Electronic Supplementary Material (ESM)

To see or not to see: Molecular evolution of the rhodopsin visual pigment in neotropical electric fishes

DOI 10.1098/rspb.**2019.1182**

Alexander Van Nynatten[1,2], Francesco H Janzen[3,4], Kristen Brochu[5], Javier A Maldonado-Ocampo[6], William GR Crampton[7], Belinda SW Chang[1,8,9], Nathan R Lovejoy[1,2,8]

[1]Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada M5S 3G5

[2]Department of Biological Sciences, University of Toronto Scarborough, Toronto, ON, Canada M1C 1A4

[3] Department of Biology, University of Ottawa, Ottawa, ON, Canada K1N 6N5

[4]Canadian Museum of Nature, Ottawa, ON, Canada K1P 6P4

[5]Department of Entomology, Penn State University, University Park, PA, USA 16802

[6]Laboratorio de Ictiología, Unidad de Ecología y Sistemática -UNESIS-, Departamento de Biología, Facultad de Ciencias, Pontificia Universidad Javeriana, Bogotá, Colombia

[7]Department of Biology, University of Central Florida, Orlando, FL, USA 32816

[8]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada M5S 3B2

[9]Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, ON, Canada M5S 3B2

Authors for correspondence: Belinda SW Chang – belinda.chang@utoronto.ca, Nathan R Lovejoy – lovejoy@utsc.utoronto.ca

Supplementary Methods

We amplified and sequenced rhodopsin from 147 gymnotiform species (**table S4**) using forward and reverse primers Rh193F (CNTATGAATAYCCTCAGTACTACC) and Rh1073R (CCRCAGCACARCGTGGTGATCATG) from [1]. Each amplification reaction contained 1X mixed Taq buffer, 2.5 mM MgCl2, 0.8 mM dNTPs, 0.5 μM of each primer, 0.05 U of Taq DNA polymerase, recombinant, approximately 20-30 ng of template DNA extract, and enough sterile, nuclease-free water to adjust to the final reaction volume of 30 μL. Touchdown PCR cycling conditions for rhodopsin included an initial heating to 94°C for 4 min, then 6 cycles of DNA denaturation at 95°C for 30s, primer annealing at 54°C for 30s (decreasing by 1°C each cycle), and DNA extension at 72°C for 1 min. Once completed, a second series of heating cycles took place with 34 cycles of DNA denaturation at 95°C for 30s, primer annealing at 48°C for 30s, and DNA extension at 72°C for 1 min. A final extension phase at 72°C for 5 min took place at the end of the cycling program.

The gymnotiform dataset was aligned using the DECIPHER package in R [2], and a rhodopsin gene tree was generated using PhyML [3] using the best fitting nucleotide substitution model, HKY + I + G, as determined by ModelFinder [4] (**figure S2**). Support for each node was determined by 100 bootstrap replicates in PhyML.

We downloaded sequences from Genbank to generate a large rhodopsin dataset using BlastPhyME [5], and we pruned it to consist only of rhodopsin sequences belonging to species represented in the robust multi-gene species tree generated in [6], keeping all sequences longer than 700 bp for characiphysian fishes but only representative species for other major vertebrate lineages spanning Gnathostomata for computational tractability (**table S5**).We used the topology from [6] (**figure S3**) for subsequent analyses.

Random-sites models in PAML (M0, M1a and M2a) have one to three site classes respectively [7]. M2a has one site class where d_N/d_S can exceed one. Support for positive selection ($d_N/d_S > 1$) is tested by comparing the fit of M2a to M1a. Models M7 and M8 are an extension of M1a and

M2a with ten site classes with d_N/d_S estimates defined by a beta distribution [7]. The Branch-sites test was used to compare d_N/d_S estimates for the branch leading to the gymnotiform clade with the rest of the tree [8]. This model allows positive selection at a subset of sites only on the selected branch or clade (herein called the foreground) and its fit is compared to a nested null model assuming no positive selection by fixing the d_N/d_S estimate for the foreground to equal one [8]. Clade models C and D allow d_N/d_S to differ in a subset of sites in pre-determined phylogenetic partitions, but does not restrict the estimate to be greater than one [9]. The fit of these models were compared with nested null models M2aREL and M3 where the d_N/d_S estimate for the divergent site class in CmC and CmD is collapsed to uniformity across the phylogeny. We used clade models to compare d_N/d_S in gymnotiform rhodopsin with other vertebrates using the vertebrate rhodopsin dataset. We also used clade models to compare d_N/d_S among gymnotiform lineages, using the gymnotiform rhodopsin dataset. We partitioned clades by electric organ discharge (EOD) type: we partitioned by wave-type and pulse-type EODs, and also compared the electric eel (able to produce high-voltage EODs) with all other gymnotiform species (**Figure S2**). FUBAR, RELAX and the adaptive branch-site REL (aBSREL) models were accessed through the Datamonkey webserver [10-14]. RELAX was used to test if selection pressures have relaxed or intensified on the gymnotiform clade with cypriniforms used as the reference. The aBSREL model was used on a dataset comprised only of gymnotiforms without the specification of any foreground lineage. Sites under positive selection were identified using Bayes Empirical Bayes [15]. Ancestral reconstructions were reconstructed in PAML [16] using the best fitting empirical amino acid substitution model, mtZOA+I+G4, for the 50 species vertebrate dataset and the 147 species Gymnotiformes dataset as determined by ModelFinder [4]. Outgroups were included in ancestral reconstructions for the gymnotiforms dataset to help polarize the amino acid identities in gymnotiforms (see **figure S2**). Ancestral amino acid sequences were also inferred from ancestral codon sequences reconstructed using model m3 in PAML. Amino acid substitutions were modelled onto the rhodopsin meta II active-state structure [17] using UCSF Chimera [18].

Table S1. PAML analyses of the gymnotiform rhodopsin dataset.

