Supplementary Materials

Appendix I: Multinucleotide Mutation Caveat

Our SWS2a analyses identified a burst of positive selection along a branch near the base of the South American cichlid radiation, which was also supported by the BUSTED analysis. Reconstruction of the nonsynonymous amino acid substitutions along this branch revealed two codons with multiple differences (or multinucleotide mutations; Venkat et al. 2018). Sparsely sampled datasets may have codons with multiple differences along branches that are not a result of DNA translation error (Venkat et al. 2018), but rather reflect an absence of taxa that may possess the intervening codon change. Denser species sampling could mitigate spurious statistical conclusions arising from multinucleotide mutations, since codons with multiple differences are likely prevalent along longer branches in phylogenies with distantly related species. In the case of the ancestral cichlid SWS2a branch, our sampling between nodes i and ii is reflective of the contemporary species diversity, so the inclusion of additional species would likely not improve the incidence of codons with multiple differences. Another avenue for mitigating inflated estimates of selection along branches with MNMs is to simply remove the sites in question, and test if the result holds (e.g., Dapper and Peyseur 2019). When we excluded all MNMs along this ancestral branch, our results were consistent (Table S11).

Tables

Table S1. Species information and accession numbers (see Excel file)

Table S2. Exon-specific primers used to amplify putative stop codon-containing regions in Neotropical cichlid SWS1

Name	Sequence	Target	Target Size	Species used to design primer
FH_E1_F3	GGGCCTTTTACTTGCAAGCT	Exon 1	242	*Apistogramma cacatuoides*
FH_E1_R3	GGCAGCCTCCATAGCACA	Exon 1	242	*Apistogramma cacatuoides*
FH_E4_F1	TCGGCCTCCACCCAGAAG	Exon 4	100	*Crenicara latruncularium*
FH_E4_R1	GGTGACAAGTCGATGATCTTTGT	Exon 4	100	*Crenicara latruncularium*
MA_E2_F2	CGTGACTTCTTGGTCCTTGG	Exon 2	168	*Mikrogeophagus altispinosus*
MA_E2_R2	GCTCCAGCCAAAGAAAGGT	Exon 2	168	*Mikrogeophagus altispinosus*
MA_E3_F2	CTGAAGGTCTGGGCTGCT	Exon 3	150	*Mikrogeophagus altispinosus*
MA_E3_R2	GAGCACCAAGTAGCTGGGAGT	Exon 3	150	*Mikrogeophagus altispinosus*
MA_E3_R3	CTGGGAGTAGCAGTCAGAGG	Exon 3	153	*Mikrogeophagus altispinosus*
UF_E4_F1	TCGGCCTCCACCCAGAAG	Exon 4	162	*Uaru fernandezyepezi*
UF_E4_R1	TTTCATTTGCGGAGTAGGCG	Exon 4	162	*Uaru fernandezyepezi*
UF_E4_F2	TCGGCCTCCACCCAGAAG	Exon 4	150	*Uaru fernandezyepezi*
UF_E4_R2	TGGTGACGGTCGATAATCTTTG	Exon 4	159	*Uaru fernandezyepezi*
UA_E4_F3	TCGGCCTCCACCCAGAAG	Exon 4	156	*Uaru amphiacanthoides*
UA_E4_R3	TGGAATGGTGACGAGTCGAT	Exon 4	156	*Uaru amphiacanthoides*

PCR Conditions

Reactions were performed in a total volume of 50uL: 33.4 uL double distilled H_20, 0.8uL Phusion polymerase (New England Biolabs), 5.0uL 10X Phusion buffer (New England Biolabs), 2.5uL of each primer (20uM), 5uL template DNA, 0.8uL dNTPs (12.5 uM each)

Thermocycling conditions were as follows: 95°C (3m); 35 cycles of 95°C (30s), 50°C (30s), 72°C (1m 45s); 72°C (10m). Amplified products were visualized on a 2% agarose gel stained with RedSafe and fragments were extracted using a QUIAquick Gel Extraction Kit (QIAGEN). Sequencing in the forward and reverse direction was performed by the Centre for Genome Evolution and Function (CAGEF) and chromatograms were examined in 4Peaks.

