UC Irvine
UC Irvine Previously Published Works

Title
Early Duplication and Functional Diversification of the Opsin Gene Family in Insects

Permalink
https://escholarship.org/uc/item/2pk85368

Journal
Molecular Biology and Evolution, 21(8)

ISSN
0737-4038

Authors
Spaethe, Johannes
Briscoe, Adriana D

Publication Date
2004-08-01

DOI
10.1093/molbev/msh162

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Early Duplication and Functional Diversification of the Opsin Gene Family in Insects

Johannes Spaethe and Adriana D. Briscoe

Comparative and Evolutionary Physiology Group, Department of Ecology and Evolutionary Biology, University of California, Irvine

Recent analysis of the complete mosquito *Anopheles gambiae* genome has revealed a far higher number of opsin genes than for either the *Drosophila melanogaster* genome or any other known insect. In particular, the analysis revealed an extraordinary opsin gene content expansion, whereby half are long wavelength–sensitive (LW) opsin gene duplicates. We analyzed this genomic data in relationship to other known insect opsins to estimate the relative timing of the LW opsin gene duplications and to identify "missing" paralogs in extant species. The inferred branching patterns of the LW opsin gene family phylogeny indicate at least one early gene duplication within insects before the emergence of the orders Orthoptera, Mantodea, Hymenoptera, Lepidoptera, and Diptera. These data predict the existence of one more LW opsin gene than is currently known from most insects. We tested this prediction by using a degenerate PCR strategy to screen the hymenopteran genome for novel LW opsin genes. We isolated two LW opsin gene sequences from each of five bee species, *Bombus impatiens*, *B. terrestris*, *Diadasia afflicta*, *D. rinconis*, and *Osmia rufa*, including 1.1 to 1.2 kb from a known (*LW Rh1*) and 1 kb from a new opsin gene (*LW Rh2*). Phylogenetic analysis suggests that the novel hymenopteran gene is orthologous to *A. gambiae GPRop7*, a gene that is apparently missing from *D. melanogaster*. Relative rate tests show that LW Rh2 is evolving at a slower rate than LW Rh1 and, therefore, may be a useful marker for higher-level hymenopteran systematics. Site-specific rate tests indicate the presence of several amino acid sites between LW Rh1 and LW Rh2 that have undergone shifts in selective constraints after duplication. These sites and others are discussed in relationship to putative structural and functional differences between the two genes.

Introduction

Insect compound eyes contain light-sensitive receptors in which phototransduction takes place. The light-sensitive part of the receptor is the visual pigment (rhodopsin), which is embedded in specialized parts of the membrane of the retinula cell. It is usually composed of a chromophore derived from 11-*cis* retinal (Gleadall, Hariyama, and Tsukahara 1989; Seki and Vogt 1998) and an opsin apoprotein to which the chromophore is attached (Filipek et al. 2003). Isolated 11-*cis* retinal is most sensitive to light of about 380 nm (Han et al. 1998), but its absorption maximum can be tuned by the opsin over a wide spectral range, from about 360 to 640 nm (Kochendoerfer et al. 1999). On the basis of their physiological properties, insect photoreceptors can be assigned to three spectral classes: short wavelength–sensitive (maximally sensitive for light less than 400 nm, the UV part of the spectrum), medium wavelength–sensitive (400 to 500 nm, blue light) and long wavelength (LW)–sensitive (>500 nm, green light). The spectral sensitivity of photoreceptors is basically determined by their visual pigments but can also be affected by the presence of screening or sensitizing pigments (Gärtner 2000). Most insect species have all three spectral receptor types in the retina.

Phylogenetic studies of the protein sequences of insect visual pigments have revealed that opsins fall into four major subfamilies, three of which match the physiologically identified short wavelength–sensitive, medium wavelength–sensitive, and long wavelength–sensitive visual pigment groups (for review, see Briscoe and Chittka [2001]). The blue-green opsins constitute the fourth group, which is most closely related to the long wavelength pigments, and includes such members as the *Drosophila melanogaster* majority pigment, Rh1 and the ocellar-specific, Rh2. The largest expansion of insect opsin genes reported so far has been in the LW opsin group of lepidopterans and dipterans. In the butterfly *Papilio glaucus*, for instance, four LW opsin genes were found but only a single blue and a single UV opsin gene were found (Briscoe 1998, 2000). One of these LW opsins (*PglRh4*) is now known to be expressed exclusively outside of the retina (Briscoe and Nagy 1999), whereas all the others have been localized to the eye (Kitamoto, Ozaki, and Arikawa 2000; Kitamoto et al. 1998). A search for G protein–coupled receptors in the recently published *Anopheles gambiae* genome identified 12 different opsin genes, of which seven belong to the LW *Rh* subfamily (Hill et al. 2002). The spatial distribution of these opsins has yet to be described.

In contrast to lepidopterans and dipterans, only one opsin in each spectral group has been reported in hymenopterans (Bellingham et al. 1997; Chang et al. 1996; Townson et al. 1998). All three of these transcripts were isolated from eye-specific mRNAs. One of these genes, the *LW Rh* gene, is extensively used to reconstruct phylogenetic relationships among hymenopteran tribes, families, and species (Ascher, Danforth, and Ji 2001; Cook et al. 2002; Kawakita et al. 2003; Mardulyn and Cameron 1999) as well as to compare patterns of nucleotide substitution between nuclear and mitochondrial genes (Lin and Danforth 2004). Physiological data, however, suggest that some hymenopterans should have more that one long wavelength–sensitive visual pigment in the retina. Peitsch et al. (1992), for instance, utilized intracellular recordings to determine the spectral properties of the visual receptors of 43 bee and wasp species and found several species in the families Andrenidae, Xiphydriidae, and Tenthredinidae with a fourth receptor type sensitive in the yellow to red part of the light

Key words: photoreceptor, visual pigment, color vision, rhodopsin.

E-mail: abriscoe@uci.edu.

Mol. Biol. Evol. 21(8):1583–1594. 2004
doi:10.1093/molbev/msh162
Advance Access publication May 21, 2004

Molecular Biology and Evolution vol. 21 no. 8 © Society for Molecular Biology and Evolution 2004; all rights reserved.

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019

spectrum. They proposed that these fourth spectral receptors were based on visual pigments with λ_{max} between 570 and 596 nm (Peitsch et al. 1992). Although the observed distribution of these additional receptors favors their recent origin, they may also be the result of a single ancient gene duplication followed by loss in multiple lineages.

We were interested in determining whether the phylogenetic relationships of the recently reported *Anopheles* opsins relative to other insect opsin gene family members would provide evidence of additional hymenopteran opsins. (The *Anopheles* sequences were originally analyzed only with respect to other dipterans; see Hill et al. [2002] Supplementary Material). We, therefore, conducted phylogenetic analyses of all available insect LW opsin sequences and found that the branching pattern of the long wavelength opsin group suggested at least one early duplication of this gene in insects. We tested this hypothesis by conducting a PCR search for new and undetected LW opsin genes in bees, the most extensively investigated group among the hymenopterans. We extracted genomic DNA from *Bombus*, *Diadasia*, and *Osmia* bee species and used various combinations of degenerate primers to target possible novel genes. We found a new gene, which was present in all five species species that we studied (*Bombus impatiens*, *B. terrestris*, *Diadasia afflicta*, *D. rinconis*, and *Osmia rufa*). Phylogenetic analysis revealed that the new gene is paralogous to the already known *LW Rh* gene in hymenopterans and arose from an early gene duplication event within the insect LW-sensitive opsin gene family. We used Gu's (1999, 2001) method for identifying sites in the two bee genes that may be under altered selective constraints after duplication and have identified a number of sites that may be responsible for functional differences between the two genes. We discuss these sites in relationship to three-dimensional structural information inferred from a homology model of one of the hymenopteran opsins, as well as the possible role of the new gene in extraretinal photoreception.

Materials and Methods

Phylogenetic Analyses of Insect LW Opsins

To determine the relationship between the LW-opsin genes reported in the *Anopheles gambiae* genome sequencing project (Holt et al. 2002) and other known insect LW opsins, we conducted a search of both the literature and GenBank (using Blast) for all available members of insect and crustacean blue-green and LW-opsin gene families. Among those sequences included in our analysis were three sequences from the *Drosophila melanogaster* genome, six sequences from the mosquito *Culex pipiens* (Tian et al. 2001), and two sequences from the mosquito *Aedes agypti*, that had been reported in print (Graf et al. 1996) but were not present in GenBank. We selected for inclusion in our initial phylogenetic analysis, sequences that represented complete or near-complete cDNAs so we would have the longest alignment possible for analysis. (We, therefore, left out *Papilio glaucus* PglRh4, which lacks approximately 38 amino acids from the N-terminus). We did, however, check for the possibility of other paralogous genes from these and other taxa, which we might have excluded simply because of their short length, by downloading all available sequences and building trees with smaller alignments and found no other paralogous genes (data not shown). Altogether we included 33 amino acid sequences from 16 different species in eight orders in our phylogenetic analysis (for species names and GenBank accession numbers, see legend of figure 1). Amino acid sequences were aligned using ClustalW (Thompson, Higgins, and Gibson 1994) in the Alignment Explorer in MEGA version 3.0 (Kumar, Tamura, and Nei 2003).

