MBE Advance Access published March 5, 2014

Divergent Positive Selection in Rhodopsin from Lake and Riverine Cichlid Fishes

Ryan K. Schott,[†,1] Shannon P. Refvik,[†,1] Frances E. Hauser,[1] Hernán López-Fernández,[1,2] and Belinda S.W. Chang[*,1,3,4]

[1]Department of Ecology & Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada

[2]Department of Natural History, Royal Ontario Museum, Toronto, Ontario, Canada

[3]Department of Cell & Systems Biology, University of Toronto, Toronto, Ontario, Canada

[4]Centre for the Analysis of Genome Evolution & Function, University of Toronto, Toronto, Ontario, Canada

[†]These authors contributed equally to this work.

***Corresponding author:** E-mail: belinda.chang@utoronto.ca.

Associate editor: Juliette de Meaux

Abstract

Studies of cichlid evolution have highlighted the importance of visual pigment genes in the spectacular radiation of the African rift lake cichlids. Recent work, however, has also provided strong evidence for adaptive diversification of riverine cichlids in the Neotropics, which inhabit environments of markedly different spectral properties from the African rift lakes. These ecological and/or biogeographic differences may have imposed divergent selective pressures on the evolution of the cichlid visual system. To test these hypotheses, we investigated the molecular evolution of the dim-light visual pigment, rhodopsin. We sequenced rhodopsin from Neotropical and African riverine cichlids and combined these data with published sequences from African cichlids. We found significant evidence for positive selection using random sites codon models in all cichlid groups, with the highest levels in African lake cichlids. Tests using branch-site and clade models that partitioned the data along ecological (lake, river) and/or biogeographic (African, Neotropical) boundaries found significant evidence of divergent selective pressures among cichlid groups. However, statistical comparisons among these models suggest that ecological, rather than biogeographic, factors may be responsible for divergent selective pressures that have shaped the evolution of the visual system in cichlids. We found that branch-site models did not perform as well as clade models for our data set, in which there was evidence for positive selection in the background. One of our most intriguing results is that the amino acid sites found to be under positive selection in Neotropical and African lake cichlids were largely nonoverlapping, despite falling into the same three functional categories: spectral tuning, retinal uptake/release, and rhodopsin dimerization. Taken together, these results would imply divergent selection across cichlid clades, but targeting similar functions. This study highlights the importance of molecular investigations of ecologically important groups and the flexibility of clade models in explicitly testing ecological hypotheses.

Key words: **Neotropical cichlids, evolution of vision, visual pigment, evolution of protein structure and function, codon substitution model, d_N/d_S, clade model.**

Introduction

Aquatic organisms contend with complex photic environments where incident brightness, depth, and turbidity affect the nature of light available for vision (Lythgoe 1979). In fish species, aspects of visual pigment function, particularly wavelength of maximum absorbance (λ_{max}), have been found to correlate with spectral environment, suggesting that the environment can impose selective pressures on the visual system (Bowmaker 1995). This is especially evident in cichlid fishes, where molecular evolution and differential expression of visual pigment genes have been implicated in the adaptive radiation of the African rift lake taxa (Seehausen et al. 2008). In these fishes, mutations causing shifts in λ_{max} of cone visual pigments are associated with depth and male nuptial coloration, promoting sympatric divergence of color morphs within cichlid populations through sensory drive (Carleton et al. 2005; Seehausen et al. 2008) and affecting ecological divergence during adaptive radiation (Spady et al. 2005; Terai et al. 2006; Miyagi et al. 2012). African lake cichlids have thus emerged as a model system for the study of visual system molecular evolution (Spady et al. 2005; Carleton 2009; Nagai et al. 2011), but studies to date have not incorporated analyses of any riverine lineages, potentially limiting our understanding of visual evolution of cichlids living in starkly different environments.

Neotropical cichlid fishes are ubiquitous throughout the ecologically varied riverine habitats of South and Central America (Reis et al. 2003; López-Fernández et al. 2012, 2013) and are a monophyletic clade sister to the African cichlids (Stiassny 1991; Farias et al. 2000; Sparks and Smith 2004). In contrast to the recent diversification of African cichlids, this group underwent ancient adaptive radiation in the rivers of South America (Arbour and López-Fernández 2013; López-Fernández et al. 2013). Although Neotropical cichlids are

Article

almost exclusively riverine, there is a genus of Neotropical cichlids that has more recently radiated within Central American crater lakes (e.g., Barluenga et al. 2006). Neotropical cichlids are less species-rich than African rift lake cichlids; however, they are characterized by high levels of morphological, ecological, and reproductive diversity, much like their African relatives (e.g., Wimberger et al. 1998; Barlow 2000; López-Fernández et al. 2012, 2013). Much of this diversity is represented in the tribe Geophagini: with 18 genera, this clade includes piscivorous species, substrate sifters, and water-column feeders, as well as species that mouth brood their young (e.g., López-Fernández et al. 2005, 2012, 2013). The riverine habitats of Neotropical cichlid fishes also exhibit stark contrasts in spectral quality with black, white, and clear water types that result in differential degrees of light attenuation among habitats (Sioli 1984; Cooke et al. 2012). How these differences in water type and life history may have affected the evolution and ecology of vision in Neotropical cichlids, and how this compares with African lake and riverine cichlids, is not known.

Vision is initiated with the absorption of light by visual pigment complexes, which consist of a chromophore (retinal) covalently bound to an opsin protein (Wald 1968), a member of the G protein-coupled receptor (GPCR) superfamily (Hofmann et al. 2009). There are five major classes of visual opsins in vertebrates, each of which maximally absorbs a characteristic wavelength of light. Four classes of cone opsins mediate bright-light vision, whereas rhodopsin (RH1) mediates dim-light vision (Bowmaker 2008). Prior to this study, the only Neotropical cichlid in which visual pigment genes had been sequenced was the Trinidadian geophagine *Crenicichla frenata* (Weadick et al. 2012), due to its relevance as a guppy predator (Houde 1997; Magurran 2005). *Crenicichla frenata* was found to have only five opsin genes compared with the eight in African cichlids, and both rhodopsin and one cone opsin were found to be under positive selection (Weadick et al. 2012). Previous work has also found that opsin genes are under varying selective pressures in African cichlids, including strong positive selection on rhodopsin (Spady et al. 2005; Sugawara et al. 2010). Because the visual systems of Neotropical cichlids and African riverine cichlids have not been as well studied as the African rift lake cichlids, it is unclear whether the patterns of positive selection seen in *C. frenata* are due to differences in the selective pressure of lake and river habitats (ecology) or differences in phylogenetic and/or geographic history between African and Neotropical clades (biogeography).

Codon-based likelihood models, which have been the target of much recent development (reviewed in Anisimova and Kosiol 2009), have proven extremely useful in examining selective pressures in a variety of systems (Yang et al. 2000; Swanson et al. 2001; Fay and Wu 2003; Bakewell et al. 2007; Briscoe et al. 2010; Shen et al. 2010; Khan et al. 2011; Moury and Simon 2011; Yang and dos Reis 2011; Rennison et al. 2012). Increasing attention, however, has been directed to testing for divergent selection among clades using either branch-site models (Spady et al. 2005; Ramm et al. 2008; Yoshida et al. 2011; Smith et al. 2012; Badouin et al. 2013; Veilleux et al. 2013) or clade models (Zhao et al. 2009; Yoshida et al. 2011; Weadick and Chang 2012a, 2012b). Although the use of the branch-site test in identifying instances of episodic selection along prespecified lineages has been extensively tested (Yang and dos Reis 2011), its use for detecting divergent selection among clades has not been evaluated and rarely has the same data set been examined using both the branch-site and clade models (Yoshida et al. 2011).

Recent criticisms of codon-based likelihood models have suggested that certain positively selected sites have no functional significance (Yokoyama et al. 2008). This criticism was specific to an analysis of selection in rhodopsin and assumed that spectral tuning was the only aspect of rhodopsin function that could be under positive selection. Rhodopsin, however, provides a unique opportunity to explore a variety of functions of positively selected sites because its crystal structure has been resolved in a number of conformations, including both inactive and active states (Palczewski et al. 2000; Choe et al. 2011). Moreover, rhodopsin has been experimentally studied in detail for many functional aspects other than simply spectral tuning, for example, thermal stability (Janz et al. 2003; Yan et al. 2002), retinal uptake and release rates (Hildebrand et al. 2009; Hunt et al. 2009; Piechnick et al. 2012), and physicochemical properties of ultrafast retinal isomerization (Prokhorenko et al. 2006; Gozem et al. 2012). Applying this extensive body of research on rhodopsin structure and function to the study of positively selected sites in cichlids will provide new insights into the adaptive evolution of visual pigments in these fish.

To begin investigating evolutionary differences between African and Neotropical cichlid visual pigments, we sequenced rhodopsin (RH1) from 30 species of Neotropical riverine cichlid, as well as two African riverine species, and compared them with available sequences for African cichlids and *C. frenata*. We hypothesize that differences in biogeographic history and/or ecological differences between lake and river habitats among Neotropical riverine, African riverine, and African rift lake cichlids may have resulted in divergent selective pressures on the rhodopsin gene. To test this hypothesis, we used a variety of codon-based sites and branch models of molecular evolution, including recently developed multiclade models (Yoshida et al. 2011), as well as newly implemented improvements to existing models (Chang et al. 2012; Weadick and Chang 2012a), and compared the results from the various methods. Models that incorporate independently estimated rates of synonymous substitutions were also utilized to determine what effect, if any, the addition of variable synonymous rates had on the results. This collection of models was used to explicitly test for differences in selective pressure between African and Neotropical clades and between lake and river habitats.

Results

Rhodopsin Sequencing and Phylogenetic Analyses

An 859 base pair fragment of rhodopsin (representing 81% of the gene, including the seven transmembrane helices) was amplified from 30 species of Neotropical river cichlid and

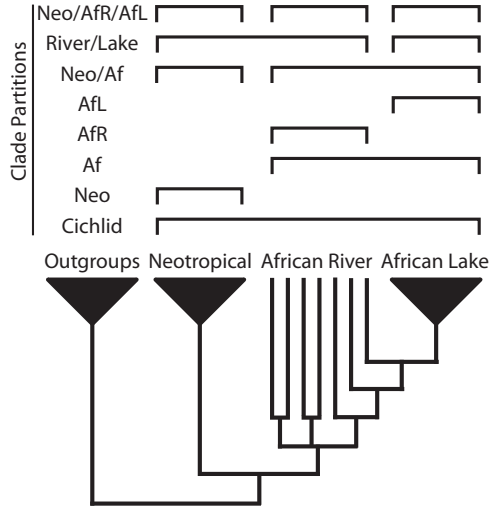

Fig. 1. Simplified phylogeny of cichlids based on the Bayesian and ML RH1 gene trees (supplementary figs. S1 and S2, Supplementary Material online). Each of the clade partitions tested by the branch-site and clade models are demarcated and named. In each case, an additional partition exists that includes all taxa that are not part of the illustrated partitions. Af, African cichlids; AfL, African lake cichlids, AfR, African river cichlids; Neo, Neotropical cichlids.

two African riverine cichlids. In addition to this, 38 African lake and riverine cichlid sequences, the sequence from the Neotropical river cichlid *C. frenata*, as well as sequences from eight noncichlid outgroups, were obtained from Genbank (supplementary table S1, Supplementary Material online).

Results from the phylogenetic analyses of the complete RH1 alignment (cichlids and eight outgroup taxa; fig. 1, supplementary figs. S1 and S2, Supplementary Material online) were generally congruent with accepted species relationships for Neotropical (Smith et al. 2008; López-Fernández et al. 2010), African river (Schwarzer et al. 2009), and African lake cichlids (Genner and Turner 2012; Muschick et al. 2012; Wagner et al. 2013). The Bayesian and maximum likelihood (ML) topologies were similar, but not identical, especially with regard to the topology of some of the African lake species, which have very short branch lengths (supplementary figs. S1 and S2, Supplementary Material online). Both results were used in subsequent analyses to confirm the robustness of the results to minor changes in topology.