Table S3. Results of PAML random sites analyses on each SWS opsin

					Parameters†						
Opsin	**Model**	**np**	**lnL**	**k**	**ω_0/p**	**ω_1/q**	**ω_2/ω_p**	**Null**	**LRT**	**df**	***P***
SWS1	M0	115	-5334.0	3.0		0.20		n/a			
	M1a	116	-5292.3	3.0	0.12 (87%)	1 (13%)		M0	83.4	1	**0.00**
	M2a	118	-5292.0	3.0	0.12 (86%)	1 (12%)	1 (2%)	M1a	0.6	2	0.74
	M3	119	-5287.8	3.0	0.09 (79%)	0.7 (20%)	4.5 (0%)	M0	92.5	4	**0.00**
	M7	116	-5293.7	3.0	0.48	1.74		n/a			
	M8a	117	-5290.6	3.0	1.1	6.4	1 (9%)	n/a			
	M8	118	-5289.8	3.0	0.54	2.02	4.3 (0%)	M7	7.9	2	**0.02**
								M8a	1.7	1	0.31
SWS2b	M0	109	-4893.9	2.4		0.25		n/a			
	M1a	110	-4758.6	2.3	0.04 (81%)	1 (19%)		M0	270.6	1	**0.00**
	M2a	112	-4752.2	2.3	0.04 (80%)	1 (19%)	3.3 (1%)	M1a	13.0	2	**0.00**
	M3	113	-4752.1	2.4	0.04 (80%)	1 (19%)	3.3 (1%)	M0	284.0	4	**0.00**
	M7	110	-4763.4	2.3	0.1	0.3		n/a			
	M8a	111	-4759.4	2.3	0.44	9.5	1 (18%)	n/a			
	M8	112	-4753.8	2.4	0.12	0.46	2.8 (2%)	M7	19.2	2	**0.00**
								M8a	11.2	1	**0.00**
SWS2a	M0	121	-5313.2	3.4		0.3		n/a			
	M1a	122	-5034.7	2.7	0.02 (85%)	1 (15%)		M0	557.0	1	**0.00**
	M2a	124	-4987.5	3.0	0.03 (85%)	1 (8%)	3.1 (7%)	M1a	94.4	2	**0.00**
	M3	125	-4987.5	3.0	0.03 (85%)	1 (8%)	3.1 (7%)	M0	651.4	4	**0.00**
	M7	122	-5052.5	2.9	0.01	0.3		n/a			
	M8a	123	-5034.7	3.1	2.3	99	1 (16%)	n/a			
	M8	124	-4988.9	3.0	0.09	0.82	2.9 (8%)	M7	127.2	2	**0.00**
								M8a	91.6	1	**0.00**

NOTE.—np, number of parameters; lnL, ln likelihood; *K*, transition/transversion ratio; df, degrees of freedom; n/a, not applicable.

†ω values of each site class are shown for models M0–M3 ($\omega_0 - \omega_2$) 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.

Table S4. PAML random sites results for Central and South American cichlid subsets of the SWS1 opsin