Sequence composition heterogeneity may affect the reconstruction of phylogenetic trees. Therefore we tested 1st + 2nd nucleotide positions, 3rd nucleotide positions, and protein sequences for composition homogeneity among lineages by application of the disparity Index test (Kumar and Gadagkar 2001). Both 1st + 2nd positions and 3rd positions appear to have evolved in a significantly nonhomogeneous fashion, that is, nearly all mosquito opsins were significantly different from all other sequences at 1st + 2nd positions (data not shown). Because only the protein sequences were evolving with homogeneous patterns, amino acid sequences were used in the phylogenetic analysis. Both maximum-parsimony (MP) and neighbor-joining (NJ) analyses were conducted (PAUP* and MEGA 3.0). A total of 271 parsimony-informative amino acid sites were included, and a step matrix of amino acid changes derived from a larger data set of G-protein–coupled receptors (Rice 1994) was employed as a weighting scheme (available as Supplementary Material online) to account for unequal probabilities of amino acid change in the MP analysis. The reliability of the MP trees was tested by bootstrap analysis in PAUP* (Swofford 1998). For the NJ analysis, a total of 350 aligned amino acid sites (complete protein sequence) was used with poisson correction and complete deletion of gaps. Similar analyses were performed by adding to the alignment the 10 new bee gene sequences reported below. However, instead of MP, maximum-likelihood (ML) analysis of nucleotide 2nd positions (GTR + Γ model), as implemented in PhyML (Guindon and Gascuel 2003), and quartet-puzzling analysis of amino acid sequences in PAUP* were performed to verify that the novel *LW Rh2* gene is indeed part of the LW-sensitive opsin class.

DNA Extraction

All specimens were caught in the field except the *B. impatiens* specimen, which was obtained from a colony bought from a commercial breeder (Koppert, distributed by Plant Sciences, Inc., Watsonville, Calif.). The *Diadasia rinconis* and *D. afflicta* specimens were gifts of Jack Neff (table 1). All individuals were frozen in liquid nitrogen, except *Osmia rufa*, which was field collected directly into RNAlater (Ambion). One individual bee per species was used for DNA extraction. The heads were removed and homogenized in 400 μl of 50 mM Tris-Cl (pH 8.0), 20 mM EDTA, and 2% SDS and digested with 2 μl proteinase K. After incubating overnight at 37°C, 100 μl of 5 M NaCl was added and the tubes were placed on ice for 40 min. The tubes were then centrifuged for 15 min at 14,000 rpm

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019

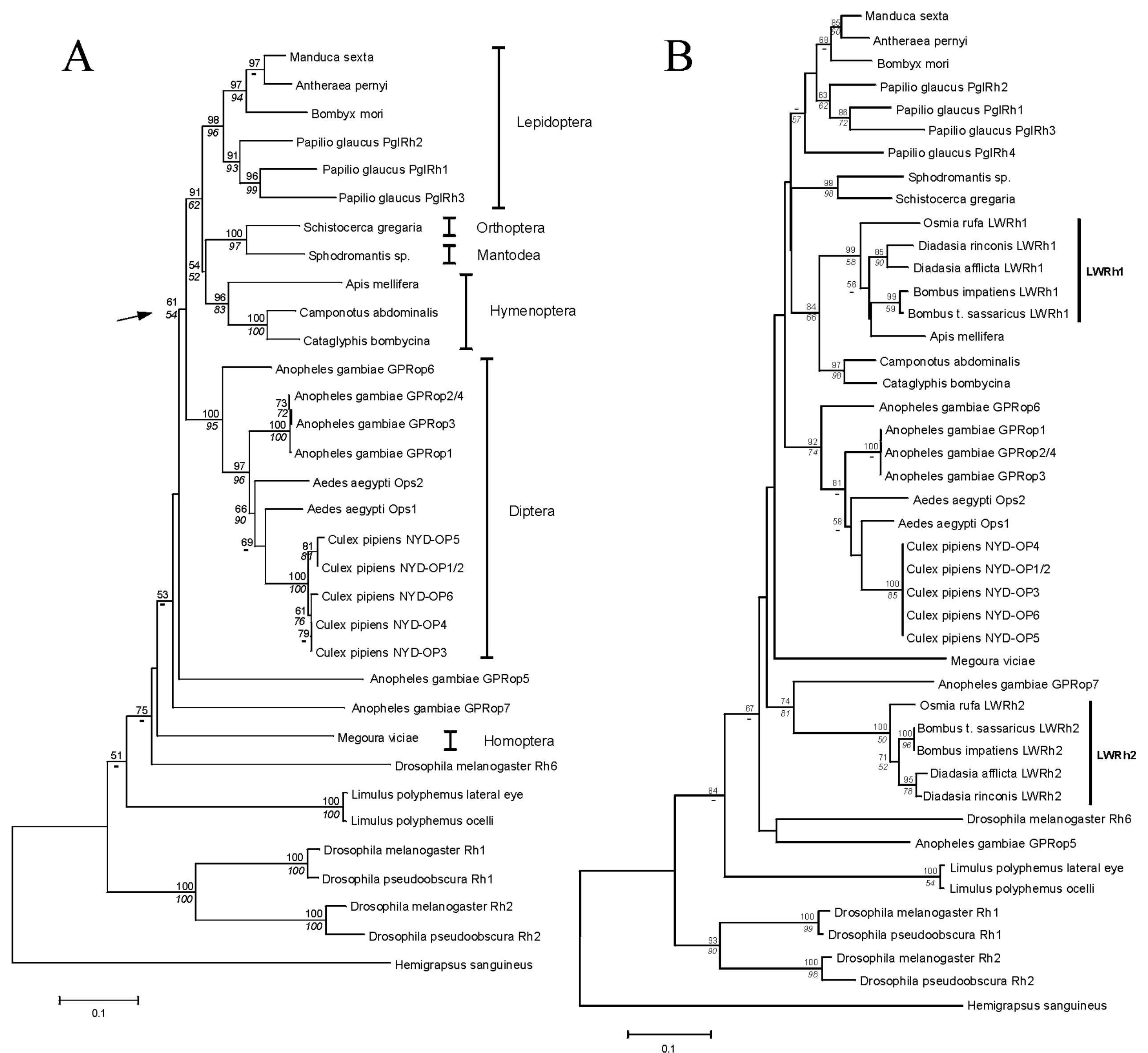

FIG. 1.—Phylogeny of arthropod blue-green and long wavelength–sensitive opsins. (*A*) The neighbor-joining (NJ) tree (shown) was constructed using an amino acid alignment and Poisson correction. Bootstrap support values (out of 500 replicates) for the NJ tree are shown above the nodes and bootstrap support (out of 100 replicates) for the step matrix–weighted MP tree are below. (*B*) Phylogeny of insect LW and blue-green opsins based upon 206 amino acid sites. Values above the nodes represent support out of 500 NJ bootstrap replicates with Poisson correction; below the nodes, out of 1,000 quartet-puzzling steps. Only values of 50% or higher are shown. GenBank accession numbers for the sequences used in the reconstruction are as follows: *Anopheles gambiae* (Scaffold Nos., GPRop1 AAAB01008987 gi|19612317:15485449–15486561, GPRop2 AAAB01008987 gi|19612317:15548577–15547465, GPRop3 AAAB01008987 gi|19612317:15549903–15551016, GPRop4 AAAB01008987 gi|19612317:155485608–15547465, GPRop5 AAAB01008987 gi|19612317:15552323–15553525, GPRop6, AAAB01008987 gi|19612317:15555428–15556641; GPRop7, AAAB01008859 gi|19611897:9719568–9718344); *Antheraea pernyi* (AB073299); *Apis mellifera* (U26026); *Bombyx mori* (AB064496); *Camponotus abdominalis* (U32502); *Cataglyphis bombycinus* (U32501); *Culex pipiens* (NYD-OP1/2, AY297441; NYD-OP3 AY297443; NYD-OP4, AY299444; NYD-OP5, AY299337; NYD-OP6, AY299338); *Drosophila melanogaster* (Rh1, K02315; Rh2, M12896; Rh6, Z86118); *Drosophila pseudoobscura* (Rh1, X65877; Rh2, X65878); *Hemigrapsus sanguineus* (BcRh1, D50583); *Limulus polyphemus* (lateral eye, L03781; ocelli, L03782); *Manduca sexta* (Manop1, L78080); *Megoura viciae* (AF189714); *Papilio glaucus* (PglRh1, AF077189; PglRh2, AF077190; PglRh3, AF067080; PglRh4, AF077193); *Schistocerca gregaria* (X80071); *Sphodromantis sp.* (X71665). *Aedes aegypti* Ops1 and Ops2 sequences are published in Graf et al. (1996).

and 4°C. Afterward the cellular debris was removed, and DNA was extracted by use of phenol/chloroform.