Molecular Evolutionary Analyses

To explore variation in selective pressure among cichlid rhodopsins, we used codon-based models of evolution as implemented in PAML (Yang 2007) and HYPHY (Pond et al. 2005) on various data subsets (and partitions) that reflect differences in ecology and/or biogeographical history: lake/river, African/Neotropical (see Materials and Methods for details). The random sites PAML models confirmed that there is variation in ω (the ratio of nonsynonymous to synonymous substitution rate, d_N/d_S; M3 vs. M0, $P < 0.0001$, table 1, supplementary tables S2 and S3, Supplementary Material online) for all alignment subsets, as expected for functional protein coding genes. These models also found significant evidence for a positively selected class of sites (M2a vs. M1a, M7 vs. M8, $P < 0.0001$; table 1, supplementary tables S2 and S3, Supplementary Material online) for all subsets except African riverine cichlids under the Bayesian topology, with ω values significantly greater than 1 (M8 vs. M8a, $P < 0.0001$; table 1, supplementary tables S2 and S3, Supplementary Material online). Under the ML topology, the African river subset was also found to be under significant positive selection (M2a vs. M1a, and M8 vs. M7 and M8a, $P < 0.05$; supplementary table S3, Supplementary Material online). The two topologies only differed in the location of a single African river taxon (*Tilapia guineensis*), but our sample size for African riverine cichlids was small ($n = 8$) and thus small changes in topology likely had a disproportionate effect. The PARRIS test (Scheffler et al. 2006), which is similar to the M2a vs. M1a comparison, but allows variation in synonymous rates (d_S), was also significant in all cases, except for African river subset, confirming that variation in d_S did not affect our inferences of positive selection in cichlids (supplementary tables S4 and S5, Supplementary Material online).

African lake cichlids had the highest level of ω for the positively selected site class (13.69/14.37 in 7.1/6.6% of sites with the Bayesian/ML topologies, respectively, under M8), whereas Neotropical cichlids had a lower, but still significantly positive, value (3.60/3.66 in 5.0/4.9% of sites with the Bayesian/ML topologies, respectively, under M8; table 1, supplementary table S3, Supplementary Material online). African river cichlids showed a pattern of positive selection intermediate between this (ω of 2.78/3.07 at 7.1/6.9% of sites for the Bayesian/ML topologies, under M8). It is interesting to note that the strength of positive selection was similar between African river cichlids and Neotropical cichlids but that the proportion of sites under positive selection was more similar between African lake and African river cichlids. A previous study of rhodopsin in African rift lake cichlids and a single African riverine outgroup reported 6.9% of sites under positive selection, with an average ω value of 14.07 (M8; Spady et al. 2005), which is very similar to our findings. Overall ω values for cichlid rhodopsin are quite high (M0, 0.25) compared with typical values found in ray-finned fish rhodopsins ($\omega = 0.07–0.08$; Rennison et al. 2012) and even protein coding genes in general ($\omega = 0.08–0.18$; Fay and Wu 2003). Instead, they are similar to values found for genes known to be under strong positive selection such as human MHC and reproductive proteins (Swanson et al. 2001).

Despite the high levels of positive selection found in this data set, branch-site and clade model analyses did not reveal evidence for a burst of selection occurring along the lineages leading to the diversification of the major clades (cichlids, African cichlids, and Neotropical cichlids; supplementary table S6, Supplementary Material online). The only significant result was found using the branch-site model on the lineage leading to the African cichlids as foreground; the clade model C (CmC) test on the same branch was not significant. Furthermore, the null model for CmC, M2a_rel, was found

Table 1. Results of Random Sites (PAML) Analyses on Various Subsets of the Cichlid RH1 Bayesian Gene Tree.

Tree[a]	Model	np	ln *L*	κ	Parameters[b]			Null	LRT	df	*P*
					ω_0/p	ω_1/q	ω_2/ω_p				
Neo	M0	55	−3,076.84	2.81	0.25			NA			
	M1a	56	−2,942.75	2.44	0.01 (84.7%)	1 (15.3%)		M0	268.17	1	0.0000
	M2a	58	−2,923.18	2.75	0.01 (85.2%)	1 (9.9%)	3.63 (4.9%)	M1a	39.15	2	0.0000
	M3	59	−2,922.87	2.76	0.02 (87.2%)	1.51 (10.5%)	5.28 (2.4%)	M0	307.94	4	0.0000
	M7	56	−2,945.05	2.38	0.01	0.08		NA			
	M8a	57	−2,942.76	2.44	0.56	99	1 (15.3%)	NA			
	M8	58	−2,923.17	2.74	0.01	0.07	3.60 (5.0%)	M7	43.77	2	0.0000
								M8a	39.18	1	0.0000
AfR	M0	15	−1,551.94	2.57	0.15			NA			
	M1a	16	−1,540.64	2.36	0 (87.6%)	1 (12.4%)		M0	22.60	1	0.0000
	M2a	18	−1,538.19	2.42	0 (92.9%)	1 (0%)	2.78 (7.1%)	M1a	4.90	2	0.0863
	M3	19	−1,538.19	2.42	0 (84.2%)	0 (8.7%)	2.78 (7.1%)	M0	27.50	4	0.0000
	M7	16	−1,540.91	2.29	0.01	0.05		NA			
	M8a	17	−1,540.64	2.36	0.01	99	1 (12.4%)	NA			
	M8	18	−1,538.19	2.42	0.01	99	2.78 (7.1%)	M7	5.43	2	0.0661
								M8a	4.90	1	0.0269
AfL	M0	59	−2,232.90	2.68	1.22			NA			
	M1a	60	−2,187.22	1.89	0 (72.0%)	1 (28.0%)		M0	91.35	1	0.0000
	M2a	62	−2,121.61	2.33	0 (63.2%)	1 (29.8%)	14.05 (7.0%)	M1a	131.22	2	0.0000
	M3	63	−2,120.61	2.38	0.07 (79.3%)	2.55 (14.9%)	17.20 (5.7%)	M0	224.58	4	0.0000
	M7	60	−2,187.33	1.91	0.01	0.01		NA			
	M8a	61	−2,187.22	1.89	0.01	2.09	1 (28.0%)	NA			
	M8	62	−2,121.65	2.32	0.01	0.01	13.69 (7.1%)	M7	131.36	2	0.0000
								M8a	131.14	1	0.0000

NOTE.—np, number of parameters; ln *L*, *ln* likelihood; κ, transition/transversion ratio; df, degrees of freedom; NA, not applicable. Additional subsets are shown in supplementary table S2, Supplementary Material online.

[a]The gene tree from the full RH1 alignment was pruned to contain only Neotropical cichlids (Neo), African river cichlids (AfR), and African lake cichlids (AfL).

[b]ω values of each site class are shown for models M0–M3 (ω_0 – ω_2) with the proportion of each site class in parentheses. For M7–M8, the shape parameters, *p* and *q*, which describe the beta distribution are listed. In addition, the ω value for the positively selected site class (ω_p, with the proportion of sites in parentheses) is shown for M8a (where ω_p is constrained to equal one) and M8.

to be a better fit than the branch-site model and the best-fitting model overall by Akaike information criterion (AIC) comparison (supplementary table S6, Supplementary Material online). This may be due to the fact that branch-site models do not allow positive selection in the "background," which is likely to be the case for this data set. Together, these results indicate that the divergent selective pressures found using the random sites models was not driven by selection as each group invaded a new environment. This suggests that instead positive selection was acting within each clade during its diversification.

CmC (Bielawski and Yang 2004) was used to analyze variation among clades. Specifically, we used this model to test for divergent selection by partitioning the data into biogeographic (African, Neotropical) or ecological (lake, river) groupings (fig. 1). In all cases, allowing for divergent selection resulted in a significant improvement over the null model ($P < 0.0001$, for all tests; table 2 and supplementary table S7, Supplementary Material online), confirming that there is significant divergent selective pressure between cichlids and the outgroups, and within cichlid groups. Comparison with a null model in which the divergent site class was constrained to equal one confirmed the presence of significant positive selection in all cases, except for the most heavily partitioned models (supplementary table S8, Supplementary Material online). Because we are interested in comparisons among models that partition the data along different ecological/biogeographic boundaries, we used the novel approach of comparing the results of different levels of partitioning to each other using likelihood ratio tests (LRTs), for those models that were fully nested, and by AIC, for those that were not, to determine which method of partitioning resulted in a better fit to the data. Among all clade models with two partitions, the best-fitting model was found to be the one that partitioned African lake cichlids from all other taxa (table 2, supplementary table S7, Supplementary Material

Table 2. Results of CmC (PAML) Analyses on the Cichlid RH1 Bayesian Gene Tree.

Model and Partition[a]	np	ln *L*	κ	Parameters[b]			ΔAIC[c]	Null	LRT	df	*P*
				ω_0	ω_1	ω_2/ω_d					
M1a	147	−6,239.50	2.24	0.25			154.86	NA			
M2a	149	−6,186.53	2.46	0.04 (80.1%)	1 (16.6%)	4.27 (3.3%)	52.91	M1a	105.95	2	0.0000
M2a_rel	149	−6,186.53	2.46	0.04 (80.1%)	1 (16.6%)	4.27 (3.3%)	52.91	M1a	105.95	2	0.0000
CmC: Cichlid	150	−6,179.22	2.44	0.04 (80.3%)	1 (14.3%)	1.33 (5.4%) Cichlid: 4.28	40.30	M2a_rel	14.61	1	0.0001
CmC: Neo	150	−6,179.29	2.49	0.04 (80.0%)	1 (16.4%)	5.92 (3.6%) Neo: 2.33	40.44	M2a_rel	14.61	1	0.0001
CmC: Af	150	−6,165.68	2.41	0.04 (80.3%)	1 (14.8%)	1.90 (4.9%) Af: 8.41	13.21	M2a_rel	41.70	1	0.0000
CmC: AfR	150	−6,186.35	2.47	0.04 (80.1%)	1 (16.7%)	4.40 (3.3%) AfR: 3.34	54.56	M2a_rel	0.35	1	0.5533
CmC: AfL	150	−6,160.32	2.40	0.04 (80.6%)	1 (13.7%)	1.60 (5.7%) AfL: 11.91	2.50	M2a_rel	52.41	1	0.0000
CmC: Neo/Af	151	−6,165.1	2.41	0.04 (80.3%)	1 (14.7%)	1.63 (5.0%) Neo: 2.15 Af: 8.40	14.05	M2a_rel Af	42.86 1.15	2 1	0.0000 0.2827
CmC: R/L	151	−6,158.07	2.40	0.04 (80.7%)	1 (13.1%)	1.12 (6.1%) R: 1.91 L: 11.68	0	M2a_rel Af AfL	56.91 15.21 4.50	2 1 1	0.0000 0.0001 0.0339
CmC: Neo/AfR/AfL	152	−6,157.48	2.40	0.04 (80.8%)	1 (12.8%)	1.16 (6.4%) Neo: 1.51 AfR: 2.97 AfL: 11.72	0.82	M2a_rel AfL R + L	58.09 5.68 1.18	3 2 1	0.0000 0.0584 0.2769

NOTE.—np, number of parameters; ln *L*, ln likelihood; κ, transition/transversion ratio; df, degrees of freedom; NA, not applicable.

[a]Partitions listed are explained in figure 1. In all cases, an additional partition exist that contains the remaining taxa (e.g., outgroups).

[b]ω values of each site class are shown with the proportion of each site class in parentheses. ω_D is divergent site class that has a separate value for each partition.

[c]The difference in AIC values was calculated compared with the overall best-fitting model, R/L, with an AIC of 12,618.14.

online). This is not surprising due to the extremely high ω found for African lake cichlids in the random sites analysis, although the ω estimated from the clade model was lower (11.91/12.49, Bayesian and ML, respectively; supplementary table S7, Supplementary Material online), likely due to the constraint on the proportion of sites in the divergent site class among partitions. The best three-partition model was a division between lake cichlids, river cichlids, and outgroups, and this was a significantly better fit than the two partition African lake model ($P < 0.0001$, supplementary table S7, Supplementary Material online). Adding an additional partition to divide the African river cichlids, African lake cichlids, Neotropical cichlids, and outgroups from each other was not a significantly better fit than the lake, river, outgroup model. This agrees with the findings of our random site analyses where the Neotropical and African river cichlids were found to have similar levels of positive selection, compared with the much higher levels found in African lake cichlids. Thus, the three-partition lake, river, outgroup was the best fitting model suggesting that ecological differences between lake and river habitats, rather than biogeographic factors, may have been the cause of divergent selective pressures among cichlids. However, future studies with additional taxonomic sampling of African riverine species and other Neotropical cichlid clades, especially Neotropical lake cichlids, may reveal more complicated patterns than suggested by our current data.