Dataset	**Model**	**np**	**lnL**	**k**	**Parameters**				**Null**	**LRT**	**df**	***p***
Central America	M0	37	-1801.1	2.7	ω:	0.20			n/a			
	M1a	38	-1799.6	2.7	p:	0.87	0.13		M0	3.0	1.0	0.08
					ω:	0.08	1.00					
	M2a	40	-1799.5	2.7	p:	0.87	0.10	0.03	M1a	0.2	2.0	0.90
					ω:	0.08	1.00	1.00				
	M3	41	-1799.3	2.6	p:	0.65	0.35	0.00	M0	3.6	4.0	0.46
					ω:	0.00	0.55	999.00				
	M7	38	-1799.4	2.6	p:	0.21	q:	0.85	n/a			
	M8	40	-1799.3		p:	0.20	q:	0.88	M7	0.2	2.0	0.90
					p1:	0.00	ω:	999.00				
South America	M0	79	-4791.0	3.1	ω:	0.20			n/a			
	M1a	80	-4753.7	3.1	p:	0.87	0.13		M0	74.7	1.0	**0.00**
					ω:	0.11	1.00					
	M2a	82	-4751.2	3.1	p:	0.87	0.13	0.00	M1a	4.9	2.0	0.09
					ω:	0.11	1.00	5.44				
	M3	83	-4748.6		p:	0.78	0.21	0.00	M0	84.8	4.0	**0.00**
					ω:	0.09	0.65	5.09				
	M7	80	-4754.8	3.0	p:	0.44	q:	1.57	n/a			
	M8a	81	-4752.3	3.0	p:	0.96	q:	5.58	n/a			
					p1:	0.09	ω:	1.00				
	M8	82	-4750.0	3.0	p:	0.50	q:	1.88	M7	9.6	2.0	**0.01**
					p1:	0.00	ω:	4.96	M8a	4.6	1.0	**0.03**

Table S5. PAML random sites results for Central and South American cichlid subsets of the SWS2b opsin

Dataset	Model	np	lnL	k	Parameters				Null	LRT	df	P
Central America	M0	21	-1526.4	1.9	ω:	0.21			n/a			
	M1a	22	-1522.1	1.8	p:	0.84	0.16		M0	8.5	1.0	**0.00**
					ω:	0.00	1.00					
	M2a	24	-1521.3	1.9	p:	0.91	0.00	0.09	M1a	1.6	2.0	0.45
					ω:	0.00	1.00	2.40				
	M3	25	-1521.3	1.9	p:	0.44	0.47	0.09	M0	10.1	4.0	**0.04**
					ω:	0.00	0.00	2.40				
	M7	22	-1522.3	1.8	p:	0.05	q:	0.02	n/a			
	M8	24	-1521.3	1.9	p:	0.05	q:	2.20	M7	2.0	2.0	0.37
					p1:	0.09	ω:	2.42				
South America	M0	87	-4631.3	2.4	ω:	0.25			n/a			
	M1a	88	-4508.3	2.3	p:	0.80	0.20		M0	246.0	1.0	**0.00**
					ω:	0.04	1.00					
	M2a	90	-4501.7	2.3	p:	0.79	0.20	0.01	M1a	13.2	2.0	**0.00**
					ω:	0.04	1.00	3.82				
	M3	91	-4501.7	2.3	p:	0.79	0.20	0.01	M0	259.3	4.0	**0.00**
					ω:	0.04	0.97	3.70				
	M7	88	-4510.3	2.3	p:	0.09	q:	0.33	n/a			
	M8a	89	-4507.7		p:	0.40	q:	6.70	n/a			
					p1:	0.19	ω:	1.00				
	M8	90	-4502.5	2.3	p:	0.10	q:	0.40	M7	15.6	2.0	**0.00**
					p1:	0.01	ω:	3.15	M8a	10.4	1.0	**0.00**

Table S6. PAML random sites results for Central and South American cichlid subsets of the SWS2a opsin