PCR, Cloning, and Sequencing

We first amplified two to three fragments from each species by polymerase chain reaction (PCR) with a *Taq* DNA polymerase (Hot Master, Eppendorf) and different combinations of four degenerate primers (80, OPSRD, LWRHF, and LWRHR) (Chang et al. 1996; Mardulyn and Cameron 1999). The fragments spanned the last two to five transmembrane domains. To obtain remaining transmembrane domains and internal sequencing primers, we used the initial fragments to design gene-specific as well as

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019

Table 1
Bee Specimens Used in This Study, GenBank Accession Numbers, and Sample Origin

Species	Gene	Accession Number	Source/Location
Bombus impatiens	*LW Rh1*	AY485302	Koppert, USA
	LW Rh2	AY485306	
Bombus terrestris sassaricus	*LW Rh1*	AY485301	Sardinia, Italy
	LW Rh2	AY485305	
Diadasia afflicta	*LW Rh1*	AY485303	Pedernales Falls State Park, Blanco County, Texas, USA
	LW Rh2	AY485308	
Diadasia rinconis	*LW Rh1*	AY485304	Pedernales Falls State Park, Blanco County, Texas, USA
	LW Rh2	AY485307	
Osmia rufa	*LW Rh1*	AY572828	Wuerzburg, Germany
	LW Rh2	AY572829	

degenerate reverse primers (see table 2 for primer sequences). These primers were paired either with LWRHF or newly designed degenerate forward primers, BLWFD or B2LWFD, based on the already published *Apis mellifera*, *Camponotus abdominalis*, and *Cataglyphis bicolor* LW opsin cDNA sequences (see legend of figure 1 for GenBank accession numbers).

For most primer combinations we used the following PCR thermocycling profile: denaturation for 2 min at 94°C, followed by 35 cycles of 94°C for 30 s, 50°C to 54°C for 30 s and 68°C for 1 min, then 10 min at 68°C. The PCR products were either run on a gel and subsequently cut out and purified by use of a Genclean kit (Q-BIOgene) or directly purified by use of a QIAquick PCR purification kit (Qiagen, Inc.). Most PCR products, except the *D. rinconis* and *B. impatiens LW Rh2* (LWRHF-LWRHR2) gene product, were cloned into a pGEM-T Easy plasmid (Promega). The plasmids were prepared in a QIAprep Spin Miniprep Kit (Qiagen) or a FastPlasmid Miniprep kit (Eppendorf), screened for inserts by *EcoRI* digestion, and then cycle sequenced (BigDye Terminator version 3.1 Cycle Sequencing Kit) on an ABI Prism 3100 DNA automated sequencer (Applied Biosystems). Sequences were imported into SeqMan (DNASTAR, Inc., Madison, Wis.), assembled into contigs, and manually edited. Sequences were then manually aligned in MacClade (Maddison and Maddison 2000). The first sequence that exhibited a high amino acid sequence similarity to other published bee LW-sensitive opsin genes was termed *LW Rh1*, and the second gene was termed *LW Rh2*.

On the basis of the sequences for *B. terrestris* and *D. afflicta*, which we cloned and sequenced first, we designed two gene-specific degenerate reverse primer (LWRH1 and LWRHR2) for the novel gene *LW Rh2* to be able to target the gene by application of a direct sequencing strategy. The reverse primer was then paired with LWRHF. A single PCR product was obtained for both *B. impatiens* and *D. rinconis* (943 bp and 967 bp, respectively) with LWRH2 as the reverse, and both products were directly sequenced following the method described above. We found that the newly designed reverse primer LWRHR2, in combination with LWRHF, exclusively amplifies the new gene in *Bombus* and *Diadasia*, whereas amplification of the *Osmia rufa LW Rh2* gene worked with LWRHF and LWRHR1.

Phylogenetic Analyses of the New Bee Sequences in Relation to Other Bees

LW opsin sequences from a large number of hymenopteran species have been published (Ascher, Danforth, and Ji 2001; Cameron and Mardulyn 2001; Cameron and Williams 2003; Cook et al. 2002; Danforth, Conway, and Ji 2003; Kawakita et al. 2003; Mardulyn and Cameron 1999). We were interested in determining whether any of these sequences are orthologous to *LW Rh2*, and, thus, whether the identification of this new gene has an impact on existing phylogenies. We, therefore, calculated synonymous substitution rates among 234 hymenopteran sequences downloaded from GenBank and the *LW Rh1* and *LW Rh2* sequence from one of our bees, *B. terrestris*. All pairwise comparisons between the *B. terrestris LW Rh2* sequence and any of the published sequences resulted in a value larger than 0.75 and indicated saturation (data are not shown). However, no saturation of the rate of synonymous substitution was found within the data set when *LW Rh2* was excluded. We,

Table 2
Nucleotide Sequence and Gene Specificity of Newly Designed Primers Used in This Study

Name	Nucleotide Sequence	Gene Specificity	Species Specificity
BLWFD	5′-CAY HTN RTN GAY GCN MAY TGG-3′	*LW Rh1*	B, D, O
B2LWFD	5′-CAY YTN ATH GAY GCN AAY TGG-3′	*LW Rh1*	B, D, O
BLWR1	5′-TGC TGT CAT ATT ACC CTC TGG A-3′	*LW Rh2*	B, D
BLWR2	5′-AGA GCT CCG TTA ATG GTC AAT GG-3′	*LW Rh1*	B
DLWR1	5′-CCC TCG GGA ACG TAT CTA AAA G-3′	*LW Rh1*	D
DLWR2	5′-CGT CGT TAG TGC TAC CTG AAG TTT-3′	*LW Rh2*	D
LWRHR2	5′-RGA WCG ATT AAA TAT TCC AAT-3′	*LW Rh2*	B, D
LWRHR1	5′-ACC CCA DAT KGT RAA RAG BGG BGT-3′	*LW Rh2*	B, D, O
OLWR1	5′-TCG TTG AAT TGT TTG TTA GAA GGA-3′	*LW Rh1*	O
OLWR2	5′-GTC CAG GCC ATA AAC CAC AAT-3′	*LW Rh2*	O

Note.—For primer position see figure 2. B = *Bombus;* D = *Diadasia;* O = *Osmia.*

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019

therefore, conducted a phylogenetic analysis with the *LW Rh1* sequences from our five bee species and 231 published hymenopteran sequences to determine the relationship of the new *LW Rh1* gene sequences and the already published LW-sensitive hymenopteran opsins. We used the NJ algorithm with p-distance and included all sites (for species names and GenBank accession numbers see Supplementary Material online at http://visiongene.bio.uci.edu/ABresearch.html). We also used the maximum likelihood algorithm with the general time reversible model + gamma as implemented in PhyML (Guindon and Gascuel 2003). The length of the sequences in the analysis was 471 nucleotides, and we excluded introns as well as four very short sequences (*Ceratina* sp., *Exomalopsis completa*, *E. rufiventris*, and *Tetraloniella* sp.).

Relative Rates Tests of the Duplicated Bee Gene Sequences

Possible site-specific rate changes that accompany gene duplication may indicate altered selective constraints (either enhanced or reduced) after diversification and provide information about potential amino acid residues that may count for functional divergence between the two genes (such changes are called type-I evolutionary functional divergence [Gu 1999]). (For a general review of functional divergence tests in protein evolution see Gaucher et al. [2002] and Yang and Bielawski [2000].) Hence, we attempted to detect possible site-specific rate changes between LW Rh1 and LW Rh2 by use of a nonhomogeneous gamma model that is implemented in the software DIVERGE version 1.04 (Gu 1999; Gu and Vander Velden 2002). For that, we calculated the coefficient of functional divergence (θ) between the two gene clusters for each site and tested the null hypothesis $\theta = 0$. Conceptually, θ is a measure of the degree of independence between the relative evolutionary rates at similar sites in the LW Rh1 and LW Rh2 clusters. The coefficient ranges from 0 to 1, where 0 indicates that the evolutionary rate is virtually the same in both gene clusters. Values of θ significantly greater than 0 indicate rate shifts between homologous sites among the two gene clusters (Gu 1999, 2001).