As recent studies have increasingly used branch-site models to study the evolution of multiple lineages or even entire clades (e.g., Spady et al. 2005; Ramm et al. 2008; Yoshida et al. 2011; Smith et al. 2012; Badouin et al. 2013; Veilleux et al. 2013), we performed the same two-partition tests (cichlid, African cichlids, African lake cichlids, and Neotropical cichlids, partitioned from the rest of the taxa) that were done with CmC using the branch-site model. The branch-site model cannot contain more than two partitions, so we could not replicate our higher partitioned tests. Each of the two-partition branch-site models was found to fit significantly better than the null model ($P < 0.0001$, table 3, supplementary table S9, Supplementary Material online) and the African lake partition was found to be the best-fitting branch-site model, similar to the CmC analysis. However, the CmC lake, river, outgroup partitioned model was still the overall best-fitting model by AIC comparison. This is likely due to the branch-site model not being able to accommodate more than two partitions and not allowing for positive selection in the background, which appears to be the case for this data set. In all cases, the fits of the CmC models were better than those of the branch-site models (although only very slightly

Table 3. Results of Branch-Site (BrS, PAML) Analyses on the RH1 Cichlid Gene Tree Highlighting Clades, in Comparison to the CmC Results.

Partition[a]	Model	np	ln *L*	κ	Parameters[b]				ΔAIC[c]	Null	LRT	df	*P*
					ω_0	ω_1	ω_{2a}	ω_{2b}					
NA	M2a_rel	149	−6,186.5	2.5	0.04 (80.1%)	1 (16.6%)	4.27 (3.3%)		50.4	NA			
Cichlid	BrS_Null	148	−6,222.8	2.2	B: 0.03 F: 0.03 (80.0%)	B: 1 F: 1 (11.4%)	B: 0.03 F: 1 (7.8%)	B: 1 F: 1 (1.1%)	120.9	NA			
	BrS_Alt	149	−6,193.3	2.4	B: 0.04 F: 0.04 (78.1%)	B: 1 F: 1 (16.8%)	B: 0.04 F: 4.16 (4.3%)	B:1 F: 4.16 (0.9%)	64.0	BrS_Null	58.9	1	0.000
	CmC	150	−6,179.2	2.4	B: 0.04 F: 0.04 (80.3%)	B: 1 F: 1 (14.3%)	B: 1.33 F: 4.28 (5.4%)		37.8	M2a_rel	14.6	1	0.000
Neo	BrS_Null	148	−6,229.3	2.2	B: 0.03 F: 0.03 (79.9%)	B: 1 F: 1 (15.8%)	B: 0.03 F: 1 (3.6%)	B: 1 F: 1 (0.72%)	134.0	NA			
	BrS_Alt	149	−6,212.4	2.3	B: 0.04 F: 0.04 (80.1%)	B: 1 F: 1 (15.7%)	B: 0.04 F: 4.10 (3.5%)	B: 1 F: 4.10 (0.69%)	102.1	BrS_Null	33.9	1	0.000
	CmC	150	−6,179.3	2.5	B: 0.04 F: 0.04 (80.0%)	B: 1 F: 1 (16.4%)	B: 5.92 F: 2.33 (3.6%)		37.9	M2a_rel	14.5	1	0.000
AfL	BrS_Null	148	−6,217.6	2.2	B: 0.03 F: 0.03 (75.0%)	B: 1 F: 1 (14.3%)	B: 0.03 F: 1 (9.0%)	B: 1 F: 1 (1.7%)	110.6	NA			
	BrS_Alt	149	−6,161.4	2.3	B: 0.03 F: 0.03 (76.2%)	B: 1 F: 1 (14.4%)	B: 0.03 F: 9.74 (7.9%)	B: 1 F: 9.74 (1.5%)	0.1	BrS_Null	112.5	1	0.000
	CmC	150	−6,160.3	2.4	B: 0.04 F: 0.04 (80.6%)	B: 1 F: 1 (13.7%)	B: 1.60 F: 11.91 (5.7%)		0	M2a_rel	52.4	1	0.000

NOTE.—Additional partitions are shown in supplementary table S9, Supplementary Material online.

[a]Paritions listed are explained in figure 1. In all cases an additional partition exits that contains the remaining taxa (referred to as the background in the branch-site model).

[b]ω values of each site class are shown with the proportion of each site class in parentheses. B and F refer to the background and foreground partitions using the branch-site terminology.

[c]Minimum overall AIC (African Lake partition of CmC; 12,620.6) was used for all comparisons.

Abbreviations—np, number of parameters; ln *L*, ln likelihood; κ, transition/transversion ratio; df, degrees of freedom; NA, not applicable.

for the African lake partition), and in three of the five cases, the fit of the branch-site model was worse than that of the random sites model (M2a/M2a_rel), which does not allow among-lineage variation (table 3, supplementary table S9, Supplementary Material online). These results highlight the weaknesses of the branch-site model when there is positive selection in the background and how a suboptimal model can be chosen by this method.

Positively Selected Sites

To examine the potential aspects of rhodopsin function that were under positive selection in cichlids, positively selected sites were identified from the random sites analyses of the African river, African lake, and Neotropical data subsets. We found 18 positively selected sites in Neotropical cichlids, 5 in African river cichlids, and 26 in African lake cichlids, identified with a posterior probability of at least 80% (or P value < 0.1) in the M8 (PAML) and FUBAR (HYPHY) models (table 4, supplementary table S10, Supplementary Material online). Surprisingly, the positively selected sites in African lake and Neotropical cichlids were largely nonoverlapping with only five of the sites shared between the two groups (table 4, supplementary table S10, Supplementary Material online). The positively selected sites in African river cichlids consisted of three sites shared with African lake cichlids and two sites unique to African river cichlids. It is intriguing that some of the sites under positive selection appear to be similar between African river and African lake species but that the strength of selection is similar between African river and Neotropical cichlids, suggesting that perhaps selective pressures are influenced more strongly by environmental factors, whereas the particular sites under positive selection are influenced by the evolutionary history of the protein. However, our sample size for African river cichlids is quite small, and this reduces both our power to detect positively selected sites and our reliability in the sites detected. Inclusion of African river sites detected by the less stringent REL model reveals two sites shared with Neotropical cichlids to the exclusion of African lake cichlids, and thus it is likely that additional sampling will reveal a more complex pattern of sites under positive selection.

Other models implemented in PAML (M2, M8) and HYPHY (REL, FUBAR, FEL) showed similar results for the identification of positively selected sites (supplementary tables S11–S13 and fig. S3, Supplementary Material online). Overall, the REL model found the largest number of positively selected sites, with the highest posterior probabilities, followed by FUBAR, M8, M2a, and then FEL (supplementary tables S11–S13 and fig. S3, Supplementary Material online). Some sites were identified by only the PAML models or the HYPHY models (table 4, supplementary tables S11–S13, Supplementary Material online). Sites identified as positively selected by HYPHY (FUBAR), but not PAML (M8), were assigned to the positively selected site class in M8 but with low posterior probability. Conversely, where M8 found sites that FUBAR did not, the ω levels estimated by FUBAR tended to

Table 4. List of Positively Selected Sites Found by the M8 BEB (M) and FUBAR (F) Analyses with a Posterior Probability of at Least 80%.

Codon	Location	Neo	AfR	AfL	Possible Effect on Rhodopsin Function	References
32	N	–	–	F		
33	N	MF	–	–		
37	TM1	–	–	MF*		
41	TM1	–	–	MF*	Near retinal channel A	Hildebrand et al. (2009)
42	TM1	–	–	MF	Near retinal channel A	Hildebrand et al. (2009)
48	TM1	–	–	F		
49	TM1	–	MF	–		
50	TM1	F	–	–		
75	TM2	F	–	–		
83	TM2	–	–	F	Spectral tuning, meta-II equilibrium, dark adaptation	Breikers et al. (2001), Sugawara et al. (2005, 2010)
95	TM2	–	–	MF*		
104	E1	–	–	F		
124	TM3	MF	–	–	Spectral tuning	Garriga et al. (1996), Lin et al. (1998), Hunt et al. (2001)
133	TM3	–	MF	MF*		
156	TM4	MF*	–		Dimerization interface	Guo et al. (2005), Fotiadis et al. (2006)
162	TM4	–	MF	MF*	Dimerization interface	Guo et al. (2005), Fotiadis et al. (2006)
163	TM4	–	–	M	Dimerization interface	Guo et al. (2005), Fotiadis et al. (2006)
165	TM4	–	–	MF*	Dimerization interface	Guo et al. (2005), Fotiadis et al. (2006)
166	TM4	–	–	MF*	Dimerization interface	Guo et al. (2005), Fotiadis et al. (2006)
169	TM4	MF*	–	MF*	Dimerization interface	Guo et al. (2005), Fotiadis et al. (2006)
172	TM4	MF	–	–	Dimerization interface	Guo et al. (2005), Fotiadis et al. (2006)
173	TM4	MF*	–	–	Dimerization interface	Guo et al. (2005), Fotiadis et al. (2006)
213	TM5	–	–	M*	Dimerization interface, near retinal channel B	Guo et al. (2005), Fotiadis et al. (2006), Hildebrand et al. (2009)
217	TM5	MF*	–	MF*	Dimerization interface	Guo et al. (2005), Fotiadis et al. (2006)
218	TM5	F	–	MF*	Dimerization interface	Guo et al. (2005), Fotiadis et al. (2006)
259	TM6	–	–	M		
260	TM6	–	F	–		
263	TM6	–	–	MF*		
270	TM6	MF*	–	M	Near retinal channel B	Hildebrand et al. (2009)
274	TM6	MF*	–	–	Near retinal channel B	Hildebrand et al. (2009)
281	E2	MF*	–	–	3D structure	Anukanth and Khorana (1994)
286	E2	MF*	–	–	Near retinal channel A	Hildebrand et al. (2009)
287	TM7	MF*	–	–	Near retinal channel A	Hildebrand et al. (2009)
290	TM7	–	–	F		
292	TM7	–	–	MF	Spectral tuning, dark adaptation	Sugawara et al. (2005)
297	TM7	–	MF	M*	Near retinal binding site 296	
298	TM7	–	–	MF*	Near retinal binding site 296	
299	TM7	MF	–	MF*	Spectral tuning	Fasick and Robinson (1998), Bischoff et al. (2012)

NOTE.–E, extracellular loop; N, N-terminus; TM, transmembrane domain. Asterisk denotes entries with a posterior probability of at least 95%. Codon site numbers follow bovine rhodopsin. Amino acid identities at the positively selected sites are shown in supplementary table S10, Supplementary Material online. Results from the full sets of positively selected site analyses are shown in supplementary tables S11–S12, Supplementary Material online.

be slightly greater than one but with elevated d_S values. However, elevated d_S driving an increase in ω did not appear to be the sole reason for detection of positive selection by M8, because these same sites, with one exception, were found by REL, which allows for variation in d_S.

Here, we map the sites found by the M8 and FUBAR analyses on African lake and Neotropical cichlids onto the crystal structure of both the dark-state and meta-II conformations of rhodopsin (Palczewski et al. 2000; Choe et al. 2011) and discuss their relevance in light of published results of site-directed mutagenesis studies of rhodopsin function.

We found that the positively selected sites map to regions in rhodopsin that are associated with both spectral and non-spectral properties and fall into several functional categories including spectral tuning, retinal entry/exit, and receptor dimerization (table 4, supplementary table S10, Supplementary Material online). Although the sites were largely nonoverlapping between African lake and Neotropical cichlids, both clades had positively selected sites in each of the functional categories.