Model	np	lnL	Parameter Estimates: d_N/d_S (proportion of sites): background d_N/d_S / foreground d_N/d_S	Null	LRT	p value
m0	293	-8566.72	0.03(1.00)			
m1a	294	-8393.80	0.01(0.96) 1.00(0.04)	m0	345.84	< 0.0001
m2a	296	-8393.80	0.01(0.96) 1.00(0.04) 27.15(0.00)	m1a	0	1.0000
m3 (2 site classes)	295	-8283.06	0.00(0.91) 0.25(0.09)			
m2aREL	296	-8280.60	0.00(0.90) 1.00(0.01) 0.21(0.09)	m3 (ns2)	4.92	0.0227
M3	297	-8267.28	0.00(0.88) 0.10(0.08) 0.38(0.05)	M2aREL	26.64	< 0.0001
m7	294	-8272.85	p = 0.07, q = 1.39			
m8	296	-8272.85	p = 0.07, q = 1.39, 2.74(0.00)	m7	0	1.0000
Partitioned models based on electric organ discharge (EOD)						
High-voltage EOD						
Branch site (null)	295	-8393.73	0.01(0.96) 1.00(0.04) 0.01/1.00(0.00) 1.00/1.00(0.00)			
Branch site	296	-8393.73	0.01(0.96) 1.00(0.04) 0.01/1.00(0.00) 1.00/1.00(0.00)	Branch site (null)	0	1.0000
CmC	297	-8280.01	0.00(0.90) 1.00(0.01) 0.21/0.10(0.09)	m2aREL	1.18	0.2774
CmD	298	-8266.14	0.00(0.88) 0.38(0.05) 0.10/0.02(0.07)	m3 (3 site classes)	2.28	0.1307
Wave-type EOD						
Branch site (null)	296	-8393.73	0.01(0.96) 1.00(0.04) 0.01/1.00(0.00) 1.00/1.00(0.00)			
Branch site	296	-8393.73	0.01(0.96) 1.00(0.04) 0.01/1.00(0.00) 1.00/1.00(0.00)	Branch site (null)	0	1.0000
CmC	297	-8280.00	0.00(0.90) 1.00(0.01) 0.23/0.18(0.09)	m2aREL	1.2	0.2733
CmD	298	-8267.12	0.00(0.88) 0.05(0.38) 0.09/0.11(0.08)	m3 (3 site classes)	0.32	0.5716

Note: lnL, ln likelihood; *LRT, likelihood ratio test result;* Br-site, Branch-site; EOD, Electric Organ Discharge

Table S2. PAML analyses of the vertebrate rhodopsin dataset.

Model	np	lnL	Parameter Estimates: d_N/d_S (proportion of sites): background d_N/d_S / foreground d_N/d_S	Null	LRT	p value
M0	99	-14551.75	0.05(1.00)			
M1	100	-14345.03	0.04(0.93), 1.00(0.07)	m0	413.44	< 0.0001
M2	102	-14345.03	0.04(0.93), 1.00(0.03), 1.00(0.04)	m1	0	1.0000
m7	100	-13947.32	p = 0.24, q = 3.07			
m8a	101	-13945.23	p = 0.25, q = 3.50			
m8	102	-13945.23	p = 0.26, q = 3.50 1.00(0.01)	m7	4.18	0.1237
				m8a	0	1.0000
M2aREL	102	-14001.94	0.01(0.73), 1.00(0.01), 0.17(0.26)			
m3	103	-13959.18	0.01(0.64), 0.10(0.28), 0.36(0.08)			
Partitioned models with Gymnotiform species separate from other vertebrate lineages						
Gymnotiformes branch						
Branch Site (null)	101	-14342.37	0.04(0.90), 1.00(0.07), 0.04/1.00(0.03), 1.00/1.00(0.00)			
Branch Site	102	-14340.37	0.04(0.91), 1.00(0.07), 0.04/83.14(0.02), 1.00/83.14(0.00)	BrS null	4	0.0455
CmC	103	-14001.94	0.01(0.73), 1.00(0.01), 0.17/0.17(0.26)	M2aREL	0	1.0000
CmD	104	-13959.02	0.01(0.64), 0.10(0.28), 0.36/0.26(0.08)	m3	0.32	0.5716
Gymnotiformes Clade						
Branch Site (null)	101	-14340.81	0.04(0.92), 1.00(0.07), 0.04/1.00(0.01), 1.00/1.00(0.00)			
Branch Site	102	-14340.59	0.04(0.92), 1.00(0.07), 0.04/1.79(0.00), 1.00/1.79(0.00)	BrS null	0.44	0.5071
CmC	103	-13988.66	0.01(0.72), 1.00(0.01), 0.18/0.08(0.27)	M2aREL	26.56	< 0.0001
CmD	104	-13951.80	0.01(0.63), 0.09(0.26), 0.33/0.12(0.11)	m3	14.76	< 0.0001
Gymnotiformes branch and clade						
CmC	104	-13988.58	0.01(0.72), 1.00(0.01), 0.19/0.16/0.08(0.27)	M2aREL	26.72	< 0.0001
CmD	105	-13951.56	0.01(0.63), 0.09(0.26), 0.34/0.23/0.12(0.11)	m3	15.24	0.0001
Cypriniformes clade						
CmC	103	-14001.45	0.01(0.73) 1.00(0.01) 0.18/0.16(0.26)	M2aREL	0.98	0.3222
CmD	104	-13958.94	0.01(0.64) 0.36(0.08) 0.10/0.09(0.28)	m3	0.48	0.4884
Gymnotiformes and Cypriniformes clade						
CmC	104	-13987.62	0.01(0.72) 1.00(0.01) 0.20/0.09/0.16(0.27)	M2aREL	28.64	< 0.0001
CmD	105	-13950.85	0.01(0.64) 0.10(0.27) 0.40/0.16/0.32(0.09)	m3	16.66	0.0002
				Gymnotiformes CmC	2.08	0.1492
				Gymnotiformes CmD	1.90	0.1675
				Cypriniformes CmC	27.66	< 0.0001
				Cypriniformes CmD	23.24	< 0.0001

Note: lnL, ln likelihood; *LRT, likelihood ratio test result;* CmC, Clade model C; CmD, Clade model D

Table S3. Ancestral amino acid and codon reconstructions at select nodes in the vertebrate rhodopsin dataset.