Homology Modeling of Hymenopteran Opsins

Identifying polymorphic amino acid sites, correlating amino acid substitutions with shifts in spectral sensitivity, and mapping such substitutions onto a three-dimensional crystal structure of rhodopsin has been shown to be a useful approach for studying the relationship between opsin structure and spectral function (Briscoe 2002; Hunt et al. 2001). In such a way, the identified amino acid sites allow predictions about potential spectral-tuning effects in homologous opsin genes even if physiological data for these opsins are still missing. We, therefore, created a homology model of the complete *Apis mellifera* LW Rh1 opsin by the same procedure as described in Briscoe (2002) and mapped onto it variable sites that have previously been implicated in spectral tuning as well as the sites identified by method of Gu (1999, 2001).

Results

Duplicated *LW Rh* Genes Isolated from *Bombus*, *Diadasia*, and *Osmia*

The results of both NJ and MP analyses indicated at least one LW gene duplication event before the diversification of mantodean, orthopteran, dipteran, hymenopteran, and lepidopteran orders (fig. 1*A*). Therefore, we decided to screen the hymenopteran genome for evidence of this early gene duplication event by application of a degenerate PCR strategy. For each of the five bee species we screened, we obtained two sequences that varied in length from 1,082 to 1,283 bp for *LW Rh1* and from 995 to 1,027 bp for *LW Rh2*. (The *LW Rh1* fragments from *Bombus terrestris* and *B. impatiens* were about 500 bp longer than previously published sequences from these species). We identified three introns in the *LW Rh1* gene and five introns in the *LW Rh2* gene, which varied in length between species (fig. 2). Two of the introns were found at conserved positions in both genes. A comparison with other insect green and blue-green opsin genes revealed that two out of three introns in *LW Rh1* and four out of five in *LW Rh2* share similar splice sites with other dipteran and lepidopteran species (Briscoe 1999; Hsu et al. 2001). We also observed that most intron splice sites started with the nucleotides GT, except the intron after amino acid position 37 in *Diadasia LW Rh2*, which started with GC, and the intron after amino acid position 121 in the same gene, which started with GA in *B. impatiens* and with GC in both *Diadasia* species.

The translation of the nucleotide sequences indicates that we obtained all seven transmembrane domains for LW Rh1 and five of seven transmembrane domains for LW Rh2. (LW Rh2 is presumably a functional gene, as we did not identify any stop codons). The deduced amino acid sequence was 284 amino acid residues for LW Rh1 and 206 amino acid residues for LW Rh2 (fig. 3). The amino acid similarity was 84.3% among the LW Rh1 and 92.2% among the LW Rh2 sequences. However, only 67.5% of aligned amino acid sites were found conserved between both genes, indicating that both genes are only distantly related.

Green and Blue-Green Opsin Family Phylogeny

As mentioned previously, the opsin gene family is subdivided into orthologous groups that generally correspond to functional differences in visual pigment spectral properties. We, therefore, established a phylogeny using NJ and quartet-puzzling analysis of amino acid sequences, and ML analysis of nucleotide 2nd positions (GTR + Γ model) as implemented in PhyML to verify that the novel *LW Rh2* gene is indeed part of the LW-sensitive opsin class. The NJ analysis revealed that both bee genes, *LW Rh1* and *LW Rh2*, form, along with other insect LW opsin sequences, a group that is clearly distinguished from the blue-green–sensitive (λmax = 420 to 480 nm) *Drosophila* and crustacean opsins (fig. 1*B*). In general, the topology of the NJ and quartet-puzzling trees that include the shorter bee LW Rh1 and LW Rh2 sequences is similar to that observed in the original analysis (fig. 1*A*) and also in an ML analysis of 2nd positions of the full-length blue-green

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019

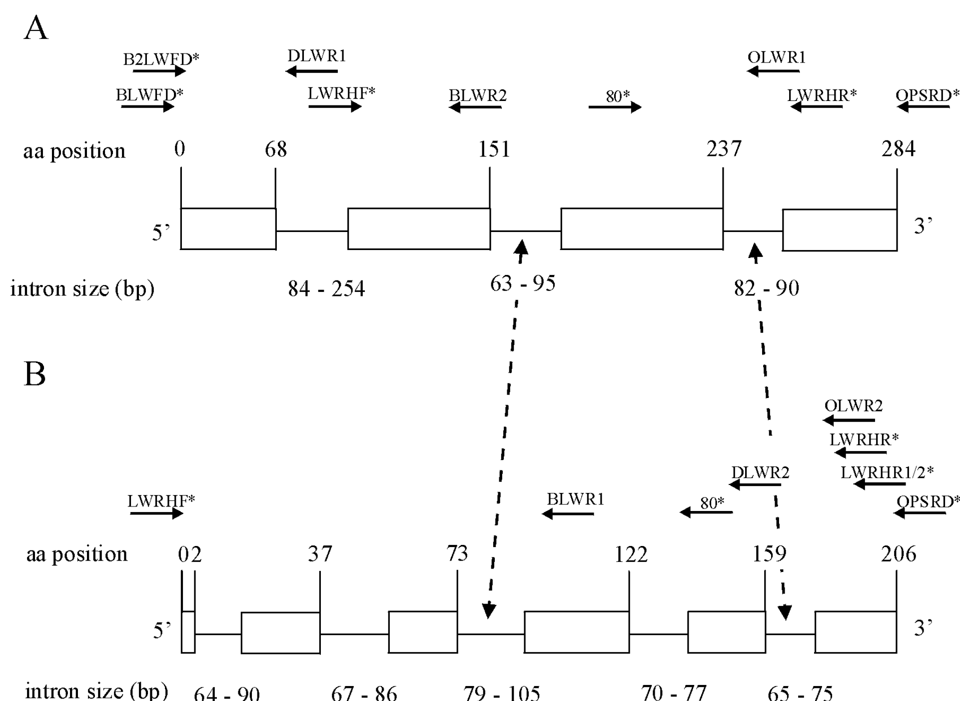

FIG. 2.—Primers used for PCR amplification and direct sequencing of genomic DNA mapped onto the (*A*) *LW Rh1* and (*B*) *LW Rh2* sequences. Degenerate primers are marked with an asterisk. Dashed lines indicate conserved introns between the two genes.

and LW-sensitive opsins (data not shown). The ML and quartet-puzzling trees of the shorter alignment are very similar in topology to the NJ tree shown in figure 1*B*, except that *Drosophila Rh1* and *Rh2* form a part of the group that includes *LW Rh2* and *Anopheles GPRop 7*, and *Drosophila Rh6* is the outgroup to all three. In all analyses, the LW Rh1 sequences cluster together with other known hymenopteran opsins (100% NJ bootstrap support, 58% quartet-puzzling support). In the NJ and quartet-puzzling trees, the LW Rh2 sequences are most closely related to the *Anopheles gambiae* GPRop7 opsin (72% NJ bootstrap support, 81% quartet-puzzling support) and are, thus, only very distantly related to all other hymenopteran LW-sensitive opsins. We note that the amino acid topology shown in figure 1*B* also recovers Rh6 as a homolog of *Anopheles* GPRop5, and that these sequences share an indel that is not found in any of the other mosquito opsins.

Hymenopteran LW-Sensitive Opsin Phylogeny

To control for the possibility of PCR contamination and to test what impact the new gene has on existing phylogenies, the LW opsin nucleotide sequences from 231 species and 13 families were analyzed together with the *LW Rh1* sequences from our five species. All new *LW Rh1* sequences presented in this study cluster together with the other members of their genus (phylogeny available as Supplementary Material online at http://visiongene.bio.uci.edu/ABresearch.html). *Diadasia afflicta* forms a well-supported branch (86% NJ bootstrap support) with *D. nigrifrons* that is also recovered in the ML analysis. *D. rinconis* is the sister group to *D. martialis*–*D. diminuta* in both the ML tree and the NJ tree. *Osmia rufa* clusters with other Megachilinae, *Hoplitis*, and *Chelostoma* spp. (61% bootstrap support). The two *Bombus* sequences cluster together with the previously published sequences from these species in both the NJ and ML trees and with all other published LW-sensitive opsin gene sequences from about 80 species of the same genus. In both species, our sequence data and the already published sequences showed variation at only six of the 711 nucleotide sites of common overlap. This analysis demonstrates that the *LW Rh1* sequences we obtained from our species are most closely related to all previously reported *LW Rh* sequences in GenBank from similar species or genera. The identification of the new *LW Rh2* gene in hymenopterans, therefore, does not impact the interpretation of previously published *LW Rh1* phylogenies.