Dataset	Model	np	lnL	k	Parameters				Null	LRT	df	*p*
Central America	M0	33	-1873.7	4.1	ω:	0.32			n/a			
	M1a	34	-1854.4	3.6	p:	0.81	0.19		M0	38.6	1.0	**0.00**
					ω:	0.00	1.00					
	M2a	36	-1844.2	3.7	p:	0.85	0.12	0.02	M1a	20.4	2.0	**0.00**
					ω:	0.04	1.00	8.70				
	M3	37	-1844.1	3.7	p:	0.74	0.23	0.02	M0	59.2	4.0	**0.00**
					ω:	0.00	0.66	8.54				
	M7	34	-1854.4	3.6	p:	0.01	q:	0.02	n/a			
	M8a	35	-1854.4	3.6	p:	0.01	q:	1.68	n/a			
					p1:	0.19	ω:	1.00				
	M8	36	-1844.1	3.7	p:	0.09	q:	0.45	M7	20.5	2.0	**0.00**
					p1:	0.02	ω:	8.59	M8a	20.5	1.0	**0.00**
South America	M0	87	-4722.4	3.2	ω:	0.31			n/a			
	M1a	88	-4493.2	2.7	p:	0.84	0.16		M0	458.4	1.0	**0.00**
					ω:	0.02	1.00					
	M2a	90	-4454.2	2.9	p:	0.85	0.08	0.07	M1a	78.0	2.0	**0.00**
					ω:	0.03	1.00	3.15				
	M3	91	-4454.2	2.8	p:	0.85	0.08	0.07	M0	536.4	4.0	**0.00**
					ω:	0.03	1.00	3.15				
	M7	88	-4501.0	2.6	p:	0.01	q:	0.03	n/a			
	M8a	89	-4491.3	2.6	p:	2.30	q:	99.00	n/a			
					p1:	0.16	ω:	1.00				
	M8	90	-4455.6	2.9	p:	0.09	q:	0.82	M7	90.8	2.0	**0.00**
					p1:	0.08	ω:	2.90	M8a	71.4	1.0	**0.00**

Table S7. Results of RELAX analyses on Neotropical cichlid SWS opsins

Opsin	Test (foreground) branches	Reference (background) branches	Model	lnL	Np	AIC	***k***	LR	p	Relaxation or Intensification of selection
SWS1	Pseudogenes	Intact sequences	Null	-5261.3	132	10788.7	1	--	--	
			Alternative	-5255.4	133	10778.8	**0.23**	11.9	**0.001**	Relaxation
	South American Cichlids	Central American Cichlids	Null	-5261.2	132	10788.5	1	--	--	
			Alternative	-5258.0	133	10884.1	**0.33**	6.45	**0.011**	Relaxation
	Dim/turbid water-dwelling cichlids	Clear water-dwelling cichlids	Null	-5261.3	132	10788.7	1	--	--	
			Alternative	-5259.0	133	10785.9	**0.65**	4.79	**0.03**	Relaxation
	Geophagini/ Astronotini clade	All other lineages	Null	-5261.3	132	10788.7	1	--	--	
			Alternative	-5256.6	133	10781.3	**0.55**	9.46	**0.002**	Relaxation
	Branch leading to major Neotropical cichlid tribes (i)	All other lineages	Null	-5261.3	132	10788.7	1	--	--	
			Alternative	-5261.3	133	10790.7	1.02	0.0	0.94	n/s
	Branch leading to Geophagini/ Astronotini (ii)	All other lineages	Null	-5261.3	132	10788.7	1	--	--	
			Alternative	-5261.1	133	10790.3	0.0	0.39	0.53	n/s
	Branch leading to Heroini/Cichlasomatini (ii)	All other lineages	Null	-5261.3	132	10788.7	1	--	--	
			Alternative	-5261.1	133	10790.3	0.0	0.39	0.53	n/s
	Branch leading to Geophagini (v)	All other lineages	Null	-5261.3	132	10788.7	1	--	--	
			Alternative	-5261.2	133	10790.5	1.37	0.26	0.61	n/s
SWS2b	Dim/turbid water dwelling cichlids	Clear water-dwelling cichlids	Null	-4837.0	124	9923.9	1	--	**--**	
			Alternative	-4835.4	125	9922.6	0.65	3.33	0.07	n/s
	South American Cichlids	Central American Cichlids	Null	-4837.0	124	9923.9	1	--	--	
			Alternative	-4837.0	125	9925.9	0.97	0.01	0.93	n/s
SWS2a	Dim/turbid water dwelling cichlids	Clear water-dwelling cichlids	Null	-5232.3	138	10742.0	1	--	--	
			Alternative	-5232.0	139	10743.9	0.91	0.14	0.71	n/s
	South American cichlids	Central American Cichlids	Null	-5232.3	138	10742.6	1	--	--	
			Alternative	-5232.3	139	10742.6	1	0	1.0	n/s

NOTE—lnL = ln likelihood; np = number of parameters; AICc = akaike information criterion corrected for sample size; *k=1* is the null hypothesis while *k<1* indicates relaxed selection and *k>1* indicates an intensification of selection. LR = likelihood ratio of alternative model relative to null model. For SWS1, test branches leading to major cichlid clades are denoted with a letter, and correspond to the branches marked in Figure 1.