Branch	Substitution	PPi | PPf	Method
6 -> 7	V209I	0.529 | 0.554	aa
7 -> 8	I209C	0.554 | 0.514	aa
9 -> 10	C209A	0.514 | 1.000	aa
7 -> 8	V209S	0.488 | 0.605	codon
9 -> 10	S209A	0.605 | 1.000	codon
6 -> 7	S213L	0.897 | 0.753	aa
4 -> 5	I213T	0.762 | 0.899	codon
7 -> 8	T213L	0.880 | 0.999	codon
7 -> 8	I214T	0.999 | 1.000	aa
7 -> 8	I214T	0.999 | 0.996	codon
9 -> 10	L216T	1.000 | 1.000	aa
9 -> 10	L216T	1.000 | 0.999	codon
3 -> 4	**I217F**	**0.989 | 0.647**	**aa**
7 -> 8	**I217F**	**0.997 | 0.996**	**codon**
3 -> 4	V218I	0.99 | 0.895	aa
9 -> 10	I218T	0.999 | 1.000	aa
3 -> 4	V218I	0.929 | 0.865	codon
9 -> 10	I218T	1.000 | 0.999	codon
9 -> 10	T220C	1.000 | 1.000	aa
4 -> 5	**F220T**	**0.728 | 0.928**	**codon**
9 -> 10	**T220C**	**0.984 | 1.000**	**codon**
3 -> 4	T229A	0.999 | 0.999	aa
6 -> 7	A229T	1.000 | 0.990	aa
3 -> 4	T229A	1.000 | 0.966	codon
6 -> 7	A229T	0.967 | 0.968	codon
9 -> 10	I256V	0.938 | 1.000	aa
9 -> 10	I256V	0.966 | 0.999	codon
4 -> 5	**V258F**	**0.998 | 0.997**	**aa**
4 -> 5	**V258F**	**0.995 | 0.964**	**codon**
3 -> 4	S260A	0.953 | 0.552	aa
3 -> 4	S260A	0.841 | 0.802	codon
7 -> 8	F261Y	0.998 | 1.000	aa
7 -> 8	F261Y	0.998 | 1.000	codon
7 -> 8	I263V	0.811 | 0.995	codon
5 -> 6	V266L	0.599 | 0.998	aa
4 -> 5	V266L	0.602 | 0.565	codon
9 -> 10	S270H	1.000 | 1.000	aa
9 -> 10	S270H	0.999 | 0.999	codon
7 -> 8	V271T	0.999 | 0.765	aa
7 -> 8	V271T	0.996 | 0.685	codon
1 -> 2	F273W	0.984 | 0.996	aa
1 -> 2	F273W	0.878 | 1.000	codon
7 -> 8	Y274W	0.997 | 0.771	aa
7 -> 8	Y274W	0.998 | 0.490	codon

Note: PPi, best supported (highest posterior probability) amino acid identity in node i; PPf, best supported amino acid identity in node f; Bold sites discussed in manuscript, Branches match those labelled in figure 1 with the first label representing node I and the second node f. mtZOA+I+G used for amino acid reconstructions and M3 used for codon reconstructions.

Tables S4. Voucher and Genbank accession numbers for Gymnotiform rhodopsin sequences generated for this study

Species	ID	Voucher	Accession #
Adontosternarchus balaenops	2612	UF 116559	MN031442.1
Adontosternarchus clarkae	2906	MCP 39341	MN031445.1
Adontosternarchus clarkae	8673	INPA 28867	MN031443.1
Adontosternarchus nebulosus	2892	MCP 39313	MN031446.1
Adontosternarchus nebulosus	8679	NA	MN031444.1
Akawaio penak	8796	ROM 83884	MN031447.1
Apteronotus albifrons	7301	MNRJ 33616	MN031448.1
Apteronotus anu	8703	NA	MN031449.1
Apteronotus apurensis	8688	UNELLEZ 41	MN031450.1
Apteronotus bonapartii	2914	NA	MN031451.1
Apteronotus eschmeyeri	4001	IAvHP 7806	MN031452.1
Apteronotus galvisi	8700	IAvHP-8635	MN031453.1
Apteronotus leptorhynchus	8704	IAvHP 9682	MN031454.1
Apteronotus macrolepis	7110	NA	MN031455.1
Apteronotus magdalenensis	4009	IAvHP 7829	MN031456.1
Apteronotus mariae	2813	IAvHP 7812	MN031457.1
Archolaemus luciae	11857	ZUEC 15782	MN031459.1
Brachyhypopomus alberti	7046	CBF 10278	MN031460.1
Brachyhypopomus batesi	2414	MCP 45312	MN031461.1
Brachyhypopomus beebei	6967	UF 177358	MN031462.1
Brachyhypopomus belindae	2132	MCP 45431	MN031463.1
Brachyhypopomus bennetti	2136	MCP 45255	MN031464.1
Brachyhypopomus bombilla	9104	UFRGS 10561	MN031465.1
Brachyhypopomus brevirostris	7019	UF 177359	MN038185.1
Brachyhypopomus diazae	305	UF174334	MN031467.1
Brachyhypopomus draco	9101	UFRGS 14562	MN031468.1
Brachyhypopomus flavipomus	2141	MCP 45265	MN031469.1
Brachyhypopomus hamiltoni	7234	MCP 45681	MN031470.1
Brachyhypopomus occidentalis	8780	NA	MN031472.1
Brachyhypopomus palenque	2433	UF 148572	MN031473.1
Brachyhypopomus pinnicaudatus	2121	MCP 45281	MN031474.1
Brachyhypopomus provenzanoi	2365	UF 177347	MN031475.1
Brachyhypopomus regani	7040	UMSS 7038	MN031476.1
Brachyhypopomus sp. 1	10344	NA	MN031477.1
Brachyhypopomus sp. 2	11994	NA	MN031478.1
Brachyhypopomus sp. 3	11999	NA	MN031479.1
Brachyhypopomus sp. 4	11995	NA	MN031480.1
Brachyhypopomus sp. 5	11997	MCP 49410	MN031466.1
Brachyhypopomus sp. 6	6966	UF 177365	MN031471.1
Brachyhypopomus sullivani	7039	UF 177341	MN031481.1
Brachyhypopomus verdii	2254	UF 148520	MN031482.1
Brachyhypopomus walteri	7048	CBF 10257	MN031483.1
Compsaraia compsa	8720	INPA 28876	MN031484.1