Several of the sites found to be positively selected in African lake and Neotropical cichlids have been identified

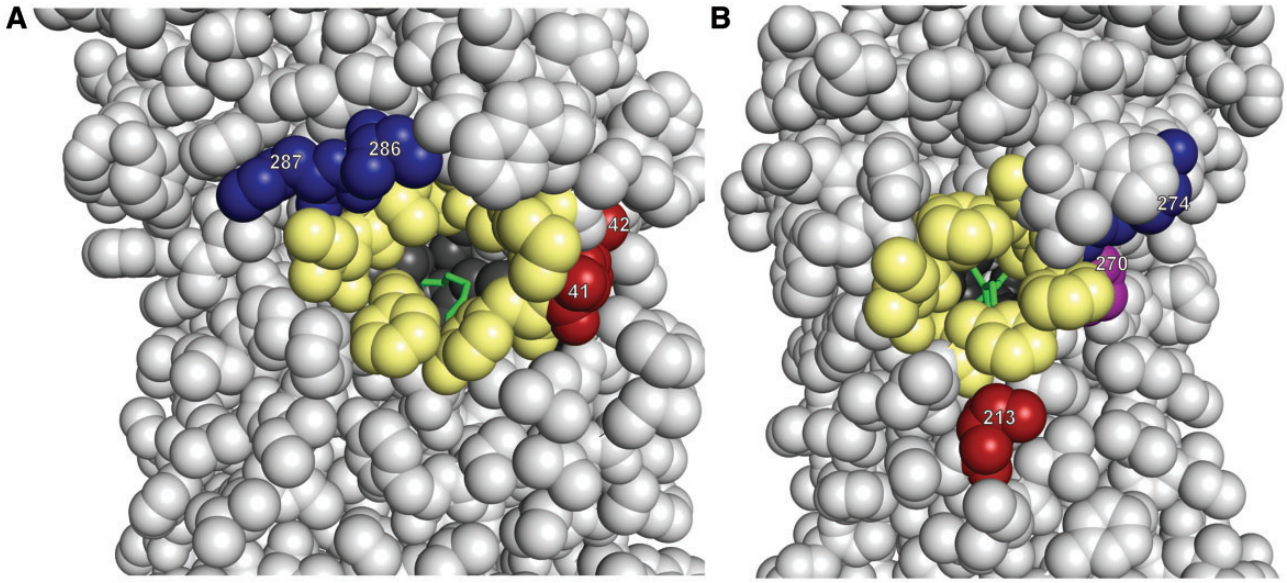

FIG. 2. Openings to retinal binding pocket in the active meta-II conformation of rhodopsin. (*A*) The opening between helices I and VII and (*B*) the opening between helices V and VI. Residues around the opening are highlighted. Residues 41, 42, and 213 are positively selected sites near the openings in African lake cichlids, and residues 274, 286, and 287 are positively selected sites near the openings in Neotropical cichlids. Residue 270 is positively selected in both African lake and Neotropical cichlids. The chromophore is shown in the center of the opening as a stick model. Sites are mapped onto PDB ID 3PXO.

by site-directed mutagenesis studies to affect spectral tuning in cichlids and other taxa. Two sites were positively selected only in African lake cichlids (83, 292), one in only Neotropical cichlids (124), and one in both groups (299). At site 292, mutation from Ala to Ser has been shown to cause a >10 nm blue shift in African lake cichlids, including a similar red shift for the reverse mutation (Sugawara et al. 2005). Mutation studies in bovine rhodopsin, which also show a large blue shift for A292S, demonstrate the consistent effects of this mutation (Fasick and Robinson 1998; Lin et al. 1998; Janz and Farrens 2001). In our data set, four African lake species were found with S292, whereas other African lake and all Neotropical and African river cichlids had A292. Mutagenesis analyses have shown that D83N causes a 2–8 nm blue shift in African lake cichlids (Sugawara et al. 2005) and a small, variable blue shift in bovine rhodopsin (Janssen et al. 1990; Nathans 1990; Fahmy et al. 1993; Weitz and Nathans 1993; Fasick and Robinson 1998; Breikers et al. 2001; Nagata et al. 2002). At site 83, African lake cichlids primarily have Asp, but three species were found with Asn. In Neotropical cichlids, however, all species except for the basal *Retroculus xinguensis* had Asn, and all African river cichlids have Asp. Although the N83 mutation is sometimes found in conjunction with S292 (Sugawara et al. 2010), this was not the case here. Site 124 was positively selected in Neotropical cichlids and varied between Ser, Ala, and Gly. Variation between these three residues has been proposed to be associated with changes in spectral tuning of rhodopsin through analysis of natural variation in deep-sea fish (Hunt et al. 2001). Although site-directed mutagenesis has not been performed between Ser, Ala, and Gly, the mutations A124R and A124T in bovine rhodopsin did result in small blue shifts (Garriga et al. 1996; Lin et al. 1998) confirming that changes in polarity at this site can affect spectral tuning. In African lake and river cichlids, this site is conserved as Gly. Both African lake and Neotropical cichlids were positively selected at site 299 and both with variation between Ala and Ser. Mutagenesis studies in bovine rhodopsin have found that A299S resulted in a 2 nm red shift (Fasick and Robinson 1998), and studies of natural variation in cetacean rhodopsins have associated this same mutation with a 5 nm red shift (Bischoff et al. 2012). African river cichlids are conserved at site 299 with Ser.

A number of sites adjacent to the entry/exit channels for retinal into the binding pocket were found to be under positive selection as well (fig. 2). The structures of the activated opsin (Park et al. 2008), G protein-interacting (Scheerer et al. 2008), and meta-II state (Choe et al. 2011) have revealed a channel through the protein that is thought to provide access to the chromophore binding pocket, with openings into the lipid bilayer between helices I and VII and between helices V and VI (Hildebrand et al. 2009). Current theories suggest that retinal traverses through this channel unidirectionally (Schadel et al. 2003; Hildebrand et al. 2009), with helices V and VI providing the opening for retinal release (Wang and Duan 2011). Positively selected sites adjacent to the entry/exit channels show variation in size and polarity in both African lake and Neotropical cichlids (table 4, supplementary table S10, Supplementary Material online). However, this variation is found at different sites in the two groups. As illustrated in figure 2, the positively selected sites 41 and 42 in African lake cichlids and 286 and 287 in Neotropical cichlids are adjacent to the opening between helices I and VII (retinal channel A), and sites 213 in African lake cichlids, 274 in Neotropical cichlids, and 270 in both groups are adjacent to the opening between helices V and VI (retinal channel B). Variation in

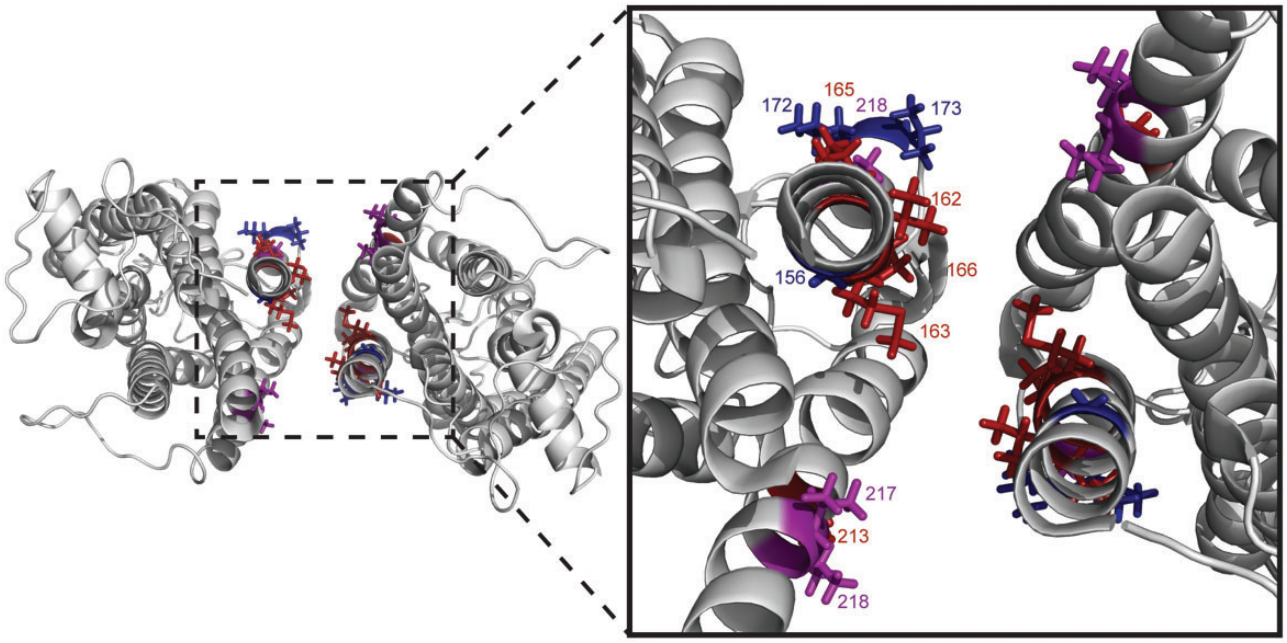

Fig. 3. Interface between rhodopsin molecules in a theoretical dimer model (Fotiadis et al. 2006). Residues 163, 163, 165, 166, and 213 are positively selected sites on helices IV and V in African lake cichlids; residues 156, 172, and 173 are positively selected sites on helix IV in Neotropical cichlids; and residues 217 and 218 are positively selected sites on helix V in both African lake and Neotropical cichlids. Sites are mapped onto the theoretical model of the rhodopsin dimer, PDB ID 1N3M.

size and polarity at these sites could affect the size and shape of the channels and thus the ability for retinal to enter and exit the binding pocket through steric effects or changes in polarity (Chen et al. 2012). The side chain of residue 286, in particular, points directly into channel A, and shows Val/Ile variation in Neotropical cichlids (fig. 2; supplementary table S10, Supplementary Material online). The additional methyl group in Ile could potentially hinder the passage of retinal and may be a good target for future mutagenesis studies aiming to determine the direction of retinal passage. The other positively selected sites are directly adjacent to the residues lining the openings (fig. 2), and therefore, variation at these sites could also affect retinal uptake or release.

The final functional category under positive selection pertains to rhodopsin sites thought to be involved in dimerization (fig. 3). Rhodopsin is known to form dimers and higher order oligomeric interactions in vivo (Fotiadis et al. 2003; Jastrzebska et al. 2006). Modeling studies suggest that the closest contact between monomers occurs between transmembrane helices IV and V (Fotiadis et al. 2006); though a recent study has found that there may also be important intermolecular contacts between helices I and VIII mediated by C316 in native disk membranes (Knepp et al. 2012). Positively selected sites found on the modeled dimerization interface are illustrated in figure 3 and fall in two distinct clusters in African and Neotropical cichlids, respectively, with a third small cluster that is positively selected in both. A single site on the dimerization interface was positively selected in African river cichlids. These sites either vary between smaller and larger hydrophobic residues or between hydrophobic and polar residues and are highly conserved in the group in which they are not positively selected (supplementary table S10, Supplementary Material online). Site 156, in particular, shows substantial size variation between Gly and Phe. Although the positively selected sites were not identified as being directly involved in binding (Guo et al. 2005; Fotiadis et al. 2006), the precise nature of the dimeric interface is not known (Morris et al. 2009; Lohse 2010), and the sites in each group are tightly clustered together facing the opposing rhodopsin partner (fig. 3), which is certainly suggestive of a function in dimerization. It is possible that these substitutions affect the affinity between members of a rhodopsin dimer or the density of rhodopsin packing, although it is not clear what effect this would have for the visual sensitivity or acuity of the animal.

Discussion

Through targeted rhodopsin sequencing and molecular evolutionary analyses, we have found evidence for positive selection in both Neotropical and African cichlid rhodopsins. We have also found evidence which suggests that divergent selective pressures on this gene may have been caused by ecological differences in lake and river habitats. We found that clade models were sufficient to detect divergent selective pressures among cichlid groups, but that the commonly used branch-site models did not perform as well, due to violation of the key assumption of no positive selection in the background. Positively selected sites in rhodopsin identified by random sites analyses were associated with three main functions: spectral tuning, chromophore entry and exit, and rhodopsin dimerization. Surprisingly, these sites were largely nonoverlapping between African lake and Neotropical

cichlids, but both groups had positively selected sites in each functional category. Here, we discuss our results in light of the evolution and ecology of cichlid fishes and also assess the methodological implications of our findings.

Positive selection on rhodopsin in Neotropical cichlids was expected given the wide variety of niches and environments these fish occupy, but the strength of the evidence is remarkable given that positive selection in African rift lake opsin genes is closely linked to both sexual dimorphism (Terai et al. 2006; Miyagi et al. 2012) and very recent adaptive radiation (Spady et al. 2005), neither of which is the case in Neotropical cichlids (López-Fernández et al. 2010, 2013). This suggests that ecological differences over long time scales are sufficient to drive detectable positive selection in the rhodopsin of cichlid fishes and indicates that visual system evolution may be important for cichlid adaptive diversification not only in the African rift lakes.

