bayesian phylogenetic analyses were performed by using MrBayes 3.1 (58). Because MrBayes 3.1 does not implement the TrN93 DNA substitution model, we used the next-less-complex (HKY85+I+G) and the next-more-complex (GTR+I+G) models in two separate analyses. Both models were run for 10^6 generations, with a sampling frequency of 10^2, using three heated and one cold chain and with a burnin of 2.5×10^3 trees.

Parallel Change Analysis and Branch-Site Tests of Selection. The Bayesian and maximum-likelihood tree topology obtained was used to perform all selection tests.

For the parallel change analysis, amino acid substitutions along each lineage in the opsin gene tree were reconstructed by using maximum parsimony in MacClade (59) and maximum likelihood in PAML (60). Parallel amino acid changes were detected along butterfly lineages that also displayed parallel phenotypic evolution in the L visual pigment λ_{max} value. The statistical significance of these amino acid changes was tested by using the method (44) implemented in the program Converge.

For the branch-site test of selection, we used branch-site models (61, 62) that allow the d_N/d_S ratio (ω) to vary both among sites and among lineages because these models may be more likely to detect positive selection affecting only a few sites. Specifically, we used test 2 of ref. 62 to construct a likelihood ratio test with Model A and Model A_1 (null model). Both models require *a priori* specification of the lineages likely to have experienced positive selection. Visual inspection of Fig. 3 allowed us to test two blue-shifted lineages that may have experienced positive selection. Sites that may be under positive

selection in the spectrally shifted lineages were identified with a Bayes Empirical Bayes approach (63).

Homology Modeling. We used homology modeling to study the relationship between the *L. arthemis astyanax* opsin structure and function using the methods described (64). The full-length *L. arthemis astyanax* L opsin protein sequence was manually aligned with the bovine template 1U19.pdb (47), and the alignment was submitted to the Swiss-Model server (www.expasy.ch/swissmod) (65). The atomic coordinates were then viewed with SwissPdb Viewer (www.expasy.ch/spdbv) (66), and candi-date spectral tuning sites were mapped onto the 3D homology model.