Table S8. Results of HYPHY BUSTED analyses on the SWS2a opsin branch under episodic selection (per CMC analyses)

Analysis	Test branch	Model	lnL	Np	AIC_c	LRT	p
BUSTED	Branch leading to Geophagini/Astronotini/ Chaetobranchini clade (ii)	Null	-5235.0	142	10756.1	--	--
		Alternative	-5266.1	143	10740.2	33.3	**0.00**

Table S9. All SWS2a codon changes occurring along the Astronotini/Chaetobranchini/Geophagini ancestral branch (branch ii, Figure 2) under episodic selection

Site	Codon change	Amino acid change	Type of change (nonsynonymous or synonymous)	Multi-nucleotide mutation?	Under divergent selection? (CMC)	Increased hydrophobicity?
49	GTA → ATA	V → I	NS	no	yes	Yes
88	TGT → GTT	C → V	NS	yes	yes	No
99	TCA → GCA	S → A	NS	no	yes	Yes
166	GCT → ATT	A → I	NS	yes	yes	Yes
194	ACA → TCA	T → S	NS	no	yes	No
207	CTT → CTG	L → L	S	no	no	n/a

Table S10. SWS2a CMC and BUSTED results with multinucleotide mutations along the Geophagini/Astronotini/Chaetobranchini branch removed

							Parameters						
Analysis type	Model/Foreground	np	lnL	AIC	ΔAIC	k	ω_0	ω_1	ω_2/ω_d	Null	LRT	df	*p*
CMC (PAML)	M2a_rel	136	-5358.7	10989.4	5.0	3.0	0.03(84%)	1(9%)	3.0(6.5%)				
	Geo/Astro/Chaeto branch (ii)	137	-5355.2	10984.4	0	2.9	0.03(84%)	1 (9%)	3.0(6.5%) G/A/C_branch: 999	M2a_rel	7.0	1.0	**0.01**
BUSTED (HYPHY)	Constrained model	154	-5622.3	11554.9	9.5	n/a	Test:0.0(8.96%)	0.0 (89.8%)	1 (1.19%)				
							BG: 0.0 (76.2%)	0.15(16.9%)	3.86 (6.94%)				
	Unconstrained (G/A/C test branch (ii))	155	-5616.0	11544.4	0	n/a	Test:0.0(8.96%)	0.0 (89.8%)	456.3 (1.19%)	n/a	12.6	1.0	**0.002**
							BG: 0.0 (76.2%)	0.15(16.8%)	3.63 (7.05%)				

Table S11. IDs for photographs in Figure 2 (top to bottom)

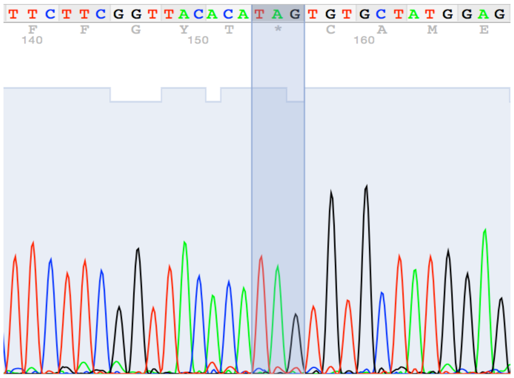
Genus and species (where applicable
Retroculus sp.
Astronotus ocellatus
Crenicara punctatum
Satanoperca sp.
Geophagus dicrozoster
Acarichthys heckelii
Crenicichla wallacii
Crenicichla lugubris
Aequidens sp.
Acaronia nassa
Andinoacara rivulatus
Pterophyllum scalare
Symphysodon sp.
Rocio octofasciata
Caquetaia spectabile
Herichthys sp.
Petenia splendida
Parachromis sp.
Amphilophus cintrinellus