The African rift lake cichlids encompass most of the family's diversity and are under the highest levels of positive selection on rhodopsin. The African and Neotropical riverine cichlids, although less diverse than the African rift lake cichlids, all showed evidence of positive selection, albeit at lower levels than their African lake counterparts. The results of the clade analyses support the hypothesis that differences in selective pressures among clades may be due to ecological differences between lake and river habitats. Lake environments are highly variable in terms of spectral quality but are also strongly partitioned with respect to habitat. As African cichlids invaded the rift lakes and diversified, specializing in various micro-habitats (reviewed in Danley and Kocher 2001), varying light levels within these habitats may have provided the opportunity for strong positive selection on visual pigments to adapt to new environments. This divergent selection on the visual system, and the associated divergence in male coloration within the rift lakes, likely promoted rapid and extensive speciation through sensory drive (Seehausen et al. 2008). Until recently, low rates of diversification in African riverine cichlids have been hypothesized to be due to temporal instability of African riverine habitats and interpreted as evidence of reduced adaptive speciation, favoring the idea that African riverine cichlid diversification resulted predominantly from vicariance events (Joyce et al. 2005; Katongo et al. 2005, 2007). Recent work, however, has started providing strong evidence for adaptive diversification in riverine cichlids both in African (Koblmüller et al. 2008; Schwarzer et al. 2011, 2012) and Neotropical cichlids (Piálek et al. 2012; López-Fernández et al. 2013; Arbour and López-Fernández 2013). Both Neotropical and African riverine environments contain markedly different water types (e.g., Lamboj 2004; Albert and Reis 2011) with spectral quality characterized by black (transparent but stained with tannins), white (turbid, with high amounts of dissolved solids), or clear (transparent, with low amounts of dissolved solids) water that create differential degrees of light attenuation among habitats (Sioli 1984; Cooke et al. 2012). These differences may have influenced selection on the dim-light visual pigment rhodopsin because the amount and quality of available light can differ dramatically across habitats. However, future studies that incorporate additional sampling of African riverine cichlids, as well as Central American crater lake cichlids, may reveal more complex patterns, which would alter our current ecological interpretations.

Coincident with divergent selective pressure between lake and riverine cichlids, there is also divergence in the identities of sites under positive selection between African lake cichlids and Neotropical riverine cichlids. Although the positively selected sites were largely nonoverlapping between the African lake and Neotropical cichlid groups, they did fall into the same three functional categories. Thus, although there is divergence between the strength of selection between the two groups, it is unclear if there is a divergence or convergence of function in the positively selected sites. Different substitutions at the molecular level may yield similar adaptive functions as a response to environmental pressures. Amino acid substitutions can produce general, nonlocal effects on rhodopsin function (Piechnick et al. 2012), and substitutions at different sites can have convergent effects on function (e.g., Hunt et al. 2001; Takenaka and Yokoyama 2007). Interestingly, African river species shared three positively selected sites with African lake species, and none with Neotropical river species, although African river species suffer from a small sample size, and only five positively selected sites were detected. This suggests that, at least in this case, the strength of the selection may be influenced more strongly by environmental or ecological factors (lake vs. river), whereas the particular sites under positive selection might be more influenced by the evolutionary history of the protein. It follows that the same functional changes may be selected for in each environment, but on different sites as dictated by the protein background. This opens an interesting avenue for future research where the functional changes associated with positive selection are experimentally tested.

Rhodopsin spectral tuning sites have frequently been found to be a target of selection in aquatic organisms (e.g., Fasick and Robinson 2000; Hunt et al. 2001; Larmuseau et al. 2010; Sivasundar and Palumbi 2010). In our cichlid data set, a number of positively selected sites have been previously found to be important for spectral tuning in site-directed mutagenesis studies. Site 292, which was positively selected in African lake cichlids, likely plays a role in adaptation to blue-shifted waters in deep-dwelling African lake species (Sugawara et al. 2005, 2010; Nagai et al. 2011), as well as in deep-sea fish (Hunt et al. 2001). Interestingly, the D83N substitution has also been noted frequently in deep-water dwelling organisms (Hunt et al. 1996, 2001; Hope et al. 1997; Fasick and Robinson 2000), including cichlids from the deepest regions of the African rift lakes (Sugawara et al. 2005, 2010). Substitutions at both sites, D83N and A292S, have recently been shown to also increase the speed of production of rhodopsin's active meta-II state upon photoactivation in deep water cichlids (Sugawara et al. 2010), as well as affecting the rate of retinal release (Bickelmann et al. 2012), which suggests a potential adaptive role in light-limiting environments by increasing the stability of meta-II.

In addition to substitutions that shift the λ_{max} of rhodopsin, we identified positively selected sites that may be

influencing nonspectral properties of rhodopsin such as the passage of retinal through the protein and the dimerization interface between rhodopsin monomers. Variation at positively selected sites lining the retinal channels in cichlids may alter either the uptake of 11-*cis*-retinal during regeneration or the release of all-*trans*-retinal during meta-II decay (Chen et al. 2012; Piechnick et al. 2012). Mutations at highly conserved sites around both channel openings in bovine rhodopsin have been shown to dramatically alter retinal uptake and release (Piechnick et al. 2012). This may affect the rate of recovery in photoreceptors, and potentially be an adaptation for increased visual sensitivity or dark adaptation, particularly at low bleaching levels (Ala-Laurila et al. 2006). Similarly, sites that affect the rhodopsin dimerization interface are likely to affect visual sensitivity, as dimers are thought to enhance transducin activation (Fotiadis et al. 2006; Jastrzebska et al. 2013). Studies in other GPCRs have shown that heterodimerization can affect various aspects of protein function, such as ligand binding (reviewed in Lohse 2010). Altogether, our findings suggest that visual protein evolution may have played an important role in the diversification of riverine cichlids in the Neotropics and highlight the importance of investigating natural sequence variation in organisms from varied environments in studies of protein evolution.

Methodological Implications

In this study, we combined several different codon-based likelihood methods to explore biogeographic and ecological hypotheses of rhodopsin evolution in cichlid fishes. Random sites models were used to determine selective pressures acting on each cichlid group and to identify individual sites under positive selection, whereas branch-site and clade models were used to localize the effects of positive selection, either to the branch leading to a clade or to the clade itself, and to test for divergent selective pressures. We argue that the combination of these methods uncovers patterns of variation not apparent when different models are used in isolation and that clade models generally performed better for our data set than branch-site models.

The branch-site model was originally designed to test for an episode of positive selection along particular branches in an otherwise conservatively evolving background (Yang et al. 2005; Zhang et al. 2005). These tests assume that there is a category of sites that switches from neutral or purifying selection to positive selection on a specific branch. Although the use of the branch-site model for this purpose has been shown to be statistically robust (Zhang et al. 2005; Yang and dos Reis 2011), the use of this model to test for divergent selection between clades is becoming increasingly popular (Spady et al. 2005; Ramm et al. 2008; Yoshida et al. 2011; Smith et al. 2012; Badouin et al. 2013; Veilleux et al. 2013). However, this method was not designed for this purpose and can produce false positives when its assumptions are violated (Suzuki 2008). For example, if the assumption of the branch-site test that there is no positive selection in the background is violated (as was the case for the data presented here), the alternative model allowing positive selection in the foreground may fit the data better, even if there are positively selected sites throughout the phylogeny, leading to false-positive results (Zhang et al. 2005; Suzuki 2008; Yoshida et al. 2011). Clade models, which were designed to detect sites that vary in the strength and form of selection among clades (Bielawski and Yang 2004; Weadick and Chang 2012a), do not suffer from these limitations, and we found that they performed better in our data set, which showed clear evidence of positive selection in the background. A potential limitation of CmC is that it tests for divergent selective pressures rather than positive selection. This issue can be overcome by constraining the clade model's divergent site class to equal one and comparing the results with an unconstrained CmC using an LRT as suggested by Chang et al. (2012), thus making it an explicit test of positive selection in the context of divergent selection between clades. On the basis of our results, we recommended the use of CmC over the branch-site model in data sets where there is likely to be variation in the strength of selection in the background.

CmC is a powerful method for detecting divergent selective pressures among clades. With the addition of multiclade models (Yoshida et al. 2011) and the approach of comparing successive clade partitions with LRTs introduced here, it is possible to test various evolutionary and ecological hypotheses, explore different models, and determine which best fits the data. Here, we used a priori knowledge of cichlid evolution and ecology to formulate hypotheses of divergent selective pressures in cichlid rhodopsin based on their ecology and biogeographic history, which were then tested using a combination of random sites, branch-site, and clade models of evolution. Using this approach, we found that a clade model incorporating divergent selective pressures resulting from ecological differences in lake and river habitats best fit our cichlid rhodopsin data set. In general, although the clade models were best suited for testing for divergent selective pressures among clades, it was useful to combine them with random sites models to determine if the positively selected sites were different in the lake and river habitats, as well as with structure/function analysis of the positively selected sites to infer whether their effects on function might also differ between clades.

Conclusions

African rift lake cichlids have emerged as a model system for the study of visual ecology and speciation by sensory drive due to their recent and extensive diversification into varied lake environments (Seehausen 2006; Terai et al. 2006; Seehausen et al. 2008; Miyagi et al. 2012). It has been hypothesized that niche partitioning along light gradients in lake habitats coupled with sexual selection for male nuptial coloration has resulted in positive selection in African lake cichlid visual pigments and contributed to their extensive diversification (Seehausen et al. 2008). Here, we have shown that Neotropical cichlids, which mainly inhabit river environments, also show strong positive selection on the dim-light visual pigment, rhodopsin. Furthermore, we have shown that selective pressure on rhodopsin in lake and river habitats is divergent both in terms of the strength of selection and the

individual sites under selection. It is important to note that divergent sites can affect similar aspects of function, based on current understanding of rhodopsin structure and function; this would certainly be an interesting avenue of future experimental research. Our analyses were greatly strengthened by an understanding of the ecological and biogeographic processes shaping the evolution of cichlids, which allowed us to formulate and test a priori hypotheses in the context of sites and lineage-specific likelihood models of evolution. Our approach combined random sites with clade models of evolution, using LRTs between successive partitions; this allowed us to determine that a model incorporating divergent selection between lake and river habitats best fit our data and to identify individual sites under positive selection in different ecological groups. The branch-site test, which is becoming increasingly popular for examining divergent selective pressures between clades, was found to perform poorly in this context. Our study demonstrates the advantages in combining different methods to investigate molecular evolution in biological systems, particularly in genes strongly tied to ecology, and emphasizes the importance of studies incorporating natural sequence variation in organisms from varying environments.

Materials and Methods

Opsin Sequencing and Data Set Preparation

Rhodopsin was amplified from tissue samples (muscle or fin) obtained from the Ichthyology collection at the Royal Ontario Museum for one to three individuals of 32 cichlid species. This included Neotropical riverine cichlids with at least one species from each genus in the tribe Geophagini, except *Acarichthys*, and three species basal to Geophagini (*R. xinguensis*, *Cichla temensis*, and *Chaetobranchus flavescens*), as well as the basal African riverine cichlids *Heterochromis multidens*, *Hemichromis fasciatus*, and *Chromidotilapia guntheri* (López-Fernández et al. 2010). DNA was extracted using standard phenol/chloroform extraction protocols, and an 859 base pair fragment of rhodopsin was amplified using the primers 193F (CNTATGAATAYCCTCAGTACTA) and Rh1039R (CCRCAGCACARCGTGGTGATCA) (Chen et al. 2003). Polymerase chain reaction was performed using standard cycling conditions. Fragments were visualized on agarose gels and extracted using a QIAquick Gel Extraction Kit (QIAGEN). Fragments were cloned into the pJET 1.2 cloning vector (Fermentas), cultured in liquid media, and isolated using GeteJET Plasmid Miniprep Kit (Fermentas). Three to four clones were sequenced per individual. Sequencing was performed in the forward and reverse directions using a 3730 Analyzer (Applied biosystems). Sequences were assembled, manually trimmed, and edited in Sequencher 5.0.4.9 (Genecodes) to produce a consensus sequence for each species. Additional cichlid rhodopsin sequences were obtained from Genbank and include all RH1 sequences available from African riverine cichlids (six additional species) as well as representatives from Lakes Malawi, Tanganyika, and Victoria (32 species). Sequences were also obtained from eight species found within the same major percomorph clade as cichlids (Ovalentaria) to act as outgroups for phylogenetic and molecular evolutionary analyses (Near et al. 2013). Sequences were aligned using PRANK (Löytynoja and Goldman 2005). Prior to molecular evolutionary analyses, the alignment was pruned to the length of the average Neotropical rhodopsin sequence. Species list and accession numbers for all sequences used in the study are provided in supplementary table S1, Supplementary Material online.