1. Land MF, Fernald RD (1992) *Annu Rev Neurosci* 15:1–29.
2. Han M, Groesbeek M, Smith SO, Sakmar TP (1998) *Biochemistry* 37:538–545.
3. Yokoyama S (1997) *Annu Rev Genet* 31:311–332.
4. Spady TC, Seehausen O, Loew ER, Jordan RC, Kocher TD, Carleton KL (2005) *Mol Biol Evol* 22:1412–1422.
5. Yokoyama S, Takenaka N (2005) *Mol Biol Evol* 22:968–975.
6. Surridge AK, Osorio D, Mundy NI (2003) *Trends Ecol Evol* 18:198–205.
7. Kelber A (1999) *J Exp Biol* 202:2619–2630.
8. Kelber A (2001) *Entomol Exp Appl* 99:231–244.
9. Mollon JD (1989) *J Exp Biol* 146:21–38.
10. Terakita A (2005) *Genome Biol* 6:213.
11. Arendt D (2003) *Int J Dev Biol* 47:563–571.
12. Jacobs GH (1993) *Biol Rev Camb Philos Soc* 68:413–471.
13. Kolb G, Scherer C (1982) *J Comp Physiol* 149:325–332.
14. Kelber A, Henique U (1999) *J Comp Physiol A* 184:535–541.
15. Kelber A, Vorobyev M, Osorio D (2003) *Biol Rev Camb Philos Soc* 78:81–118.
16. Ahnelt PK, Kolb H (2000) *Prog Retin Eye Res* 19:711–777.
17. Tan Y, Yoder AD, Yamashita N, Li WH (2005) *Proc Natl Acad Sci USA* 102:14712–14716.
18. Tan Y, Li WH (1999) *Nature* 402:36.
19. Mollon JD, Bowmaker JK, Jacobs GH (1984) *Proc R Soc London* 222:373–399.
20. Kainz PM, Neitz J, Neitz M (1998) *Vision Res* 38:3315–3320.
21. Dulai KS, von Dornum M, Mollon JD, Hunt DM (1999) *Genome Res* 9:629–638.
22. Kinoshita M, Shimada N, Arikawa K (1999) *J Exp Biol* 202:95–102.
23. Briscoe AD (2000) *J Mol Evol* 51:110–121.
24. Briscoe AD (2001) *Mol Biol Evol* 18:2270–2279.
25. Kitamoto J, Sakamoto K, Ozaki K, Mishina Y, Arikawa K (1998) *J Exp Biol* 201:1255–1261.
26. Zaccardi G, Kelber A, Sison-Mangus MP, Briscoe AD (2006) *J Exp Biol* 209:1944–1955.
27. Bowmaker JK, Hunt DM (1999) in *Adaptive Mechanisms in the Ecology of Vision*, eds Archer SN, Djamgoz MBA, Loew ER, Partridge JC, Vallerga S (Kluwer, Dordrecht, The Netherlands), pp 439–462.
28. Briscoe AD, Chittka L (2001) *Annu Rev Entomol* 46:471–510.
29. van Zandt Brower J (1958) *Evolution (Lawrence, Kans)* 12:32–47.
30. Reppert SM (2006) *Cell* 124:233–236.
31. Brower AVZ (1994) *Proc Natl Acad Sci USA* 91:6491–6495.
32. McMillan WO, Jiggins CD, Mallet J (1997) *Proc Natl Acad Sci USA* 94:8628–8633.
33. Beldade P, Brakefield PM, Long AD (2005) *Evol Dev* 7:101–107.
34. Briscoe AD, Bernard GD, Szeto AS, Nagy LM, White RH (2003) *J Comp Neurol* 458:334–349.
35. Sauman I, Briscoe AD, Zhu H, Shi D, Froy O, Stalleicken J, Yuan Q, Casselman A, Reppert SM (2005) *Neuron* 46:457–467.
36. Bernard GD (1983) *Biophys Struct Mech* 9:277–286.
37. Briscoe AD (1998) *Naturwissenschaften* 85:33–35.
38. McDonald JH, Kreitman M (1991) *Nature* 351:652–654.
39. Eyre-Walker A (2002) *Genetics* 162:2017–2024.
40. Nielsen R (2005) *Annu Rev Genetics* 39:197–218.
41. Zhang J (2006) *Nat Genet* 38:819–823.
42. Wahlberg N, Braby MF, Brower AVZ, de Jong R, Lee MM, Nylin S, Pierce NE, Sperling FAH, Vila R, Warren AD, *et al.* (2005) *Proc R Soc London Ser B* 272:1577–1586.
43. Kreitman M, Akashi H (1995) *Annu Rev Ecol Syst* 26:403–422.
44. Zhang J, Kumar S (1997) *Mol Biol Evol* 14:527–536.
45. Yang ZH, Bielawski JP (2000) *Trends Ecol Evol* 15:496–503.
46. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, *et al.* (2000) *Science* 289:739–745.
47. Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V (2004) *J Mol Biol* 342:571–583.
48. Andres A, Garriga P, Manyosa J (2003) *Biochem Biophys Res Commun* 303:294–301.
49. Terakita A, Yamashita T, Shichida Y (2000) *Proc Natl Acad Sci USA* 97:14263–14267.
50. Asenjo AB, Rim J, Oprian DD (1994) *Neuron* 12:1131–1138.
51. Merbs SL, Nathans J (1993) *Photochem Photobiol* 58:706–710.
52. Smith AC, Buchanan-Smith HM, Surridge AK, Osorio D, Mundy NI (2003) *J Exp Biol* 206:3159–3165.
53. Caine NG, Mundy NI (2000) *Proc R Soc London Ser B* 267:439–444.
54. Osorio D, Smith AC, Vorobyev M, Buchanan-Smith HM (2004) *Am Nat* 164:696–708.
55. Briscoe AD, Bernard GD (2005) *J Exp Biol* 208:687–696.
56. Palacios AG, Goldsmith TH, Bernard GD (1996) *Visual Neurosci* 13:411–421.
57. Posada D, Crandall KA (1998) *Bioinformatics* 14:817–818.
58. Ronquist F, Huelsenbeck JP (2003) *Bioinformatics* 19:1572–1574.
59. Maddison DR, Maddison WP (2005) MacClade (Sinauer, Sunderland, MA), Ver. 4.06.
60. Yang Z (1997) *Comput Appl Biosci* 13:555–556.
61. Yang Z, Nielsen R (2002) *Mol Biol Evol* 19:908–917.
62. Zhang J, Nielsen R, Yang Z (2005) *Mol Biol Evol* 22:2472–2479.
63. Yang Z, Wong WSW, Nielsen R (2005) *Mol Biol Evol* 22:1107–1118.
64. Briscoe AD (2002) *Mol Biol Evol* 19:983–986.
65. Schwede T, Kopp J, Guex N, Peitsch MC (2003) *Nucleic Acids Res* 31:3381–3385.
66. Guex N, Peitsch MC (1997) *Electrophoresis* 18:2714–2723.