Table S12. SWS1 Amino acid alignment and spectral tuning sites (Excel file)
Table S13. SWS2b Amino acid alignment and spectral tuning sites (Excel file)
Table S14. SWS2a Amino acid alignment and spectral tuning sites (Excel file)

Figures

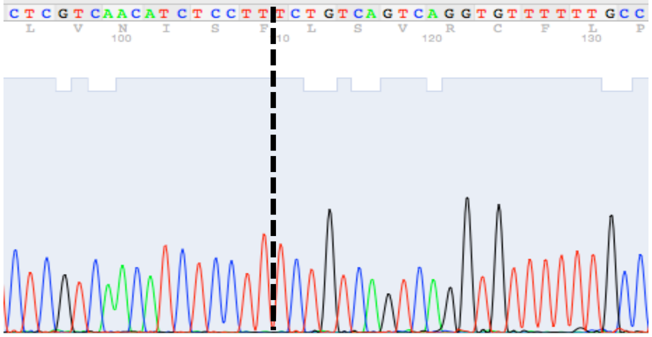
a) *Apistogramma cacatuoides* – premature stop codon

```
                              293                          318
>G_Mazarunia_charadrica_T06044  CTGGGTTACACATTGTGTGCTTTGGAG
>G_Guianacara_dacrya_T06410     CTGGGTTACACATTGTGTGCTTTGGAA
>G_Geophagus_dicrozoster_T09320 CTGGGTTATACATTGTGTGCTTTGGAG
>G_Geophagus_harreri_T11823     CTGGGTTATACATTGTGTGCTTTGGAG
>A_Apistogramma_panduro         TTCGGTTACACATTGTGTGCTATGGAG
>G_Apistogramma_cacatuoides_PG  TTCGGTTACACATAGTGTGCTATGGAG
```



b) *Apistogramma panduro* – deletion. Dotted line indicates position of the 25bp deletion

```
                                205                                                             270
>C_Nannacara_bimaculata_T12641   CTCGTCAACATTTCCTTTGCTGGCTTCATCTTTGTCACTTTCTCCGTCAGTCAGGTGTTCGTAGCC
>G_Acarichthys_heckelii_T06757   CTCGTCAACATCTCCTTTGCTGGCTTCATCTTTGTCACTTTCTCTGTCAGTCAGGTGTTCTTTGCC
>G_Mazarunia_charadrica_T06044   CTCGTCAACATCTCCTTTGCTGGCTTCATCTTTGTCACTTTCTCCGTCAGTCAGGTGTTCTTTGCC
>G_Guianacara_dacrya_T06410      CTCGTCAACATCTCCTTTGCTGGCTTCATCTTTGTCACTTTCTCTGTCAGTCAGGTGTTCTTTGCC
>G_Geophagus_dicrozoster_T09320  CTCGTCAACATCTCCTTTTCTGGCTTCATCTTTGTCACTTTCTCTGTCAGTCAGGTGTTCTTTGCC
>G_Geophagus_harreri_T11823      CTCGTCAACATCTCCTTTTCTGGCTTCATCTTTGTCACTTTCTCTGTCAGTCAGGTGTTCTTTGCC
>A_Apistogramma_panduro          CTCGTCAACATCTCCTT-------------------------TCTGTCAGTCAGGTGTTTTTTGCC
```