Phylogenetic Analyses

Rhodopsin gene trees were estimated in MrBayes 3 (Ronquist and Huelsenbeck 2003) using reversible jump Markov chain Monte Carlo with a gamma rate parameter (nst = mixed, rates = gamma), which explores the parameter space for the nucleotide model and the phylogenetic tree simultaneously, and by ML using PhyML 3 (Guindon et al. 2010) under the GTR + G + I model with a BioNJ starting tree, the best of NNI and SPR tree improvement, and aLRT SH-like branch support (Anisimova and Gascuel 2006). The Bayesian analysis was run for 5 million generations with a 25% burn-in. Convergence was confirmed by checking that the standard deviations of split frequencies approached zero and that there was no obvious trend in the log likelihood plot.

Molecular Evolutionary Analyses

To estimate the strength and form of selection acting on rhodopsin, the alignment, along with the Bayesian gene tree, was analyzed with the codeml program from the PAML 4 software package (Yang 2007) using the random sites models (M0, M1a, M2a, M3, M7, M8a, and M8), branch-site model, and CmC. The random sites and CmC analyses were repeated with the ML gene tree to ensure that minor changes in topology did not significantly alter the results. Because PAML does not incorporate rate variation in synonymous sites (d_S), we also analyzed the data using the HYPHY (Pond et al. 2005) REL, FEL, FUBAR, and PARRIS models (Kosakovsky Pond and Frost 2005; Scheffler et al. 2006; Murrell et al. 2013) implemented on the Datamonkey webserver (Delport et al. 2010), which are similar to the PAML random sites models, but allow for an independently estimated d_S. Three different subsets of the RH1 data set were analyzed with the random sites models of PAML and HYPHY to assess differences in selective pressure among the various groups. These subdivided data sets contained 1) only Neotropical cichlid RH1 sequences, 2) African riverine cichlid sequences, and 3) African lake cichlid sequences.

Comparisons between the PAML random sites models were used to test for variation in ω (M3 vs. M0) and for the presence of a positively selected class of sites (M2a vs. M1a and M8 vs. M7 and M8a). All analyses were run starting with the branch lengths estimated by MrBayes (or PhyML, as appropriate) for the complete RH1 gene tree and repeated at least four times with varying initial starting points of κ (transition to transversion rate ratio) and ω to avoid potential local optima. The model pairs were compared using an LRT with a χ^2 distribution. Sites under positive selection in the M2a and M8 models were identified by the Bayes' Empirical Bayes

(BEB) analysis implemented in PAML (Yang et al. 2005). These codon-based likelihood models are designed to detect episodic or diversifying selection but may be inadequate to detect directional selection, especially when amino acids changes associated with such selection have only occurred once. Although the PAML models (and codon-based likelihood models in general) have recently come under statistical criticisms (Friedman and Hughes 2007; Suzuki 2008; Nozawa et al. 2009a, 2009b), these criticisms have largely been refuted and the models shown to have robust statistical properties (Yang et al. 2009; Yang and dos Reis 2011; Weadick and Chang 2012a; Zhai et al. 2012; Gharib and Robinson-Rechavi 2013).

The site models of HYPHY were also implemented for each subset of the RH1 data (using the Bayesian topology, except for PARRIS which was performed using both topologies) to ensure that allowing for variation in d_S did not alter the conclusions of the analyses. The PARRIS model performs an LRT between a null model with three synonymous rate classes and two ω rate classes (constrained to equal 0 and 1, respectively) with an alternate model that adds an additional ω class that is free to vary (Scheffler et al. 2006). This provides an explicit test of positive selection under variable synonymous rates. The REL, FEL, and FUBAR models do not implement LRTs between a null and alternate model but instead estimate the ω value for each site in a comparable way to the BEB analysis of PAML (Kosakovsky Pond and Frost 2005; Murrell et al. 2013). Comparisons between the BEB analysis of the M8 model and the REL, FEL, and FUBAR models were performed by graphing the estimated ω value for each site and by comparing the sites found to be under significant positive selection.

The branch-site model (Zhang et al. 2005) and CmC (Bielawski and Yang 2004) were used to test for positive selection along particular branches using the full RH1 alignment and gene tree. These models allow ω to vary among sites and between "foreground" and background branches, or clades, specified by the user, based on a priori hypotheses of where adaptive evolution may have occurred. The branch-site model has four site classes: 0) $0 < \omega_0 < 1$ for all branches; 1) $\omega_1 = 1$ for all branches, 2a) $\omega_{2a} = \omega_{2b} \geq 1$ in the foreground and $0 < \omega_{2a} = \omega_0 < 1$ in the background, and 2 b) $\omega_{2b} = \omega_{2a} \geq 1$ in the foreground and $\omega_{2b} = \omega_1 = 1$ in the background. CmC assumes that some sites evolve conservatively across the phylogeny (two classes of sites where $0 < \omega_0 < 1$ and $\omega_1 = 1$), while a class of sites is free to evolve differently among two or more partitions (e.g., $\omega_{D1} > 0$ and $\omega_{D1} \neq \omega_{D2} > 0$), which can be branches, clades, or a mix of both. The main differences between the two models is that CmC does not assume positive selection in the divergent site class in the foreground and does not constrain that background to be under only negative and purifying selection. Instead, CmC allows separate, unconstrained estimates of ω for the third (divergent) site class for each partition. These models were used to determine whether significant differences in selection among clades highlighted by the random sites models are driven by a burst of selection in the lineage leading to each of the main clades. Analyses were conducted with the branch leading to all cichlids, African cichlids, and Neotropical cichlids, designated as the foreground. The branch-site models were compared with a null model where ω_2 is constrained to be equal to one. The null model used for CmC was M2a_rel, which does not allow divergence of ω in the third site class (but the ω value for this site class is still unconstrained; Weadick and Chang 2012a). The LRT using this model has a significantly lower false-positive rate than previous tests that compared the divergent model to the M1a model, which does not allow divergent selection, and also has one less site class (Weadick and Chang 2012a). This comparison can lead to false positives when three site classes are a better fit than two, even if there is no divergent selection (Weadick and Chang 2012a). To avoid local optima, each analysis was run at least four times with varying initial values of κ and ω.

CmC was also used to test for differences in selection among clades and to determine what division of clades best fit the data, using the full RH1 alignment and Bayesian and ML gene trees. A priori knowledge of cichlid historical biogeography and ecology were used to direct the tests. CmC was recently extended to allow for more than two partitions (Yoshida et al. 2011), allowing us to define clades in multiple, successive sets of partitions. The two-partition models were cichlids/outgroups, African cichlids/outgroups and other cichlids, African lake cichlids/outgroups and other cichlids, African river cichlids/outgroups and other cichlids, and Neotropical cichlids/outgroups and other cichlids. Three partition: African cichlids/Neotropical cichlids/outgroups, and lake cichlids (African lake cichlids)/river cichlids (African river cichlids and all Neotropical cichlids)/outgroups. Four partition: African lake cichlids/African riverine cichlids/Neotropical riverine cichlids/outgroups. These partitions are defined and depicted in figure 1. Throughout the article, we often refer to these partitions only by the foreground, whereas the background partition that contains outgroups and possibly other taxa is not mentioned. In all cases, the background partition is present and includes whichever clades and branches were not placed in one of the mentioned foreground partitions. CmC analyses are prone to local optima (Bielawski and Yang 2004; Weadick and Chang 2012a), so all models were run at least four times with varying initial κ and ω values to ensure convergence. The CmC analyses were compared with the M2a_rel null model as described earlier to test for the presence of divergent selection. The analyses with statistically significant LRTs were further analyzed to test if the ω value in the divergent site class was significantly different from one. This was done by constraining the ω of the divergent site class to be equal to 1 (Chang et al. 2012). Only a single partition could be tested at a time, so this was done sequentially for each partition thought to be under positive selection.

To statistically compare the results of our multipartition tests, we used the approach of comparing models with a difference of one partition with an LRT test and by AIC comparison to determine if the addition of a partition was a significantly better fit. The parameters of the M2a_rel null model and the two and higher partition CmC models are all nested (each adds an additional site class and associated free

parameter). In most cases, the partitions were also nested and so the models could be compared with LRTs. In the cases where the partitions were not also nested the LRT was not strictly applicable, and so AIC was used to compare the models. Using this approach, we were able to test different hypotheses based on ecology and historical biogeography.

Some studies have used branch-site models to highlight multiple lineages or entire clades (Spady et al. 2005; Ramm et al. 2008; Yoshida et al. 2011; Smith et al. 2012; Badouin et al. 2013; Veilleux et al. 2013), despite the fact that this method can lose power if selection pressures are different among foreground branches (Zhang et al. 2005) and can produce false-positive results if assumptions are violated (Suzuki 2008). We performed the same two-partition tests as were used with CmC using the branch-site model to compare the results between the two methods. The branch-site model cannot contain more than two partitions, so we could not replicate our multipartition tests. To compare the results of the branch-site model with CmC and to determine the overall best-fitting model, we used AIC comparisons.

We used the BEB method of PAML, and the REL, FEL, and FUBAR methods of HYPHY to determine which sites in the amino acid sequence were under positive selection in the rhodopsins of Neotropical, African river, and African lake cichlids (excluding any outgroups). Sites were deemed to be under positive selection if they had a posterior probability $\geq$80% ($P \leq 0.1$ for FEL). Sites with posterior probabilities $\geq$95% were deemed highly significant. To directly compare sites under positive selection from the two main positively selected clades that differ in their ecological habitat (Neotropical and African lake cichlids), we chose to use to the most robust methods of PAML and HYPHY, M8, and FUBAR (Yang 2007; Murrell et al. 2013). Sites identified as being under positive selection in African lake and Neotropical cichlids by these methods were mapped onto the meta-II (Choe et al. 2011), and theoretical dimer (Fotiadis et al. 2006) 3D structures of rhodopsin (PDB accession numbers 1U19, 3PXO, and 1N3M, respectively) using PyMOL v. 1.5.0.4 (DeLano 2002). Bovine rhodopsin numbering is used throughout.

Supplementary Material

Supplementary tables S1–S13 and figures S1–S3 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This work was supported by Natural Sciences and Engineering Research Council (NSERC) Discovery grants to B.S.W.C. and H.L.-F.; NSERC graduate fellowships to S.P.R., R.K.S., and F.E.H.; a Vision Science Research Fellowship to F.E.H.; an Ontario Graduate Scholarship to S.P.R.; a Human Frontiers Research Program grant to B.S.W.C.; and the Royal Ontario Museum through a Life in Crisis: Schad Gallery of Biodiversity Conservation Grant to H.L.-F.

References

Ala-Laurila P, Kolesnikov AV, Crouch E, Tsina SA, Shukolyukov VI, Govardovskii Y, Koutalos B, Wiggert B, Estevez ME, Cornwall MC. 2006. Visual cycle: Dependence of retinol production and removal on photoproduct decay and cell morphology. *J Gen Physiol.* 128: 153–169.

Albert JS, Reis RE, editors. 2011. Historical biogeography of Neotropical freshwater fishes. Berkeley (CA): University of California Press.

Anisimova M, Gascuel O. 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst Biol.* 55: 539–552.

Anisimova M, Kosiol C. 2009. Investigating protein-coding sequence evolution with probabilistic codon substitution models. *Mol Biol Evol.* 26:255–271.

Anukanth A, Khorana HG. 1994. Structure and function in rhodopsin. Requirements of a specific structure for the intradiscal domain. *J Biol Chem.* 269:19738–19744.

Arbour JH, López-Fernández H. 2013. Ecological variation in South American geophagine cichlids arose during an early burst of adaptive morphological and functional evolution. *Proc R Soc Lond Ser B Biol Sci.* 280:20130849.

Badouin H, Belkhir K, Gregson E, Galindo J, Sundström L, Martin SJ, Butlin RK, Smadja CM. 2013. Transcriptome characterisation of the ant *Formica exsecta* with new insights into the evolution of desaturase genes in social hymenoptera. *PLoS One* 8:e68200.

Bakewell MA, Shi P, Zhang J. 2007. More genes underwent positive selection in chimpanzee evolution than in human evolution. *Proc Natl Acad Sci U S A.* 104:7489–7494.