Compsaraia samueli	11036	ANSP 182209	MN031485.1
Distocyclus conirostris	7306	INPA 28879	MN031486.1
Eigenmannia sp. 1	10964	ANSP 194529	MN031487.1
Electrophorus electricus	9793	NA	MN031488.1
Electrophorus sp. 1	2026	MZUSP 103219	MN031489.1
Gymnorhamphichthys bogardusae	10974	ANSP 199558	MN031490.1
Gymnorhamphichthys britskii	11635	LBP 3813/22012	MN031491.1
Gymnorhamphichthys hypostomus	11851	NA	MN031493.1
Gymnorhamphichthys rondoni	2154	MCP 46936	MN031495.1
Gymnorhamphichthys rondoni	10968	NA	MN031492.1
Gymnorhamphichthys rondoni	11646	ROM 97536	MN031494.1
Gymnorhamphichthys rosamariae	12000	ANSP 191142	MN031496.1
Gymnorhamphichthys sp. 1	10965	AUM 48205	MN031497.1
Gymnotus anguillaris	9944	KU41321	MN031500.1
Gymnotus anguillaris	10545	ROM 100941	MN031498.1
Gymnotus bahianus	7244	MZUSP 102898	MN031499.1
Gymnotus chimarrao	11051	MCP 28583	MN031501.1
Gymnotus esmeraldas	10865	ZOO.A.V.Pe0310	MN031502.1
Gymnotus henni	8189	IMCN 4521	MN031503.1
Gymnotus inaequilabiatus	10977	ANSP 192991	MN031504.1
Gymnotus javari	2020	UF 122824	MN031505.1
Gymnotus jonasi	2471	UF 131410	MN031506.1
Gymnotus obscurus	2018	MZUSP 75157	MN031507.1
Gymnotus omarorum	7093	AMNH 239656	MN031508.1
Gymnotus panamensis	8210	STRI-01579	MN031509.1
Gymnotus pantanal	7076	NA	MN031510.1
Gymnotus sp. 1	10347	NA	MN031511.1
Gymnotus sp. 2	2091	AUM 36616	MN031512.1
Gymnotus sp. 3	7109	UNCAT. MCP	MN031513.1
Gymnotus sp. 4	8761	MNRJ 33642	MN031514.1
Gymnotus sp. 5	8779	MNRJ 33630	MN031515.1
Gymnotus stenoleucus	2060	UF 174329	MN031516.1
Gymnotus sylvius	7240	MZUSP 100267	MN031517.1
Gymnotus ucamara	1950	UF 126184	MN031518.1
Hypopygus lepturus	2438	NA	MN031519.1
Hypopygus minissimus	2336	UF 175389	MN031520.1
Hypopygus neblinae	2337	UF 148540	MN031521.1
Hypopygus nijsseni	2216	MCP 44651	MN031522.1
Japigny kirschbaum	8992	MHNG 2682.031	MN031523.1
Megadontognathus kaitukaensis	10970	ANSP 195961	MN031526.1
Melanosternarchus amaru	11843	MUSM 56870	MN031562.1
Microsternarchus bilineatus	2138	MCP 45480	MN031527.1
Microsternarchus sp. 1	11996	NA	MN031528.1
Microsternarchus sp. 2	11998	NA	MN031529.1
Microsternarchus sp. 3	10348	NA	MN031530.1
Orthosternarchus tamandua	2625	UF 116562	MN031531.1

Platyurosternarchus macrostoma	8726	MNRJ 33614	MN031532.1
Porotergus duende	2916	MCP 37359	MN031533.1
Porotergus gymnotus	10957	NA	MN031534.1
Porotergus sp. 1	11842	NA	MN031458.1
Procerusternarchus pixuna	11638	LBP 7006/34063	MN031535.1
Racenisia fimbriipinna	2340	UF 177352	MN031536.1
Rhabdolichops caviceps	2883	MCP 36007	MN031537.1
Rhabdolichops caviceps	8994	UNELLEZ	MN031539.1
Rhabdolichops caviceps	10990	NA	MN031538.1
Rhabdolichops eastwardi	2104	MCP 36025	MN031543.1
Rhabdolichops eastwardi	8996	INPA 28911	MN031540.1
Rhabdolichops eastwardi	9014	NA	MN031541.1
Rhabdolichops electrogrammus	2898	MCP 36029	MN031544.1
Rhabdolichops electrogrammus	9004	INPA 28863	MN031542.1
Rhabdolichops jegui	9013	ANSP 189021	MN031545.1
Rhabdolichops navalha	9030	LFCE 201004210950D_CAB03	MN031546.1
Rhabdolichops troscheli	11853	MUSM 61963	MN031547.1
Rhamphichthys apurensis	10995	ANSP 198380	MN031548.1
Rhamphichthys drepanium	11854	UF 189191	MN031549.1
Rhamphichthys heleios	11855	UF 189422	MN031550.1
Rhamphichthys rostratus	8825	INPA 46	MN031551.1
Rhamphichthys sp. 1	10999	ANSP 198379	MN031552.1
Steatogenys duidae	2147	MCP 31958	MN031553.1
Steatogenys elegans	8807	INPA 28860	MN031554.1
Steatogenys ocellatus	9107	MUSM 44772	MN031555.1
Sternarchella calhamazon	10981	ANSP 193103	MN031556.1
Sternarchella duccis	11844	MUSM 61958	MN031524.1
Sternarchella orthos	2899	MCP 49436	MN031558.1
Sternarchella raptor	2910	NA	MN031525.1
Sternarchella rex	11846	MUSM 54500	MN031557.1
Sternarchogiton labiatus	11848	MUSM 61959	MN031559.1
Sternarchogiton porcinum	10980	ANSP 182319	MN031560.1
Sternarchogiton preto	8732	INPA 28880	MN031561.1
Sternarchorhamphus muelleri	2103	MCP 41658	MN031564.1
Sternarchorhamphus sp. 1	8744	NA	MN031563.1
Sternarchorhynchus galibi	11037	ANSP 187155	MN031566.1
Sternarchorhynchus goeldii	2849	MCP 41643	MN031567.1
Sternarchorhynchus goeldii	8746	INPA 28873	MN031565.1
Sternarchorhynchus goeldii	10988	INPA 40463	MN031569.1
Sternarchorhynchus hagedornae	10969	ANSP 180637	MN031568.1
Sternarchorhynchus mareikeae	11858	ZUEC 15783	MN031570.1
Sternarchorhynchus marreroi	11013	ANSP 198345	MN031571.1
Sternarchorhynchus montanus	11849	MUSM 61960	MN031572.1
Sternarchorhynchus mormyrus	2871	MCP 41640	MN031573.1
Sternarchorhynchus retzeri	11850	MUSM 61961	MN031574.1
Sternarchorhynchus starksi	11076	MCP 47080	MN031575.1