Rate Shifts Between *LW Rh1* and *LW Rh2*

Pairwise comparisons of all 10 bee sequences by application of the Nei-Gojobori algorithm (Nei and Gojobori 1986) indicated saturation of synonymous substitutions (d_S) between the *LW Rh1* and *LW Rh2* genes (fig. 4). (We also used the Li, Wu, and Luo [1985] method, which takes into account transition and transversion bias, and synonymous substitutions were incalculable between the paralogs). Mean d_S was 0.44 and 0.47 within *LW Rh1* and *LW Rh2*, respectively; d_S between *LW Rh1* and *LW Rh2* was 0.81. The nonsynonymous substitutions per nonsynonymous site were 0.05 for *LW Rh1* and 0.02 for *LW Rh2* (we found no differences at the amino acid level between the *B. terrestris* and *B. impatiens LW Rh2* sequences). Mean d_N between *LW Rh1* and *LW Rh2* was 0.18. We used the Z-statistics implemented in MEGA 3.0 (Kumar, Tamura, and Nei 2003) to test the hypothesis that $d_S = d_N$; that is, both genes evolved neutrally and no

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019

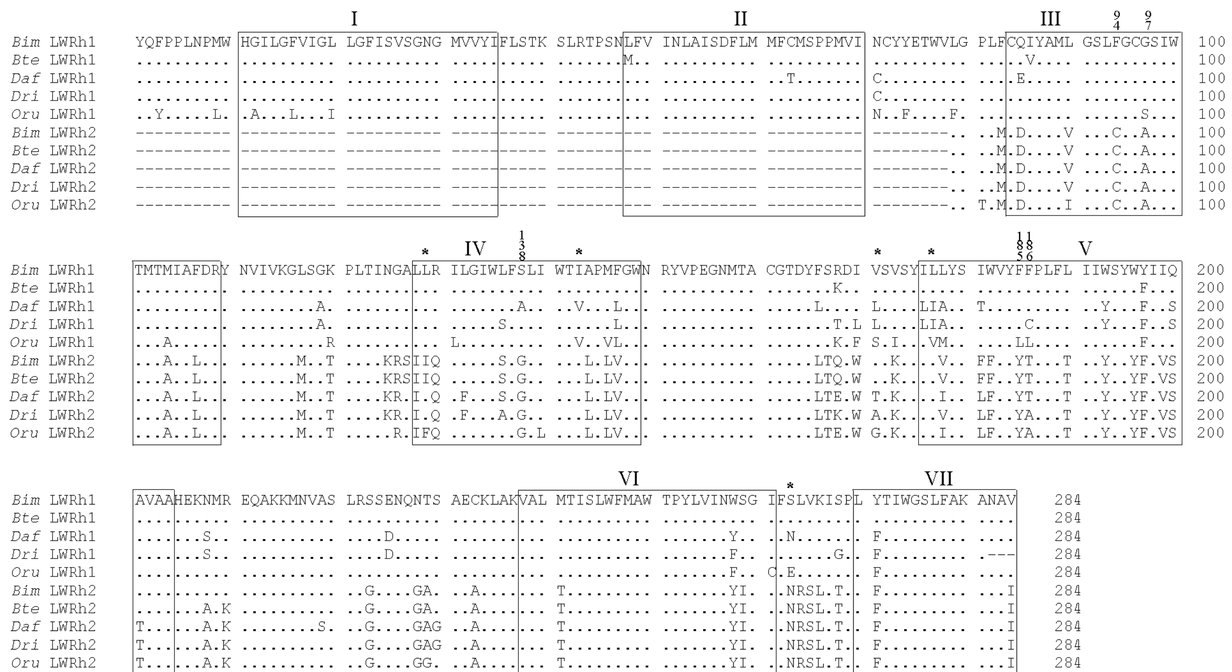

FIG. 3.—Alignment of the deduced amino acids of the two genes from five bees, *Bombus terrestris*, *B. impatiens*, *Diadasia afflicta*, *Osmia rufa*, and *D. rinconis*. Shaded blocks indicate the location of the transmembrane domains (I–VII) within the opsin gene. Asterisks indicate sites that evolved significantly faster in one of the two genes after gene duplication.

selection is operating on them. For both genes, we could reject the hypothesis and we found that d_N is significantly smaller than d_S (*LW Rh1*: Z = −13.0, $P < 0.001$; *LW Rh2*: Z = −16.8, $P < 0.001$), which indicates a strong purifying selection pressure on both genes. We also tested whether synonymous and nonsynonymous substitution rates differ between the two genes ($d_{S(LW\ Rh1)} = d_{S(LW\ Rh2)}$ and $d_{N(LW\ Rh1)} = d_{N(LW\ Rh2)}$; same Z-statistics as above). Nonsynonymous substitution rate was found to be significantly lower in *LW Rh2* compared with *LW Rh1* (Z = 2.8, $P < 0.01$), and we could not reject H_0 that $d_{S(LW\ Rh1)} = d_{S(LW\ Rh2)}$ (Z = −0.84, $P > 0.05$). Our results indicate that the rate of synonymous nucleotide substitution does not differ between both genes ($d_{S(LW\ Rh1)} = d_{S(LW\ Rh2)}$), but the effect of purifying selection pressure seems to be stronger on *LW Rh2* than on *LW Rh1* ($d_{N(LW\ Rh1)} > d_{N(LW\ Rh2)}$). The higher rate of amino acid substitutions among the LW Rh1 sequences compared with LW Rh2 after speciation is apparent in the long branches at the tips of the LW Rh1 cluster compared with the equivalent LW Rh2 branches (fig. 1*B*). These results suggest that *LW Rh2* may be a more useful gene for resolving higher-level taxonomic relationships than is *LW Rh1*.

We were also interested in using a model that takes into account lineage-specific rate differences at different sites (Gu 1999; Gu and Vander Velden 2002). We, therefore, calculated that the θ value (coefficient of functional divergence) between the LW Rh1 cluster and the LW Rh2 cluster is 0.53 (SE = 0.21) and that of the gamma shape parameter is 0.20. The estimated θ value is, therefore, significantly different from 0 (likelihood ratio test, LRT: 6.31, $P < 0.03$), and, thus, rate shifts could be detected at specific sites. Five specific sites with posterior probabilities $P^*(S_1 \mid X) > 0.70$ that are likely to be involved in type I functional divergence between the two genes were identified. One site (129) was evolving faster in LW Rh2, whereas four sites (143, 170, 177, and 263) were evolving more rapidly in LW Rh1. Our analyses have, therefore, shown that the diversification within LW Rh1 and LW Rh2 after gene duplication is accompanied by a functional divergence of type I (Gu 1999, 2001) at these sites.

Homology Modeling of *Apis mellifera LW Rh1*

Specific amino acids in the chromophore-binding pocket are known to modulate the absorption spectrum maximum of the visual pigment to longer or shorter wavelengths of light (Salcedo et al. 2003). Therefore, we used the homology model of the *Apis mellifera* LW Rh1 opsin to map variable residues near the chromophore and identified five sites in our data set (residues 94, 97, 138, 185, and 186) that are homologous to sites that are involved in spectral shifts in human, New World monkey, crayfish, or insect pigments. We also mapped the five sites identified above as undergoing type I functional divergence and found that all of these sites face the exterior of the protein, away from the chromophore-binding pocket (data not shown). The significance of these sites for functional differentiation of the two opsins is discussed below.