Barlow GW. 2000. The cichlid fishes: nature's grand experiment in evolution. Cambridge: Perseus Publishing.

Barluenga M, Stolting KN, Salzburger W, Muschick M, Meyer A. 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* 439:719–723.

Bickelmann C, Morrow JM, Müller J, Chang BSW. 2012. Functional characterization of the rod visual pigment of the echidna (*Tachyglossus aculeatus*), a basal mammal. *Vis Neurosci.* 29:211–217.

Bielawski JP, Yang Z. 2004. A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. *J Mol Evol.* 59:121–132.

Bischoff N, Nickle B, Cronin TW, Velasquez S, Fasick JI. 2012. Deep-sea and pelagic rod visual pigments identified in the mysticete whales. *Vis Neurosci.* 29:95–103.

Bowmaker JK. 1995. The visual pigments of fish. *Prog Retin Eye Res.* 15: 1–31.

Bowmaker JK. 2008. Evolution of vertebrate visual pigments. *Vision Res.* 48:2022–2041.

Breikers G, Bovee-Geurts PH, DeCaluwé GL, DeGrip WJ. 2001. A structural role for Asp83 in the photoactivation of rhodopsin. *Biol Chem.* 382:1263–1270.

Briscoe AD, Bybee SM, Bernard GD, Yuan F, Sison-Mangus MP, Reed RD, Warren AD, Llorente-Bousquets J, Chiao CC. 2010. Positive selection of a duplicated UV-sensitive visual pigment coincides with wing pigment evolution in heliconius butterflies. *Proc Natl Acad Sci U S A.* 107:3628–3633.

Carleton KL. 2009. Cichlid fish visual systems: mechanisms of spectral tuning. *Integr Zool.* 4:75–86.

Carleton KL, Parry JWL, Bowmaker JK, Hunt DM, Seehausen O. 2005. Colour vision and speciation in Lake Victoria cichlids of the genus *Pundamilia*. *Mol Ecol.* 14:4341–4353.

Chang BSW, Du J, Weadick CJW, Muller J, Bickelmann C, Yu DD, Morrow JM. 2012. The future of codon models in studies of molecular function: ancestral reconstruction, and clade models of functional divergence. In: Cannarozii GM, Schneider A, editors. Codon evolution: mechanisms and models. Oxford: Oxford University Press. p. 145–163.

Chen M-H, Kuemmel C, Birge RR, Knox BE. 2012. Rapid release of retinal from a cone visual pigment following photoactivation. *Biochemistry* 51:4117–4125.

Chen W-J, Bonillo C, Lecointre G. 2003. Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. *Mol Phylogenet Evol.* 26:262–288.
Choe H-W, Kim YJ, Park JH, Morizumi T, Pai EF, Krauss N, Hofmann KP, Scheerer P, Ernst OP. 2011. Crystal structure of metarhodopsin II. *Nature* 471:651–655.
Cooke GM, Chao NL, Beheregaray LB. 2012. Divergent natural selection with gene flow along major environmental gradients in Amazonia: insights from genome scans, population genetics and phylogeography of the characin fish *Triportheus albus*. *Mol Ecol.* 21:2410–2427.
Danley PD, Kocher TD. 2001. Speciation in rapidly diverging systems: lessons from Lake Malawi. *Mol Ecol.* 10:1075–1086.
DeLano WL. 2002. PyMOL. Palo Alto (CA): DeLano Scientific. [cited 2013 Sept 24]. Available from: http://www.pymol.org.
Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* 26:2455–2457.
Fahmy K, Jäger F, Beck M, Zvyaga TA, Sakmar TP, Siebert F. 1993. Protonation states of membrane-embedded carboxylic acid groups in rhodopsin and metarhodopsin II: a Fourier-transform infrared spectroscopy study of site-directed mutants. *Proc Natl Acad Sci U S A.* 90:10206–10210.
Farias IP, Ortí G, Meyer A. 2000. Total evidence: molecules, morphology, and the phylogenetics of cichlid fishes. *J Exp Zool.* 288:76–92.
Fasick JI, Robinson PR. 1998. Mechanism of spectral tuning in the dolphin visual pigments. *Biochemistry* 37:433–438.
Fasick JI, Robinson PR. 2000. Spectral-tuning mechanisms of marine mammal rhodopsins and correlations with foraging depth. *Vis Neurosci.* 17:781–788.
Fay JC, Wu C-I. 2003. Sequence divergence, functional constraint, and selection in protein evolution. *Annu Rev Genom Hum Genet.* 4: 213–235.
Fotiadis D, Jastrzebska B, Philippsen A, Müller DJ, Palczewski K, Engel A. 2006. Structure of the rhodopsin dimer: a working model for G protein-coupled receptors. *Curr Opin Scruct Biol.* 16:252–259.
Fotiadis D, Liang Y, Filipek S, Saperstein DA, Engel A, Palczewski K. 2003. Atomic-force microscopy: rhodopsin dimers in native disc membranes. *Nature* 421:127–128.
Friedman R, Hughes AL. 2007. Likelihood-ratio tests for positive selection of human and mouse duplicate genes reveal nonconservative and anomalous properties of widely used methods. *Mol Phyl Evol.* 42: 388–393.
Garriga P, Liu X, Khorana HG. 1996. Structure and function in rhodopsin: correct folding and misfolding in point mutants at and in proximity to the site of the retinitis pigmentosa mutation Leu-125 –> Arg in the transmembrane helix C. *Proc Natl Acad Sci U S A.* 93:4560–4564.
Genner M, Turner G. 2012. Ancient hybridization and phenotypic novelty within Lake Malawi's cichlid fish radiation. *Mol Biol Evol.* 29: 195–206.
Gharib WH, Robinson-Rechavi M. 2013. The branch-site test of positive selection is surprisingly robust but lacks power under synonymous substitution saturation and variation in GC. *Mol Biol Evol.* 30: 1675–1686.
Gozem S, Schapiro I, Ferré N, Olivucci M. 2012. The molecular mechanism of thermal noise in rod photoreceptors. *Science* 337: 1225–1228.
Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 59:307–321.
Guo W, Shi L, Filizola M, Weinstein H, Javitch JA. 2005. Crosstalk in G protein-coupled receptors: changes at the transmembrane homodimer interface determine activation. *Proc Natl Acad Sci U S A.* 102: 17495–17500.
Hildebrand PW, Scheerer P, Park JH, Choe H-W, Piechnick R, Ernst OP, Hofmann KP, Heck M. 2009. A ligand channel through the G protein-coupled receptor opsin. *PLoS One* 4:e4382.
Hofmann KP, Scheerer P, Hildebrand PW, Choe H-W, Park JH, Heck M, Ernst OP. 2009. A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem Sci.* 34:540–552.
Hope AJ, Partridge JC, Dulai KS, Hunt DM. 1997. Mechanisms of wavelength tuning in the rod opsins of deep-sea fishes. *Proc R Soc Lond Ser B Biol Sci.* 264:155–163.
Houde AE. 1997. Sex, color, and mate choice in guppies. Princeton (NJ): Princeton University Press.
Hunt DM, Carvalho LS, Cowing JA, Davies WL. 2009. Evolution and spectral tuning of visual pigments in birds and mammals. *Phil Trans R Soc B Biol Sci.* 364:2941–2955.
Hunt DM, Dulai KS, Partridge JC, Cottrill P, Bowmaker JK. 2001. The molecular basis for spectral tuning of rod visual pigments in deep-sea fish. *J Exp Biol.* 204:3333–3344.
Hunt DM, Fitzgibbon J, Slobodyanyuk SJ, Bowmaker JK. 1996. Spectral tuning and molecular evolution of rod visual pigments in the species flock of cottoid fish in Lake Baikal. *Vision Res.* 36:1217–1224.
Janssen JJ, De Caluwé GL, de Grip WJ. 1990. Asp83, Glu113 and Glu134 are not specifically involved in Schiff base protonation or wavelength regulation in bovine rhodopsin. *FEBS Lett.* 260:113–118.
Janz JM, Farrens DL. 2001. Engineering a functional blue-wavelength-shifted rhodopsin mutant. *Biochemistry* 40:7219–7227.
Janz JM, Fay JF, Farrens DL. 2003. Stability of dark state rhodopsin is mediated by a conserved ion pair in intradiscal loop E-2. *J Biol Chem.* 278:16982–16991.
Jastrzebska B, Fotiadis D, Jang G-F, Stenkamp RE, Engel A, Palczewski K. 2006. Functional and structural characterization of rhodopsin oligomers. *J Biol Chem.* 281:11917–11922.
Jastrzebska B, Orban T, Golczak M, Engel A, Palczewski K. 2013. Asymmetry of the rhodopsin dimer in complex with transducin. *FASEB J.* 27:1572–1584.
Joyce DA, Lunt DH, Bills R, Turner GF, Katongo C, Duftner N, Sturmbauer C, Seehausen O. 2005. An extant cichlid fish radiation emerged in an extinct Pleistocene lake. *Nature* 435:90–95.
Katongo C, Koblmüller S, Duftner N, Makasa L, Sturmbauer C. 2005. Phylogeography and speciation in the *Pseudocrenilabrus philander* species complex in Zambian Rivers. *Hydrobiologia* 542:221–233.
Katongo C, Koblmüller S, Duftner N, Mumba L, Sturmbauer C. 2007. Evolutionary history and biogeographic affinities of the serranochromine cichlids in Zambian rivers. *Mol Phylogenet Evol.* 45:326–338.
Khan MMG, Rydén A-M, Chowdhury MS, Hasan MA, Kazi JU. 2011. Maximum likelihood analysis of mammalian p53 indicates the presence of positively selected sites and higher tumorigenic mutations in purifying sites. *Gene* 483:29–35.
Knepp AM, Periole X, Marrink S-J, Sakmar TP, Huber T. 2012. Rhodopsin forms a dimer with cytoplasmic helix 8 contacts in native membranes. *Biochemistry* 51:1819–1821.
Koblmüller S, Sefc KM, Duftner N, Katongo C, Tomljanovic T, Sturmbauer C. 2008. A single mitochondrial haplotype and nuclear genetic differentiation in sympatric colour morphs of a riverine cichlid fish. *J Evol Biol.* 21:362–367.
Kosakovsky Pond SL, Frost SDW. 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol.* 22:1208–1222.
Lamboj A. 2004. The cichlid fishes of western Africa. Bornheim (Germany): Birgit Schmettkamp Verlag.
Larmuseau MHD, Vancampenhout K, Raeymaekers JAM, van houdt JKJ, Volckaert FAM. 2010. Differential modes of selection on the rhodopsin gene in coastal Baltic and North Sea populations of the sand goby, *Pomatoschistus minutus*. *Mol Ecol.* 19:2256–2268.
Lin SW, Kochendoerfer GG, Carroll KS, Wang D, Mathies RA, Sakmar TP. 1998. Mechanisms of spectral tuning in blue cone visual pigments: visible and Raman spectroscopy of blue-shifted rhodopsin mutants. *J Biol Chem.* 273:24583–24591.
Lohse MJ. 2010. Dimerization in GPCR mobility and signaling. *Curr Opin Pharm.* 10:53–58.
López-Fernández H, Arbour JH, Winemiller KO, Honeycutt RL. 2013. Testing for ancient adaptive radiations in neotropical cichlid fishes. *Evolution* 67:1321–1337.

López-Fernández H, Honeycutt RL, Stiassny MLJ, Winemiller KO. 2005. Morphology, molecules, and character congruence in the phylogeny of South American geophagine cichlids (Perciformes, Labroidei). *Zool Scripta*. 34:627–651.

López-Fernández H, Winemiller KO, Honeycutt RL. 2010. Multilocus phylogeny and rapid radiations in Neotropical cichlid fishes (Perciformes: Cichlidae: Cichlinae). *Mol Phylogenet Evol.* 55: 1070–1086.

López-Fernández H, Winemiller KO, Montaña C, Honeycutt RL. 2012. Diet-morphology correlations in the radiation of South American Geophagine cichlids (Perciformes: Cichlidae: Cichlinae). *PLoS One* 7: e33997.

Löytynoja A, Goldman N. 2005. An algorithm for progressive multiple alignment of sequences with insertions. *Proc Natl Acad Sci U S A.* 102:10557–10562.

Lythgoe JN. 1979. The ecology of vision. Oxford: Oxford University Press.