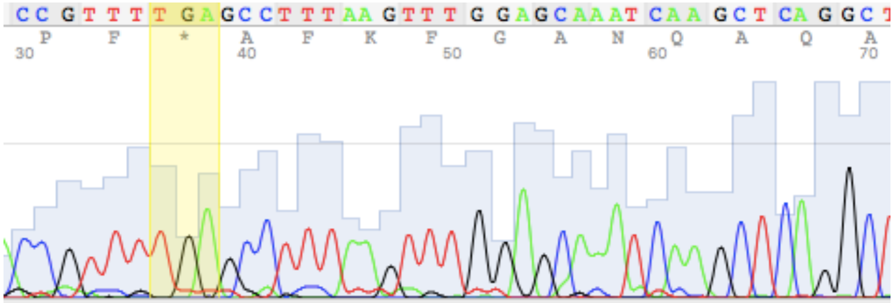


c) *Crenicara latruncularium* – premature stop codon in exon 4

```
                                   715                                750
>C_Aequidens_tetramerus_T11731   GCAGAGAAGGAGGTATCGAGGATGATCATTGTGATG
>C_Krobia_petitella_T08034       GCAGAGAAGGAGGTATCGAGGATGATCATTGTGATG
>C_Nannacara_anomala_T13932      GCGGAGAAGGAGGTATCGAGGATGATCATTGTGATG
>C_Nannacara_bimaculata_T12641   GCGGAGAAGGAGGTATCGAGGATGATCATTGTGATG
>G_Mazarunia_charadrica_T06044   GCAGAAAAGGAGGTATCGAGGATGATCATTGTGATG
>G_Crenicara_latruncularium_PG   GCAGAAAAGGAGGTATAGAGGATGATCATTGTGATG
```

d) *Mikrogeophagus altispinosus* – premature stop codon

```
                              404                               435
>B_Retroculus_sp_T11750          CCGTTTGGAGCCTTTAAGTTTGGAGCAAATCATGC
>B_Retroculus_xinguensis_T15042  CCGTTTGGAGCCTTTAAGTTTGGAGCAAATCATGC
>C_Cichlasoma_dimerus_T15879     CCATTTGGAGCCTTTAAGTTTGGAGCAAATCACGC
>H_Uaru_fernandezyepezi_PG       CCATTTGGAGCCTTTAAGTTTGGAGCAAATCACGC
>A_Astronotus_sp_PG              CCATTTGGAGCCTTTAAGTTTGGAGCAAACCATGC
>G_Mikrogeophagus_altispinosus   CCGTTTTGAGCCTTTAAGTTTGGAGCAAATCAAGC
```



e) *Uaru fernandezyepezi* – premature stop codon

```
                                 709                            738
>H_Astatheros_macracanthus_T11918  CAGAAGGCGGAGAAGGAGGTATCGAGGATG
>H_Rocia_octofasciatus_T13078      CAGAAGGCGGAGAAGGAGGTATCGAGGATG
>H_Tomocichla_tuba_T11909          CAGAAGGCGGAGAAGGAGGTATCGAGGATG
>H_Herotilapia_multispinosa_T12226 CAGAAGGCGGAGAAGGAGGTATCCAGGATG
>H_Symphysodon_discus_T11830       CAGAAGGCGGAGAAGGAGGTATCGAGGATG
>H_Uaru_fernandezyepezi_PG*        CAGAAGGTGGAGAAGTAGGTATCGAGGATG
```

`*Exon did not amplify`

Figure S1. Nonsense mutations and/or indels identified in Neotropical cichlid SWS1 opsin genes