Discussion

Among the many practical applications of phylogenetic reconstruction to the analysis of genomic data is its clear utility in predicting the discovery of "lost" orthologous genes in genomes yet to be sequenced. In this

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019

Nonsynonymous substitution per nonsynonymous site (d_n)

		1	2	3	4	5	6	7	8	9	10
1	*B. impatiens LWRh1*		0.006	0.051	0.049	0.065	0.182	0.182	0.191	0.190	0.191
2	*B. terrestris LWRh1*	0.050		0.054	0.049	0.062	0.183	0.183	0.190	0.190	0.191
3	*D. afflicta LWRh1*	0.547	0.521		0.026	0.082	0.172	0.172	0.185	0.182	0.184
4	*D. rinconis LWRh1*	0.533	0.507	0.083		0.081	0.174	0.174	0.186	0.183	0.191
5	*O. rufa LWRh1*	0.550	0.517	0.508	0.540		0.193	0.194	0.195	0.194	0.197
6	*B. impatiens LWRh2*	0.800	0.821	0.796	0.853	0.756		0.000	0.023	0.023	0.035
7	*D. terrestris LWRh2*	0.798	0.819	0.794	0.851	0.772	0.054		0.023	0.023	0.035
8	*D. rinconis LWRh2*	0.846	0.835	0.782	0.825	0.821	0.590	0.582		0.011	0.033
9	*D. afflicta LWRh2*	0.833	0.843	0.758	0.808	0.840	0.590	0.582	0.094		0.039
10	*O. rufa LWRh2*	0.852	0.862	0.844	0.844	0.760	0.476	0.503	0.579	0.611	

Synonymous substitution per synonymous site (d_s)

FIG. 4.—P-distances of synonymous and nonsynonymous substitutions among the 10 sequences using the Nei-Gojobori algorithm with complete deletion of gaps or missing data. Shaded block indicates the saturated synonymous substitution rates between the two genes.

study, we have used phylogenetic analysis of arthropod opsin sequences to identify an early gene duplication event within the insect LW opsin family (fig. 1*A*). The minimum time estimate for the occurrence of one of these duplication events is before the radiation of lepidopteran, dipteran, hymenopteran, orthopteran, and mantodean orders (>300 MYA) (Wiegmann et al. 2003). This branching pattern suggests the existence of at least one LW opsin gene beyond what has been reported so far in these groups. We tested this hypothesis by application of a PCR-based screen and degenerate primers to isolate a novel LW opsin gene (*LW Rh2*) from five species of bees that is paralogous and only distantly related to the *LW Rh1* sequences we isolated from those same species and all previously published *LW Rh* sequences in hymenopterans (figs. 2 and 3). We found that all lepidopteran, hymenopteran (LW Rh1), and orthopteran as well as mantodean opsin sequences form a sister group to the mosquito opsins, *Aedes* Ops1 and Ops2, *Anopheles* GPRop1, GPRop2/4, GPRop3 and GPRop6, and *Culex* OPS1-6 (fig. 1*A*). At the same time, *Anopheles* GPRop7 is most closely related to LW Rh2 (fig. 1*B*). These trees indicate that the new bee LW Rh2 opsin, together with *Anopheles* GPRop7, forms a very ancient subclass within the green-sensitive opsin genes that arose from a gene duplication event that took place before all the included insect orders diverged.

Given the large number of papers in recent years that have utilized the *LW Rh1* gene as a molecular marker (Hsu et al. 2001; Jiggins 2003; Ortiz-Rivas, Moya, and Martinez-Torres 2004), particularly for hymenopteran systematics (Mardulyn and Cameron 1999; Ascher, Danforth, and Ji 2001; Cook et al. 2002; Kawakita et al. 2003), the impact of a second hymenopteran LW opsin gene on all such previously published studies is worth briefly considering. First, we have shown that all available hymenopteran LW opsin sequences in GenBank belong to the LW Rh1 group. Therefore, our results do not challenge current opsin phylogenies. Second, we have found that *LW Rh2* is evolving at a much slower rate than LW Rh1 (fig. 1*B*). This finding suggests that this new gene might usefully resolve higher-level phylogenetic relationships among hymenopterans. Third, we have developed gene-specific primers that permit the direct sequencing of this gene from genomic DNA (table 2). The availability of these primers should make direct testing of this hypothesis possible.

The impact of *LW Rh2* on our original rationale for examining the hymenopteran genome for extra LW opsin copies is also worth considering. Does the discovery of *LW Rh2* shed light on the origins of the extra LW photoreceptors found in the retina of some hymenopteran species by Peitsch et al. (1992) in so far as this new gene may have given rise to those receptors? We think that the answer is likely to be no because *LW Rh1* is expressed in the retina, whereas *LW Rh2* is likely extraretinal in origin. Therefore, it is more parsimonious to assume that the retinal receptors described by Peitsch et al. (1992) evolved from other retinal opsins, such as *LW Rh1*, rather than from a (more distantly related) extraretinal opsin. Duplicated LW opsins that have given rise to new retinal photoreceptor classes have been found in butterflies (e.g., *PglRh1* and *PglRh3*) (Kitamoto et al. 1998), and as shown in figure 1, these genes are closely related to each other. Furthermore, these gene duplications evolved after the split between the common ancestor of nymphalid and papilionid butterflies (Briscoe 2001). Therefore, in the absence of screening pigments, the extra receptors found in the Andrenidae, Xiphydriidae, and Tenthredinidae are very likely the result of recent, independent duplications of the *LW Rh1* gene (see discussion below).

Where then might the new *LW Rh2* gene be expressed? From the original study reporting the cloning of the *Apis mellifera LW Rh1* gene from eye-specific cDNA (Chang et al. 1996), it is clear that this gene is expressed in the photoreceptor cells in the retina. We have attempted RT-PCR using the gene-specific *LW Rh2* primers on cDNA synthesized from *Bombus impatiens* head tissue mRNA. It is also clear from these experiments that the *LW Rh2* transcript is not very abundant, because we were unable to amplify this gene from our cDNA. By contrast, we easily amplified, cloned, and sequenced the *LW Rh1* transcript and localized it by in situ hybridization to the retina (Spaethe and Briscoe, unpublished data). We speculate that *LW Rh2* may have a specialized function as an extraretinal opsin, expressed in a small number of light-sensitive neurons in the brains of bees. What evidence do we have to support this hypothesis? First, an extraretinal opsin, *PglRh4*, has been cloned from

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019

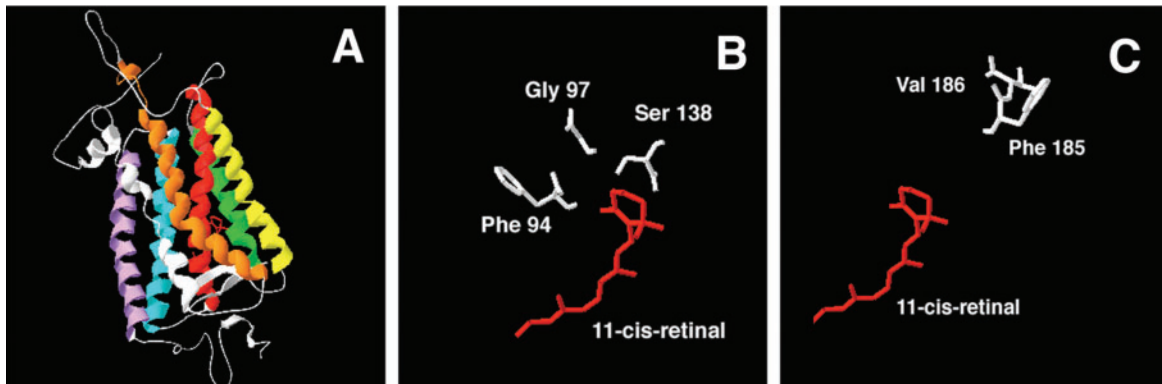

FIG. 5.—Homology model of *Apis mellifera* LW Rh1 opsin and candidate tuning sites. (*A*). Ribbon diagram of three-dimensional structure of *Apis mellifera* long wavelength–sensitive opsin; bovine rhodopsin (Palczewski et al. 2000) was used as a template. Seven transmembrane domains are indicated by colored helices. Loop domains are shown in gray. Magnified view of the same homology model showing 11-*cis*-retinal chromophore and putative candidate amino acid spectral-tuning sites (*B*) 94, 97, and 138, and (*C*) 185 and 186, which differ between the duplicated hymenopteran genes. Residues shown are those found in the *Apis LW Rh1* sequence.

butterflies (Briscoe and Nagy 1999). This transcript is expressed at exceedingly low levels, such that it was undetectable on a Northern blot of mRNA from a single individual butterfly head (Briscoe 2000). (Low mRNA abundance seems to be the case also for *Anopheles GPRop7*, [C. A. Hill, personal communication]). Second, polyclonal antibodies specific to this protein have been developed and show that this opsin is localized to the optic lobes of butterflies and moths (Briscoe and Nagy 1999; Lampel, Briscoe and Wasserthal, unpublished data.). This opsin appears to have evolved before the radiation of lepidopterans (moths and butterflies) (fig. 1*B*). Therefore, it is possible that extraretinal opsins evolved early in invertebrate evolution, but gene conversion and gene loss in some lineages has obscured this development. (Gene conversion, for example, has played a role in the evolution of some dipteran LW opsins [Spaethe and Briscoe, unpublished observation]). This finding makes sense from a functional point of view because generalized light detection is a necessary component for the entrainment of photoperiodic and circadian rhythms. We note that the *Drosophila melanogaster* Rh6 opsin has been detected in the extraretinal eyelet structure (Yasuyama and Meinertzhagen 1999) and that this class of photoreceptor has been proposed to be part of the light sensitive input pathway for the photic entrainment of the circadian clock (Malpel, Klarsfeld, and Rouyer 2002; Shimizu, Yamakawa, and Iwasa 2001). We speculate that *LW Rh2* together with *Anopheles GPRop7* represent a novel and ancient class of extraretinal opsins.