Sternarchorhynchus stewarti	7352	MUSM 33822	MN031576.1
Sternarchorhynchus yepezi	11014	ANSP 198401	MN031577.1
Sternopygus aequilabiatus	2820	IAvHP 7825	MN031578.1
Sternopygus arenatus	9038	MNRJ 734	MN031579.1
Sternopygus astrabes	2203	NA	MN031580.1
Sternopygus branco	2108	MCP 32246	MN031581.1
Sternopygus dariensis	9043	IAvHP-8477	MN031584.1
Sternopygus macrurus	2112	UF 122829	MN031585.1
Sternopygus macrurus	9065	MNRJ 33649	MN031582.1
Sternopygus macrurus	9086	NA	MN031583.1
Sternopygus obtusirostris	2114	MCP 32261	MN031586.1
Sternopygus pejeraton	9089	UNELLEZ	MN031587.1
Sternopygus xingu	11648	NA	MN031588.1

Table S5. Taxa used in vertebrate rhodopsin dataset, with Genbank accession numbers.

Species	accession
Albulichthys albuloides	KC690011.1
Amia calva	AF137208.1
Callorhinchus milii	NM_001292252.1
Chanos chanos	FJ197072.1
Ctenopharyngodon idella	KX224231.1
Danio tinwini	JQ614153.1
Danionella mirifica	FJ531347.1
Gonorynchus greyi	EU409632.1
Gyrinocheilus aymonieri	EU409663.1
Lefua echigonia	FJ197028.1
Lepisosteus oculatus	XM_006630940.2
Microdevario kubotai	JQ614189.1
Microrasbora rubescens	JF966216.1
Minytrema melanops	FJ197034.1
Mugil cephalus	Y18668.1
Neoceratodus forsteri	EF526295.1
Oncorhynchus nerka	AY214156.1
Acipenser fulvescens	KX145850.1
Rasbora borapetensis	HM223982.1
Rasbora pauciperforata	JQ614284.1
Securicula gora	HM224015.1
Sinibotia robusta	JN177209.1
Solea solea	EU638009.1
Tanichthys micagemmae	HM224017.1
Umbra limi	JN230999.1
Vaillantella maassi	FJ197031.1
Zeus faber	EU638023.1
Hepsetus odoe	JX470079.1
Ameiurus nebulosus	KX146011.1
Pygocentrus nattereri	XM_017683920.1
Diplomystes nahuelbutaensis	JN230990.1
Hydrolycus armatus	JX470081.1
Apteronotus albifrons	JN230983.1
Sternarchorhynchus stewarti 7352	MN031576.1
Sternopygus macrurus 2112	MN031585.1
Eigenmannia virescens	KX260614.1
Rhabdolichops jegui 9013	MN031545.1
Electrophorus electricus 9793	MN031488.1
Brachyhypopomus brevirostris 7019	MN038185.1
Hypopygus lepturus 2438	MN031519.1
Rhamphichthys rostratus 8825	MN031551.1
Gymnorhamphichthys hypostomus 11851	MN031493.1
Gymnorhamphichthys rondoni 2154	MN031495.1
Homo sapiens	KJ849294.1
Engraulis mordax	KT201125.1
Monodelphis domestica	XM_001366188.2
Xenopus tropicalis	NM_001097334.2
Lates microlepis	JQ938016.1
Latimeria chalumnae	XM_005997817.2
Mus musculus	XM_017321493.1