Magurran AE. 2005. Evolutionary ecology: the Trinidadian guppy. Oxford: Oxford University Press.

Miyagi R, Terai Y, Aibara M, Sugawara T, Imai H, Tachida H, Mzighani SI, Okitsu T, Wada A, Okada N. 2012. Correlation between nuptial colors and visual sensitivities tuned by opsins leads to species richness in sympatric Lake Victoria cichlid fishes. *Mol Biol Evol.* 29: 3281–3296.

Morris MB, Dastmalchi S, Church WB. 2009. Rhodopsin: structure, signal transduction and oligomerisation. *Int J Biochem Cell Biol.* 41: 721–724.

Moury B, Simon V. 2011. dN/dS-based methods detect positive selection linked to trade-offs between different fitness traits in the coat protein of potato virus Y. *Mol Biol Evol.* 28:2707–2717.

Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K. 2013. FUBAR: a fast, unconstrained Bayesian approximation for inferring selection. *Mol Biol Evol.* 30:1196–1205.

Muschick M, Indermaur A, Salzburger W. 2012. Convergent evolution within an adaptive radiation of cichlid fishes. *Curr Biol.* 22: 1–7.

Nagai H, Terai Y, Sugawara T, Imai H, Nishihara H, Hori M, Okada N. 2011. Reverse evolution in RH1 for adaptation of cichlids to water depth in Lake Tanganyika. *Mol Biol Evol.* 28:1769–1776.

Nagata T, Oura T, Terakita A, Kandori H, Shichida Y. 2002. Isomer-specific interaction of the retinal chromophore with threonine-118 in rhodopsin. *J Phys Chem A.* 106:1969–1975.

Nathans J. 1990. Determinants of visual pigment absorbance: role of charged amino acids in the putative transmembrane segments. *Biochemistry* 29:937–942.

Near TJ, Dornburg A, Eytan RI, Keck BP, Smith WL, Kuhn KL, Moore JA, Price SA, Burbrink FT, Friedman M, et al. 2013. Phylogeny and tempo of diversification in the superradiation spiny-rayed fishes. *Proc Natl Acad Sci U S A.* 110:12738–12743.

Nozawa M, Suzuki Y, Nei M. 2009a. Reliabilities of identifying positive selection by the branch-site and the site-prediction methods. *Proc Natl Acad Sci U S A.* 106:6700–6705.

Nozawa M, Suzuki Y, Nei M. 2009b. Response to Yang et al.: problems with Bayesian methods of detecting positive selection at the DNA sequence level. *Proc Natl Acad Sci U S A.* 106:E96.

Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, et al. 2000. Crystal structure of rhodopsin: a G protein-coupled receptor. *Sci Signal.* 289: 739–745.

Park JH, Scheerer P, Hofmann KP, Choe H-W, Ernst OP. 2008. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* 454:183–187.

Piálek L, Říčan O, Casciotta J, Almirón A, Zrzavý J. 2012. Multilocus phylogeny of *Crenicichla* (Teleostei: Cichlidae), with biogeography of the *C. lacustris* group: species flocks as a model for sympatric speciation in rivers. *Mol Phylogenet Evol.* 62:46–61.

Piechnick R, Ritter E, Hildebrand PW, Ernst OP, Scheerer P, Hofmann KP, Heck M. 2012. Effect of channel mutations on the uptake and release of the retinal ligand in opsin. *Proc Natl Acad Sci U S A.* 109: 5247–5252.

Pond SLK, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679.

Prokhorenko VI, Nagy AM, Waschuk SA, Brown LS, Birge RR, Miller RJD. 2006. Coherent control of retinal isomerization in bacteriorhodopsin. *Science* 313:1257–1261.

Ramm SA, Oliver PL, Ponting CP, Stockley P, Emes RD. 2008. Sexual selection and the Adaptive evolution of mammalian ejaculate proteins. *Mol Biol Evol.* 25:207–219.

Reis RE, Kullander SO, Ferraris CJ. 2003. Check list of the freshwater fishes of South and Central America. Porto Alegre (Brazil): Edipucrs.

Rennison DJ, Owens GL, Taylor JS. 2012. Opsin gene duplication and divergence in ray-finned fish. *Mol Phylogenet Evol.* 62: 986–1008.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.

Schadel SA, Heck M, Maretzki D, Filipek S, Teller DC, Palczewski K, Hofmann KP. 2003. Ligand channeling within a G-protein-coupled receptor. The entry and exit of retinals in native opsin. *J Biol Chem.* 278:24896–24903.

Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauss N, Choe H-W, Hofmann KP, Ernst OP. 2008. Crystal structure of opsin in its G-protein-interacting conformation. *Nature* 455:497–502.

Scheffler K, Martin DP, Seoighe C. 2006. Robust inference of positive selection from recombining coding sequences. *Bioinformatics* 22: 2493–2499.

Schwarzer J, Misof B, Ifuta SN, Schliewen UK. 2011. Time and origin of cichlid colonization of the lower Congo rapids. *PLoS One* 6: e22380.

Schwarzer J, Misof B, Schliewen UK. 2012. Speciation within genomic networks: a case study based on *Steatocranus* cichlids of the lower Congo rapids. *J Evol Biol.* 25:138–148.

Schwarzer J, Misof B, Tautz D, Schliewen UK. 2009. The root of the East African cichlid radiations. *BMC Evol Biol.* 9:186.

Seehausen O. 2006. African cichlid fish: a model system in adaptive radiation research. *Proc R Soc Lond Ser B Biol Sci.* 273: 1987–1998.

Seehausen O, Terai Y, Magalhaes IS, Carleton KL, Mrosso HDJ, Miyagi R, van der Sluijs I, Schneider MV, Maan ME, Tachida H, et al. 2008. Speciation through sensory drive in cichlid fish. *Nature* 455: 620–626.

Shen Y-Y, Liu J, Irwin DM, Zhang Y-P. 2010. Parallel and convergent evolution of the dim-light vision gene RH1 in bats (Order: Chiroptera). *PLoS One* 5:e8838.

Sioli H, editor. 1984. The Amazon: limnology and landscape ecology of a mighty tropical river and its basin. Dordrecht (The Netherlands): Dr. W Junk.

Sivasundar A, Palumbi SR. 2010. Parallel amino acid replacements in the rhodopsins of the rockfishes (*Sebastes* spp.) associated with shifts in habitat depth. *J Evol Biol.* 23:1159–1169.

Smith LW, Chakrabarty P, Sparks JS. 2008. Phylogeny, taxonomy, and evolution of Neotropical cichlids (Teleostei: Cichlidae: Cichlinae). *Cladistics* 24:625–641.

Smith SA, Jann OC, Haig D, Russell GC, Werling D, Glass EJ, Emes RD. 2012. Adaptive evolution of Toll-like receptor 5 in domesticated mammals. *BMC Evol Biol.* 12:122.

Spady TC, Seehausen O, Loew ER, Jordan RC, Kocher TD, Carleton KL. 2005. Adaptive molecular evolution in the opsin genes of rapidly speciating cichlid species. *Mol Biol Evol.* 22:1412–1422.

Sparks JS, Smith LW. 2004. Phylogeny and biogeography of cichlid fishes (Teleostei: Perciformes: Cichlidae). *Cladistics* 20:501–517.

Stiassny MLJ. 1991. Phylogenetic intrarelationships of the family Cichlidae: an overview. In: Keenleyside MH, editor. Cichlid fishes: behaviour, ecology, and evolution. New York: Springer. p. 1–35.

Sugawara T, Imai H, Nikaido M, Imamoto Y, Okada N. 2010. Vertebrate rhodopsin adaptation to dim light via rapid meta-II intermediate formation. *Mol Biol Evol.* 27:506–519.

Sugawara T, Terai Y, Imai H, Turner GF, Koblmüller S, Sturmbauer C, Shichida Y, Okada N. 2005. Parallelism of amino acid changes at the RH1 affecting spectral sensitivity among deep-water cichlids from

Lakes Tanganyika and Malawi. *Proc Natl Acad Sci U S A.* 102: 5448–5453.
Suzuki Y. 2008. False-positive results obtained from the branch-site test of positive selection. *Genes Genet Syst.* 83:331–338.
Swanson WJ, Yang Z, Wolfner MF, Aquadro CF. 2001. Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. *Proc Natl Acad Sci U S A.* 98:2509–2514.
Takenaka N, Yokoyama S. 2007. Mechanisms of spectral tuning in the RH2 pigments of Tokay gecko and American chameleon. *Gene* 399: 26–32.
Terai Y, Seehausen O, Sasaki T, Takahashi K, Mizoiri S, Sugawara T, Sato T, Watanabe M, Konijnendijk N, Mrosso HDJ, et al. 2006. Divergent selection on opsins drives incipient speciation in Lake Victoria cichlids. *PLoS Biol.* 4:e433.
Veilleux CC, Louis EE Jr, Bolnick DA. 2013. Nocturnal light environments influence color vision and signatures of selection on the *OPN1SW* opsin gene in nocturnal lemurs. *Mol Biol Evol.* 30:1420–1437.
Wagner CE, Keller I, Wittwe S, Selz OM, Mwaiko S, Greuter L, Sivasundar A, Seehausen O. 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Mol Ecol.* 22:787–798.
Wald G. 1968. Molecular basis of visual excitation. *Science* 162:230–239.
Wang T, Duan Y. 2011. Retinal release from opsin in molecular dynamics simulations. *J Mol Recognit* 24:350–358.
Weadick CJ, Chang BSW. 2012a. An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. *Mol Biol Evol.* 29:1297–1300.
Weadick CJ, Chang BSW. 2012b. Complex patterns of divergence among green-sensitive (RH2a) African cichlid opsins revealed by Clade model analyses. *BMC Evol Biol.* 12:206.
Weadick CJ, Loew ER, Rodd FH, Chang BSW. 2012. Visual pigment molecular evolution in the Trinidadian pike cichlid (*Crenicichla frenata*): a less colorful world for Neotropical cichlids? *Mol Biol Evol.* 29: 3045–3060.
Weitz CJ, Nathans J. 1993. Rhodopsin activation: effects of the metarhodopsin I-metarhodopsin II equilibrium of neutralization or introduction of charged amino acids within putative transmembrane segments. *Biochemistry* 32:14176–14182.
Wimberger PH, Reis RE, Thornton KR. 1998. Mitochondrial phylogenetics, biogeography, and evolution of parental care and mating systems in *Gymnogeophagus* (Perciformes: Cichlidae). In: Malabarba LR, Reis RE, Vari RP, Lucena ZM, Lucena CAS, editors. Phylogeny and classification of Neotropical fishes. Porto Alegre (Brazil): Edipucrs. p. 509–518.
Yan ECY, Kazmi MA, De S, Chang BSW, Seibert C, Marin EP, Mathies RA, Sakmar TP. 2002. Function of extracellular loop 2 in rhodopsin: glutamic acid 181 modulates stability and absorption wavelength of metarhodopsin II †. *Biochemistry* 41:3620–3627.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 24:1586–1591.
Yang Z, dos Reis M. 2011. Statistical properties of the branch-site test of positive selection. *Mol Biol Evol.* 28:1217–1228.
Yang Z, Nielsen R, Goldman N. 2009. In defense of statistical methods for detecting positive selection. *Proc Natl Acad Sci U S A.* 106: E95–E96.
Yang Z, Nielsen R, Goldman N, Pedersen AM. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155:431–449.
Yang Z, Wong WSW, Nielsen R. 2005. Bayes Empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol.* 22: 1107–1118.
Yokoyama S, Tada T, Zhang H, Britt L. 2008. Elucidation of phenotypic adaptations: molecular analyses of dim-light vision proteins in vertebrates. *Proc Natl Acad Sci U S A.* 105:13480–13485.
Yoshida I, Sugiura W, Shibata J, Ren F, Yang Z, Tanaka H. 2011. Change of positive selection pressure on HIV-1 envelope gene inferred by early and recent samples. *PLoS One* 6:e18630.
Zhai W, Nielsen R, Goldman N, Yang Z. 2012. Looking for Darwin in genomic sequences— validity and success of statistical methods. *Mol Biol Evol.* 29:2889–2893.
Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol.* 22:2472–2479.
Zhao H, Ru B, Teeling EC, Faulkes CG, Zhang S, Rossiter SJ. 2009. Rhodopsin molecular evolution in mammals inhabiting low light environments. *PLoS One* 4:e8326.