Pattern of Amino Acid Substitution and Candidate Spectral-Tuning Sites

Finally, we consider whether the observed amino acid variation between the two hymenopteran proteins is likely to result in spectral tuning or other functional differences. By use of the Gu (1999, 2001) method, we identified five sites as undergoing type I functional divergence (amino acids 129, 143, 170, 177, 263). Each of these sites, when mapped onto the homology model of the *Apis mellifera* LW Rh1 opsin, faces the exterior of the protein, and the significance of these sites for rhodopsin function is unknown. Eighteen fixed (or nearly fixed) differences between LW Rh1 and LW Rh2 exist. A number of these sites (e.g., 117, 120, 228, and 229) are located in cytoplasmic loops II and III and are presumably involved in the recognition of the G-protein (Teller et al. 2001). This finding suggests that functional differences in G-protein binding or activation may exist between the two opsins. Five variable sites (residues 94, 97, 138, 185, and 186) were found to be homologous to sites that are involved in spectral shifts in human, New World monkey, crayfish, or insect pigments (fig. 3).

Amino acid sites 94 and 97 have undergone parallel/convergent changes that are correlated with spectral shifts in peak sensitivity in several butterfly opsins (Briscoe 2001) and have also undergone similar changes in crayfish (Crandall and Cronin 1997) and bees (Briscoe 2002). We observed a Phe to Cys substitution at amino acid 94 and a Gly to Ala substitution at amino acid 97 between LW Rh1 and LW Rh2 (fig. 3), which suggests that these sites are undergoing correlated evolution. Mapping residue 94 onto a homology model of the full-length *Apis mellifera* LW opsin showed that this amino acid is located in the third transmembrane domain of the protein and faces the chromophore-binding pocket, which makes it highly likely to have an effect on spectral tuning in the bee opsins (fig. 5*B*). (Indeed, site-directed mutagenesis of this site was shown to affect wavelength regulation [Zhukovsky, Robinson, and Oprian 1992]).

Residue 138 in transmembrane domain IV is variable between LW Rh1 and LW Rh2 as well as within LW Rh1 (Ser/Ala to Gly) (figs. 3 and 5 *B*). Site-directed mutagenesis experiments (Asenjo, Rim, and Oprian 1994) have shown that a Ser to Ala substitution in human cone pigments causes a 5-nm blue shift caused by the loss of a side-chain hydroxyl group. Thus, the Ser to Ala substitution detected in *D. afflicta* strongly indicates the possibility of a blue shift of the LW Rh1 opsin in this species compared with the other six *Diadasia* species from which sequence data are available.

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019

The Phe to Tyr substitution at amino acid site 185 in transmembrane domain V is also found in butterflies and crayfish and is correlated with a red shift (Briscoe 2001; Crandall and Cronin 1997) (fig. 5*C*). In New World monkeys, an Ile to Phe substitution at residue 229 (which is equivalent to amino acid position 185 in our sequences) appears to be responsible for a 2-nm blue shift (Shyue et al. 1998). In addition, residue 186 (corresponding to human red cone pigment amino acid 230), which is highly variable in our four bee species and might be coevolving together with the adjacent residue 185, has been shown to cause a 1-nm blue shift in human cone pigments (Asenjo, Rim, and Oprian 1994; Merbs and Nathans 1993). Altogether we identified three residues (F94C, S138A, and F185Y) at which identical amino acid substitutions have been reported in other vertebrate and insect species and were shown to affect spectral tuning. Convergent evolution of vertebrate and invertebrate opsin spectral-tuning mechanisms has recently been experimentally verified by Salcedo et al. (2003), at least for UV vision. These results suggest that spectral diversification both within and between LW Rh1 and LW Rh2 may have played a role in the evolution of these opsins.

Conclusion

Studies that provide sequence data from entire genomes have revealed a much higher number of opsin genes than behavioral or physiological studies suggest and indicate that the LW-sensitive opsin class is much more functionally diverse than previously assumed. We have used phylogenetic analysis as a tool for predicting the existence of one "lost" gene in hymenopterans and other insects, and we have used a PCR-based screen to confirm its existence in five species of bees. Relative rate tests indicate that the new gene is evolving at a slower rate than the known hymenopteran LW gene and, therefore, may be useful for the reconstruction of higher-level hymenopteran phylogenies. Spatial expression studies have shown that some of the LW opsins are partially or even exclusively expressed outside of the retina and are proposed to be part of the light sensitive input pathway for the photic entrainment of the circadian clock. We speculate that the hymenopteran LW Rh2 opsin together with *Anopheles* GPRop7 represent a novel and ancient class of extraretinal opsins. Functional changes in both G-protein–binding sites and spectral-tuning sites are likely to have occurred after duplication of the hymenopteran LW opsins.

Acknowledgments

We thank Malou Sison-Mangus and Balaji Ramanujam for technical assistance with direct sequencing and implementation of PhyML, Sudhir Kumar for permission to use MEGA 3.0, and two anonymous reviewers for helpful comments on our manuscript. This research was funded in part by research grants from UCI and NSF (IBN-0346765) to A.D.B.; J.S. was supported by a fellowship from the German Academic Exchange Service (DAAD).

Literature Cited

Ascher, J. S., B. N. Danforth, and S. Ji. 2001. Phylogenetic utility of the major opsin in bees (Hymenoptera: Apoidea): a reassessment. Mol. Phylogenet. Evol. **19**:76–93.

Asenjo, A. B., J. Rim, and D. D. Oprian. 1994. Molecular determinants of human red/green color discrimination. Neuron **12**:1131–1138.

Bellingham, J., S. Wilkie, A. Morris, J. Bowmaker, and D. Hunt. 1997. Characterization of the ultraviolet-sensitive opsin gene in the honey bee, *Apis mellifera*. Eur. J. Biochem. **243**:775–781.

Briscoe, A. D. 1998. Molecular diversity of visual pigments in the butterfly *Papilio glaucus*. Naturwissenschaften **85**:33–35.

———. 1999. Intron splice sites of *Papilio glaucus PglRh3* corroborate insect opsin phylogeny. Gene **230**:101–109.

———. 2000. Six opsins from the butterfly *Papilio glaucus:* molecular phylogenetic evidence for multiple origins of red-sensitive visual pigments in insects. J. Mol. Evol. **51**:110–121.

———. 2001. Functional diversification of lepidopteran opsins following gene duplication. Mol. Biol. Evol. **18**: 2270–2279.

———. 2002. Homology modeling suggests a functional role for parallel amino acid substitutions between bee and butterfly red- and green-sensitive opsins. Mol. Biol. Evol. **19**:983–986.

Briscoe, A. D., and L. Chittka. 2001. The evolution of color vision in insects. Annu. Rev. Entomol. **46**:471–510.

Briscoe, A. D., and L. Nagy. 1999. Spatial expression of opsins in the retina and brain of the tiger swallowtail *Papilio glaucus*. Am. Zool. **39**:254B.

Cameron, S. A., and P. Mardulyn. 2001. Multiple molecular data sets suggest independent origins of highly eusocial behavior in bees (Hymenoptera:Apinae). Syst. Biol. **50**:194–214.

Cameron, S. A., and P. H. Williams. 2003. Phylogeny of bumble bees in the New World subgenus *Fervidobombus* (Hymenoptera: Apidae): congruence of molecular and morphological data. Mol. Phylogenet. Evol. **28**:552–563.

Chang, B. S. W., D. M. Ayers, W. C. Smith, and N. E. Pierce. 1996. Cloning of the gene encoding honeybee long-wavelength rhodopsin: a new class of insect visual pigments. Gene **173**:215–219.

Cook, J. M., A. Rokas, M. Pagel, and G. N. Stone. 2002. Evolutionary shifts between host oak sections and host-plant organs in *Andricus* gallwasps. Evolution **56**:1821–1830.

Crandall, K., and T. Cronin. 1997. The molecular evolution of visual pigments of freshwater crayfishes (Decapoda: Cambaridae). J. Mol. Evol. **45**:524–534.

Danforth, B. N., L. Conway, and S. Ji. 2003. Phylogeny of eusocial *Lasioglossum* reveals multiple losses of eusociality within a primitively eusocial clade of bees (Hymenoptera: Halictidae). Syst. Biol. **52**:23–36.

Filipek, S., R. E. Stenkamp, D. C. Teller, and K. Palczewski. 2003. G protein-coupled receptor rhodopsin: a prospectus. Ann. Rev. Physiol. **65**:851–879.