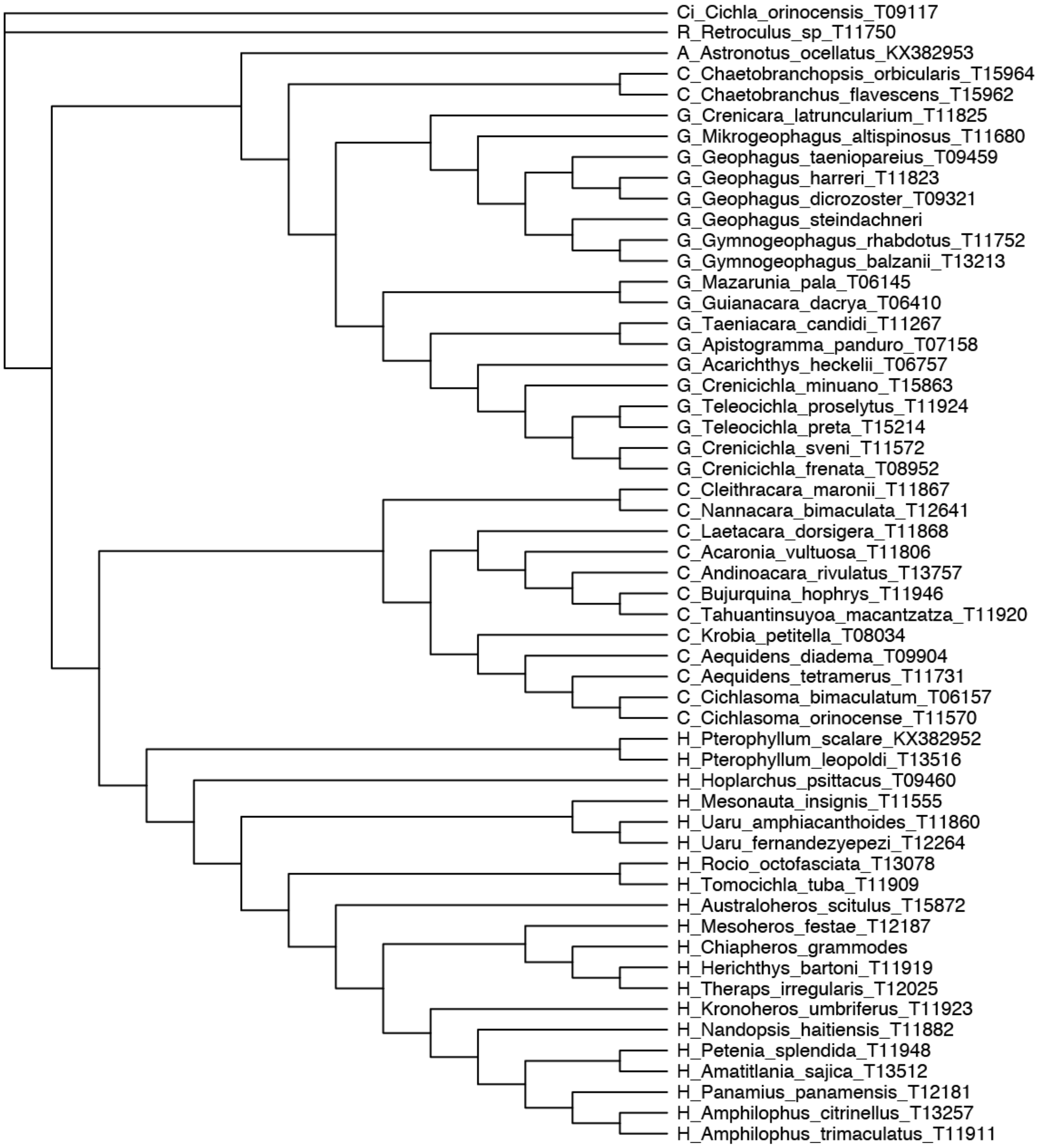


Figure S2. Neotropical cichlid species tree used for SWS2b analyses. Abbreviations denote major Neotropical cichlid tribes: Ci = Cichlini; R = Retroculini; A = Astronotini; C = Chaetobranchini; G = Geophagini; C = Cichlasomatini; H = Heroini. Numbers denote Royal Ontario Museum tissue vouchers.

References

Dapper AL, Payseur BA (2019) Molecular evolution of the meiotic recombination pathway in mammals. *Evolution*, **73**, 2368–2389.

Venkat A, Hahn MW, Thornton JW (2018) Multinucleotide mutations cause false inferences of lineage-specific positive selection. *Nature Ecology & Evolution*, **2**, 1280–1288.