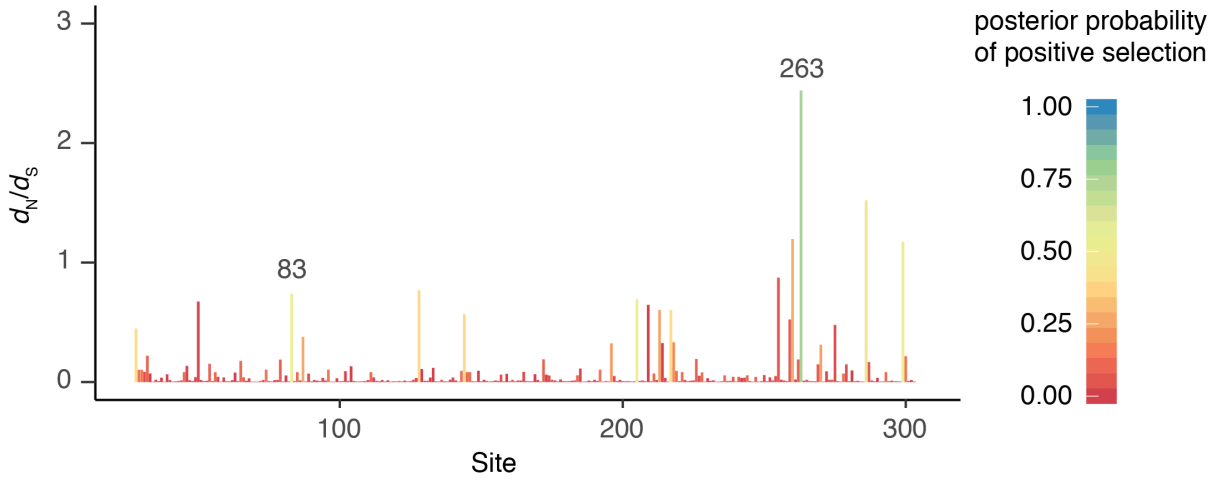

Figure S1. Bar chart showing the d_N/d_S estimates for each site in gymnotiform rhodopsin using FUBAR. Each site is coloured to match the posterior probability of positive selection at that site. Sites with posterior probabilities of positive selection above 0.5 are labelled.


Danio_rerio BC164171.1
Ameiurus_nebulosus JN230991.1
Malapterurus_electricus JX470084.1
Pangasianodon_hypophthalmus JX255590.1
Distichodus_antonii JN230985.1
Hydrolycus_armatus JX470081.1
Phenacogrammus_interruptus FJ197073.1
Sternopygus astrabes 2203
Sternopygus obtusirostris 2114
Sternopygus aequilabiatus 2820
Sternopygus pejeraton 9089
Sternopygus arenatus 9038
Sternopygus dariensis 9043
Sternopygus xingu 11648
Sternopygus branco 2108
Sternopygus macrurus 9065
Sternopygus macrurus 9086
Sternopygus macrurus 2112
Japigny kirschbaum 8992
Distocyclus conirostris 7306
Archolaemus luciae 11857
Eigenmannia sp. 1 10964
Rhabdolichops navalha 9030
Rhabdolichops troscheli 11853
Rhabdolichops jegui 9013
Rhabdolichops electrogrammus 9004
Rhabdolichops electrogrammus 2898
Rhabdolichops eastwardi 8996
Rhabdolichops eastwardi 9014
Rhabdolichops eastwardi 2104
Rhabdolichops caviceps 10990
Rhabdolichops caviceps 2883
Rhabdolichops caviceps 8994
Gymnotus sp. 3 7109
Gymnotus anguillaris 10545
Gymnotus anguillaris 9944
Gymnotus sp. 1 10347
Gymnotus sp. 2 2091
Gymnotus jonasi 2471
Gymnotus stenoleucus 2060
Gymnotus javari 2020
Gymnotus sp. 5 8779
Gymnotus inaequilabiatus 10977
Gymnotus esmeraldas 10865
Gymnotus henni 8189
Gymnotus panamensis 8210
Gymnotus sylvius 7240
Gymnotus bahianus 7244
Gymnotus ucamara 1950
Gymnotus chimarrao 11051
Gymnotus omarorum 7093
Gymnotus obscurus 2018
Gymnotus pantanal 7076
Gymnotus sp. 4 8761
Electrophorus electricus 9793
Electrophorus sp. 1 2026
Orthosternarchus tamandua 2625
Sternarchorhamphus sp. 1 8744
Sternarchorhamphus muelleri 2103
Adontosternarchus clarkae 8673
Adontosternarchus clarkae 2906
Adontosternarchus balaenops 2612
Adontosternarchus nebulosus 8679
Adontosternarchus nebulosus 2892
Platyurosternarchus macrostoma 8726
Apteronotus apurensis 8688
Porotergus duende 2916
Porotergus gymnotus 10957
Apteronotus bonapartii 2914
Porotergus sp. 1 11842
Sternarchogiton prieto 8732
Compsaraia compsa 8720
Compsaraia samueli 11036
Sternarchogiton porcinum 10980
Sternarchogiton labiatus 11848
Melanosternarchus amaru 11843
Sternarchella rex 11846
Sternarchella duccis 11844
Sternarchella orthos 2899
Sternarchella raptor 2910
Sternarchella calhamazon 10981
87
Megadontognathus kaitukaensis 10970
Apteronotus anu 8703
Apteronotus leptorhynchus 8704
Apteronotus mariae 2813
Apteronotus albifrons 7301
Apteronotus eschmeyeri 4001
Apteronotus magdalenensis 4009
Apteronotus galvisi 8700
Apteronotus macrolepis 7110
Sternarchorhynchus goeldii 2849
Sternarchorhynchus goeldii 8746
Sternarchorhynchus goeldii 10988
Sternarchorhynchus hagedornae 10969
Sternarchorhynchus stewarti 7352
Sternarchorhynchus mormyrus 2871
Sternarchorhynchus mareikeae 11858
Sternarchorhynchus marreroi 11013
Sternarchorhynchus galibi 11037
Sternarchorhynchus montanus 11849
Sternarchorhynchus starksi 11076
Sternarchorhynchus retzeri 11850
Sternarchorhynchus yepezi 11014
Steatogenys duidae 2147
Steatogenys elegans 8807
Steatogenys ocellatus 9107
Hypopygus lepturus 2438
Hypopygus minissimus 2336
Hypopygus neblinae 2337
Hypopygus nijsseni 2216
Rhamphichthys heleios 11855
Rhamphichthys sp. 1 10999
Rhamphichthys drepanium 11854
Rhamphichthys apurensis 10995
Rhamphichthys rostratus 8825
Gymnorhamphichthys sp. 1 10965
Gymnorhamphichthys hypostomus 11851
Gymnorhamphichthys rosamariae 12000
Gymnorhamphichthys bogardusae 10974
Gymnorhamphichthys britskii 11635
Gymnorhamphichthys rondoni 10968
Gymnorhamphichthys rondoni 11646
Gymnorhamphichthys rondoni 2154
Akawaio penak 8796
Racenisia fimbriipinna 2340
Microsternarchus sp. 1 11996
Microsternarchus sp. 3 10348
Microsternarchus sp. 2 11998
Microsternarchus bilineatus 2138
Procerusternarchus pixuna 11638
Brachyhypopomus palenque 2433
Brachyhypopomus diazae 305
Brachyhypopomus occidentalis 8780
Brachyhypopomus sullivani 7039
Brachyhypopomus bombilla 9104
Brachyhypopomus regani 7040
Brachyhypopomus provenzanoi 2365
Brachyhypopomus batesi 2414
Brachyhypopomus sp. 1 10344
Brachyhypopomus bennetti 2136
Brachyhypopomus walteri 7048
Brachyhypopomus sp. 5 11997
Brachyhypopomus pinnicaudatus 2121
Brachyhypopomus flavipomus 2141
Brachyhypopomus beebei 6967
Brachyhypopomus belindae 2132
Brachyhypopomus verdii 2254
Brachyhypopomus draco 9101
Brachyhypopomus hamiltoni 7234
Brachyhypopomus alberti 7046
Brachyhypopomus sp. 6 6966
Brachyhypopomus sp. 3 11999
Brachyhypopomus sp. 2 11994
Brachyhypopomus sp. 4 11995
hv
0.1