Gärtner, W. 2000. Invertebrate visual pigments. Pp. 298–387 *in* D. G. Stavenga, W. J. DeGrip, and E. N. Pugh Jr., eds. Handbook of biological physics. Elsevier, Amsterdam.

Gaucher, E. A., X. Gu, M. M. Miyamoto, and S. A. Benner. 2002. Predicting functional divergence in protein evolution by site-specific rate shifts. Trends Biochem. Sci. **27**:315–321.

Gleadall, I., T. Hariyama, and Y. Tsukahara. 1989. The visual pigment chromophores in the retina of insect compound eyes, with special reference to the Coleoptera. J. Insect Physiol. **35**:787–795.

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019

Graf, R., A. Godknecht, M. Nakano, X. Li, U. Ackermann, and P. Helbling. 1996. Cloning and expression analysis of *Aedes aegypti* opsin: adaptation of an in situ hybridization protocol for mosquitoes. Insect Mol. Biol. **5**:173–180.

Gu, X. 1999. Statistical methods for testing functional divergence after gene duplication. Mol. Biol. Evol. **16**:1664–1674.

———. 2001. Maximum-likelihood approach for gene family evolution under functional divergence. Mol. Biol. Evol. **18**:453–464.

Gu, X., and K. Vander Velden. 2002. DIVERGE: phylogeny-based analysis for functional-structural divergence of a protein family. Bioinformatics **18**:500–501.

Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52**:696–704.

Han, M., M. Groesbeek, S. O. Smith, and T. P. Sakmar. 1998. Role of the C9 methyl group in rhodopsin activation: characterization of mutant opsins with the artificial chromophore 11-cis-9-demethylretinal. Biochemistry **37**:538–545.

Hill, C. A., A. N. Fox, R. J. Pitts, L. B. Kent, P. L. Tan, M. A. Chrystal, A. Cravchik, F. H. Collins, H. M. Robertson, and L. J. Zwiebel. 2002. G protein-coupled receptors in *Anopheles gambiae*. Science **298**:176–178.

Holt, R. A., G. M. Subramanian, A. Halpern et al. (and 120 co-authors) 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science **298**:129–149.

Hsu, R., A. D. Briscoe, B. S. W. Chang, and N. E. Pierce. 2001. Molecular evolution of a long wavelength opsin in mimetic *Heliconius* butterflies (Lepidoptera: Nymphalidae). Biol. J. Linn. Soc. **72**:435–449.

Hunt, D. M., S. E. Wilkie, J. K. Bowmaker, and S. Poopalasundaram. 2001. Vision in the ultraviolet. Cell Mol. Life Sci. **58**:1583–1598.

Jiggins, F. M. 2003. Male-killing Wolbachia and mitochondrial DNA: selective sweeps, hybrid introgression and parasite population dynamics. Genetics **164**:5–12.

Kawakita, A., T. Sota, J. S. Ascher, M. Ito, H. Tanaka, and M. Kato. 2003. Evolution and phylogenetic utility of aligned gaps within intron sequences of three nuclear genes in bumble bees (*Bombus*). Mol. Biol. Evol. **20**:87–92.

Kitamoto, J., K. Ozaki, and K. Arikawa. 2000. Ultraviolet and violet receptors express identical mRNA encoding an ultraviolet-absorbing opsin: identification and histological localization of two mRNAs encoding short-wavelength-absorbing opsins in the retina of the butterfly *Papilio xuthus*. J. Exp. Biol. **203**:2887–2894.

Kitamoto, J., K. Sakamoto, K. Ozaki, Y. Mishina, and K. Arikawa. 1998. Two visual pigments in a single photoreceptor cell: identification and histological localization of three mRNAs encoding visual pigment opsins in the retina of the butterfly *Papilio xuthus*. J. Exp. Biol. **201**:1255–1261.

Kochendoerfer, G. G., W. L. Steven, T. P. Sakmar, and R. A. Maties. 1999. How color visual pigments are tuned. Trends Biochem. Sci. **24**:300–305.

Kumar, S., and S. R. Gadagkar. 2001. Disparity index: a simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. Genetics **158**:1321–1327.

Kumar, S., K. Tamura, and M. Nei. 2003. MEGA3: An integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. **5**:150–163.

Li, W. H., C. I. Wu, and C. C. Luo. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol. Biol. Evol. **2**:150–174.

Lin, C. P., and B. N. Danforth. 2004. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. Mol. Phylogenet. Evol. **30**:686–702.

Maddison, W. P., and D. R. Maddison. 2000. MacClade 4: analysis of phylogeny and character evolution. Sinauer Associates, Sunderland, Mass.

Malpel, S., A. Klarsfeld, and F. Rouyer. 2002. Larval optic nerve and adult extra-retinal photoreceptors sequentially associate with clock neurons during *Drosophila* brain development. Development **129**:1443–1453.

Mardulyn, P., and S. A. Cameron. 1999. The major opsin in bees (Insecta: Hymenoptera): a promising nuclear gene for higher level phylogenetics. Mol. Phylogenet. Evol. **12**: 168–176.

Merbs, S. L., and J. Nathans. 1993. Role of hydroxyl-bearing amino acids in differentially tuning the absorption spectra of the human red and green cone pigments. Photochem. Photobiol. **58**:706–710.

Nei, M., and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous substitutions. Mol. Biol. Evol. **3**:418–426.

Ortiz-Rivas, B., A. Moya, and Martinez-Torres D. 2004. Molecular systematics of aphids (Homoptera : Aphididae): new insights from the long-wavelength opsin gene. Mol. Phylogenet. Evol. **30**:24–37.

Palczewski, K., T. Kumasaka, T. Hori et al. (12 co-authors). 2000. Crystal structure of rhodopsin: A G protein-coupled receptor. Science **289**:739–745.

Peitsch, D., A. Feitz, H. Hertel, J. de Souza, D. F. Ventura, and R. Menzel. 1992. The spectral input systems of hymenopteran insects and their receptor-based colour vision. J. Comp. Physiol. A **170**:23–40.

Rice, K. A. 1994. The origin, evolution and classification of G protein-coupled receptors. PhD thesis, Harvard University, Cambridge, Mass.

Salcedo, E., L. Zheng, M. Phistry, E. Bagg, and S. G. Britt. 2003. Molecular basis for ultraviolet vision in invertebrates. J. Neurosci. **23**:10873–10878.

Seki, T., and K. Vogt. 1998. Evolutionary aspects of the diversity of visual pigment chromophores in the class Insecta. Comp. Biochem. Physiol. **119**:53–64.

Shimizu, I., Y. Yamakawa, Y. Shimazaki, and T. Iwasa. 2001. Molecular cloning of *Bombyx* cerebral opsin (*Boceropsin*) and cellular localization of its expression in the silkworm brain. Biochem. Biophys. Res. Commun. **287**:27–34.

Shyue, S. K., S. Boissinot, H. Schneider et al. (10 co-authors). 1998. Molecular genetics of spectral tuning in new world monkey color vision. J. Mol. Evol. **46**:697–702.

Swofford, D. L. 1998 PAUP*: phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, Mass.

Teller, D., T. Okada, C. Behnke, K. Palczewski, and R. Stenkamp. 2001. Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). Biochemistry **40**: 7761–7772.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

Tian, H. S., C. L. Zhu, X. H. Gao, L. Ma, B. Shen, X. L. Li, and G. L. Wu. 2001. Cloning and identification of deltamethrin-resistance or susceptibility associated genes of *Culex pipiens pallens*. Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi **19**:193–197.

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019

Townson, S., B. Chang, E. Salcedo, L. Chadwell, N. Pierce, and S. G. Britt. 1998. Honeybee blue- and ultraviolet-sensitive opsins: cloning, heterologous expression in *Drosophila*, and physiological characterization. J. Neurosci. **18**:2412–2422.

Wiegmann, B. M., D. K. Yeates, J. L. Thorne, and H. Kishino. 2003. Time flies, a new molecular time-scale for Brachyceran fly evolution without a clock. Syst. Biol. **52**:745–756.

Yang, Z. H., and J. P. Bielawski. 2000. Statistical methods for detecting molecular adaptation. Trends Ecol. Evol. **15**: 496–503.

Yasuyama, K., and I. Meinertzhagen. 1999. Extraretinal photoreceptors at the compound eye's posterior margin in *Drosophila melanogaster*. J. Comp. Neurol. **412**:193–202.

Zhukovsky, E. A., P. R. Robinson, and D. D. Oprian. 1992. Changing the location of the Schiff base counterion in rhodopsin. Biochemistry **31**:10400–10405.

David Irwin, Associate Editor

Accepted May 5, 2004

Downloaded from https://academic.oup.com/mbe/article-abstract/21/8/1583/1060638 by guest on 24 December 2019