Figure S2. Rhodopsin gene tree for gymnotiforms generated using PhyML. Branch lengths equal to the number of nucleotide substitutions per site. Tree rooted using *Danio rerio*. Bootstrap values shown to the right of each node. Gymnotiform clades using wave-type and pulse type electric organ discharges (EOD) are indicated in blue and red respectively. The high-voltage EOD generating lineage, the electric eel, is indicated with an (*hv*) [19]. Gray box indicates the clade comprised of deep-channel gymnotiforms exhibiting positive selection of rhodopsin.

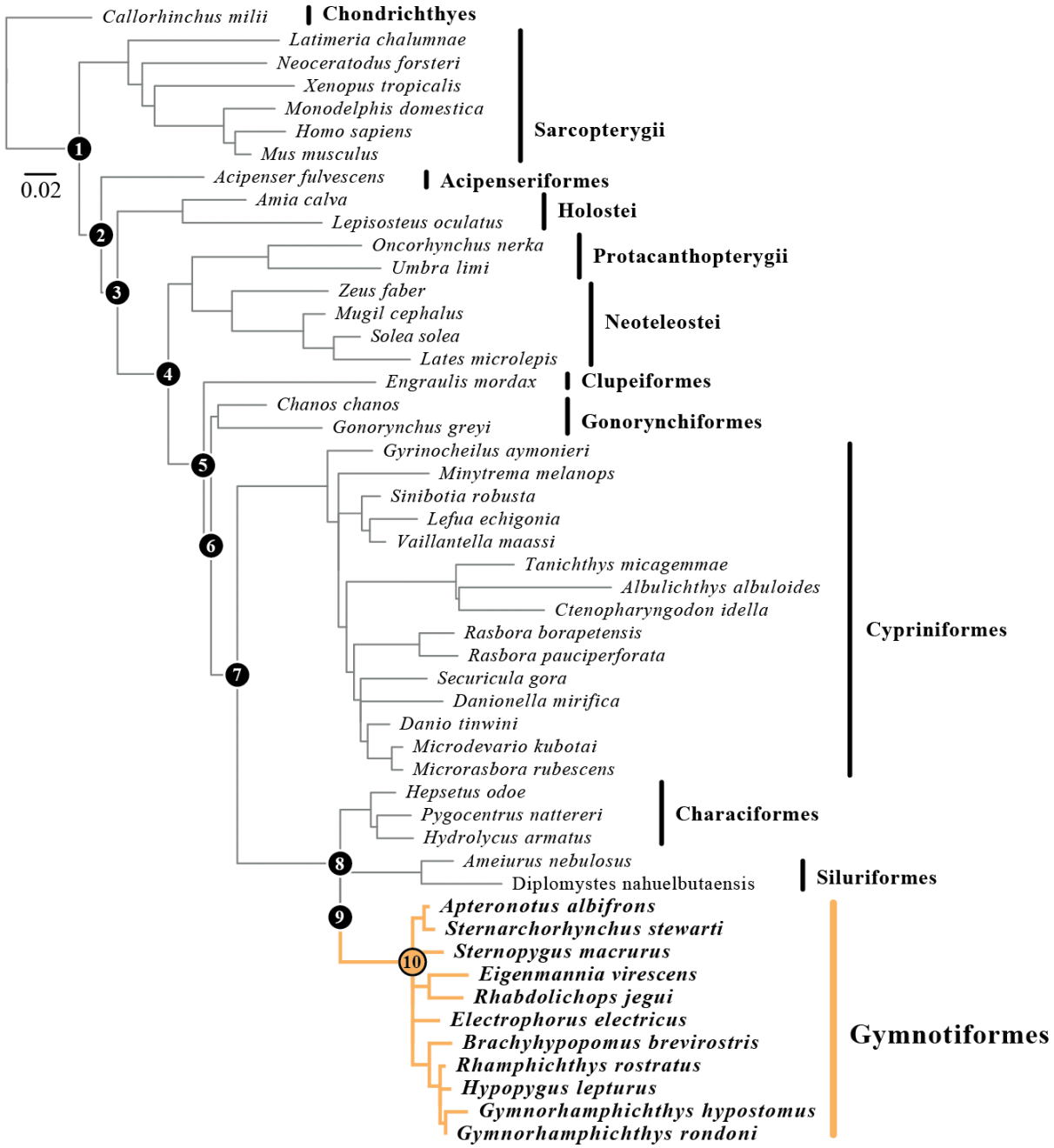

Figure S3. Vertebrate tree used to compare the molecular evolution of rhodopsin in gymnotiforms to other vertebrates and reconstruct ancestral sequences. Topology matches the species relationships in the robust multi-gene species tree generated in [5]. Branch lengths scaled by the number of amino acid substitutions per site using the mtZOA+I+G4 model of amino acid evolution. Nodes enumerated from the common ancestor of fishes and humans to the common ancestor of gymnotiforms (see Table S3).

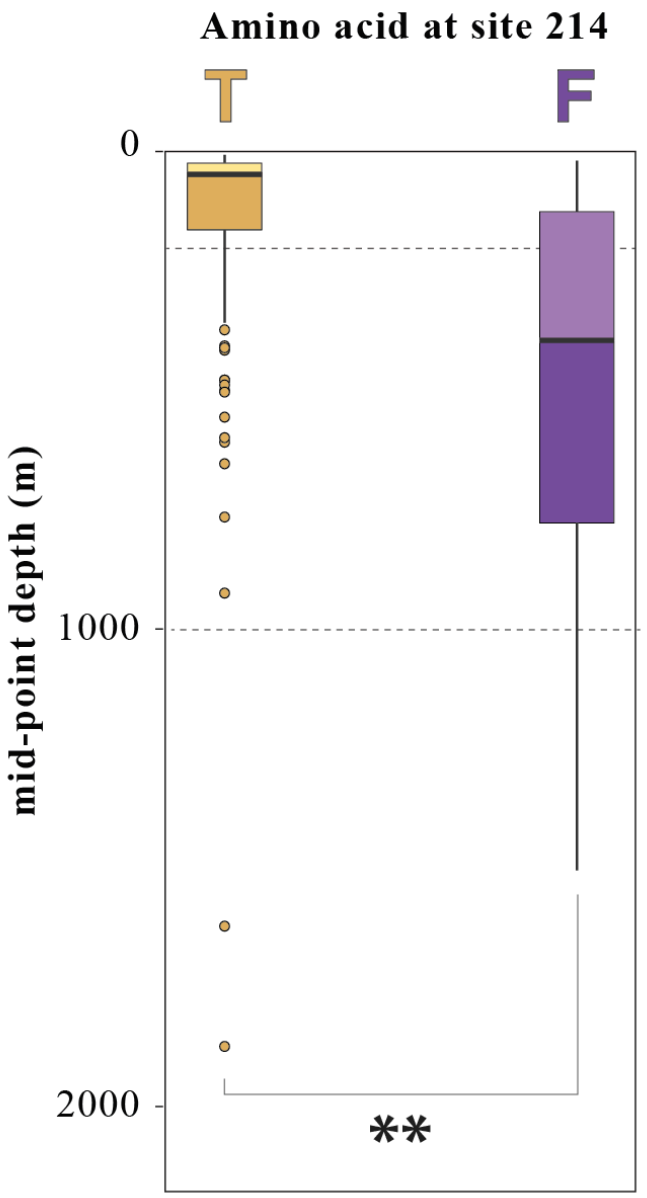

Figure S4. Boxplots showing midpoint depths for fishes with different amino acid identities at site 214 (n =169 species). Fish sequences were obtained from Genbank, and depth data were collected using the R Fishbase package [20]. Mesopelagic zone (200-1000m) indicated by dashed lines. ** indicates significance ($p < 0.001$) for Mann-Whitney U test.

Supplementary References

1. Chen, W. J., Lavoué, S. & Mayden, R. L. 2013 Evolutionary origin and early biogeography of otophysan fishes (Ostariophysi: Teleostei). *Evolution* **67**, 2218–2239.

2. Wright, E. S. 2015 DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinform.* **16**, 322.

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

4. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Haeseler, von, A. & Jermiin, L. S. 2017 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587.

5. Schott, R. K., Gow, D. A. & Chang, B. S. W. 2019 BlastPhyMe: A toolkit for rapid generation and analysis of protein-coding sequence datasets. *bioRxiv*, 059881. (doi:10.1101/059881)

6. Betancur-R, R., Wiley, E. O., Arratia, G., Acero, A., Bailly, N., Miya, M., Lecointre, G. & Orti, G. 2017 Phylogenetic classification of bony fishes. *BMC Evol. Biol.* **17**, 162.

7. Yang, Z., Nielsen, R., Goldman, N. & Pedersen, A.-M. K. 2000 Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**, 431–449.

8. Yang, Z. & Nielsen, R. 2002 Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol. Biol. Evol.* **19**, 908–917.

9. Bielawski, J. P. & 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.

10. Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L. & Scheffler, K. 2014 RELAX: detecting relaxed selection in a phylogenetic framework. *Mol. Biol. Evol.* **32**, 820–832.

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

12. Pond, S. L. K. & Muse, S. V. 2005 HyPhy: hypothesis testing using phylogenies. In *Statistical methods in molecular evolution*, pp. 125–181. Springer.

13. Kosakovsky Pond, S. L., Murrell, B., Fourment, M., Frost, S. D., Delport, W. & Scheffler, K. 2011 A random effects branch-site model for detecting episodic diversifying selection. *Mol. Biol. Evol.* **28**, 3033–3043.

14. Pond, S. L. K. & Frost, S. D. 2005 Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* **21**, 2531–2533.

15. Yang, Z., Wong, W. S. & Nielsen, R. 2005 Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* **22**, 1107–1118.

16. Yang, Z. 2007 PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591.

17. Choe, H.-W., Kim, Y. J., Park, J. H., Morizumi, T., Pai, E. F., Krauss, N., Hofmann, K. P., Scheerer, P. & Ernst, O. P. 2011 Crystal structure of metarhodopsin II. *Nature* **471**, 651.

18. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. 2004 UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.

19. Crampton, W. G. R. (2019). Electroreception, electrogenesis and electric signal evolution. *J. Fish Biol.* ***204***(12), 185.

20. Boettiger, C., Lang, D. T. & Wainwright, P. C. 2012 rfishbase: exploring, manipulating and visualizing FishBase data from R. *J. Fish Biol.* **81**, 2030–2039.