Supporting Figures and Tables for:

The molecular origin and evolution of dim-light vision in mammals

Constanze Bickelmann, James M. Morrow, Jing Du, Ryan K. Schott, Ilke van Hazel, Steve Lim, Johannes Müller, and Belinda S.W. Chang

Contents:

Table S1. Rhodopsin accession numbers of 27 taxa used in this study, of which 26 were obtained from NCBI Genbank.

Species name	Common name	Accession numbers
Alligator mississippiensis	American alligator	U23802.1
Ambystoma tigrinum	Tiger salamander	U36574.1
Anolis carolinensis	Green anole	L31503.1
Bos taurus	Cattle	NM_001014890.1
Bufo bufo	European toad	U59921.1
Caluromys philander	Fat-tailed dunnart	AY159786.2
Canis lupus familiaris	Dog	NM_001008276.1
Cavia porcellus	Guinea pig	EF457995
Cricetulus griseus	Chinese hamster	X61084.1
Felis catus	Domestic cat	NM_001009242.1
Gallus gallus domesticus	Chicken	NM_001030606.1
Homo sapiens	Human	NM_000539.2
Latimeria chalumnae	Coelacanth	AF131256.1
Loxodonta africana	African elephant	AY686752.1
Macaca fascicularis	Rhesus macaque	XM_001094250.1
Neoceratodus forsteri	Australian lungfish	EF526295
Ornithorhynchus anatinus	Platypus	EF050076.1
Oryctolagus cuniculus	European rabbit	NM_001082349.1
Otolemur crassicaudata	Galago	AB112594.2
Rana temporaria	European common frog	U59920.1
Sminthopsis crassicaudata	Bare-tailed woolly opossum	AY313946.1
Sus scrofa	Wild boar	NM_214221.1
Tachyglossus aculeatus	Short-beaked echidna	JX103830
Thamnophis proximus	Garter snake	Yang (2010)
Trichechus manatus	West-Indian manatee	AF055319.1
Uta stansburiana	Common side-blotched lizard	DQ100323.1
Ursus maritimus	Polar bear	AY883926.1

Table S2. Comparison of random site codon models used for reconstruction of the amniote, mammalian, and therian ancestral rhodopsins. M3 and M7 were the best-fitting models and resulted in identical ancestral sequences.[1] 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 and 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 M8.

Abbreviations—**np**, number of parameters; ***ln*L**, *ln* Likelihood; ***K***, transition/transversion rate ratio; **LRT**, likelihood ratio test statistic; **df**, degrees of freedom; ***P***, *P*-value; **n/a**, not applicable.

Model	np	*ln*L	*K*	Parameters[1]			Null	LRT	df	*P*
				ω_0/p	ω_1/q	ω_2/ω_p				
M0	53	-10643.20	2.32	0.054	-	-	n/a	-	-	-
M1a	54	-10529.53	2.37	0.045 (95.4%)	1 (4.6%)	-	M0	227.35	2	**0.000**
M2a	56	-10529.53	2.37	0.045 (95.4%)	1 (0.4%)	1 (4.2%)	M1a	0.00	2	1.000
M3	57	-10359.83	2.30	0.002 (58.3%)	0.081 (29.5%)	0.289 (12.2%)	M0	316.02	4	**0.000**
M7	54	-10361.14	2.29	0.240	3.389	-	n/a	-	-	-
M8	56	-10367.91	2.29	0.250	2.933	1 (0.9%)	M8a	-13.54	2	n/a

Table S3. Specific primers used in site-directed mutagenesis PCR.

Primer	Sequence (5'-3')
V39A-F	GTA TAG CGC CCT GGC CGC CTA CAT GTT C
V39A-R	TGC CAG GGC TCG GCC AGG TAG
V49L-F	TAC ATG TTC ATG CTG ATC CTG CTG GGC TTC
V49L-R	GGC GGC CAG CAC GCT ATA CTG

Table S4. List of amino acid sites reconstructed with low posterior probability (under 80%) with the M3 model. The identity of the sites found in bovine and echidna are shown for reference along with the best and next best sites reconstructed for each ancestor. Abbreviations—**AA**, amino acid; **PP**, posterior probability.

Site	Bovine	Echidna	Amniota				Mammalia				Theria			
			Best		Next Best		Best		Next Best		Best		Next Best	
	AA	AA	AA	PP (%)	AA	PP (%)	AA	PP (%)	AA	PP (%)	AA	PP (%)	AA	PP (%)
11	V	I	V	57.8	I	41.6	V	63.4	I	36.1	-	-	-	-
39	M	V	-	-	-	-	V	68.1	A	24.9	V	48.9	F/I/A	13.1/12.9/8.7
46	L	M	M	63.1	L	36	-	-	-	-	-	-	-	-
49	M	M	-	-	-	-	V	57	M/L	24.7/13.6	-	-	-	-
88	F	L	L	51.9	F	48.1	F	52.3	L	47.3	-	-	-	-
107	P	P	P	70.2	A/L/V	13.1/10.9/4.1	-	-	-	-	-	-	-	-
111	N	N	N	75.5	Y	24	-	-	-	-	-	-	-	-
137	V	I	-	-	-	-	V	54.9	I	42.5	V	52.3	I	44.3
162	V	I	-	-	-	-	-	-	-	-	I	65	V	35
217	I	T	T	73.3	I	25.6	T	72.3	I	27.2	I	59.5	T	38.4
218	V	I	I	64.9	V	34.9	I	55.7	V	43.2	-	-	-	-
286	I	I	I	61.8	T	38.2	-	-	-	-	-	-	-	-
290	I	V	V	59.5	I	39.3	V	58.2	L/I	23/16.1	-	-	-	-
299	A	A	-	-	-	-	-	-	-	-	A	76.9	S	22.2

Table S5. Tests of positive selection on the amniote, mammalian, and therian branches using the branch-site model.[1] 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 (either the amniote, mammal, or therian branch).

Abbreviations—**np**, number of parameters; ***ln*L**, ln Likelihood; ***K***, transition/transversion ratio; **LRT**, likelihood ratio test statistic; **df**, degrees of freedom; ***P***, *P*-value; **n/a**, not applicable.

Foreground Branch	Model	np	*ln*L	*K*	Parameters[1]				Null	LRT	df	*P*
					ω_0	ω_1	ω_{2a}	ω_{2b}				
Amniota	BrS_Null	55	-10561.23	2.36	B: 0.045 F: 0.045 (60.4%)	B: 1 F: 1 (2.6%)	B: 0.045 F: 1 (35.5%)	B: 1 F: 1 (1.5%)	n/a	-	-	-
	BrS_Alt	56	-10561.05	2.36	B: 0.045 F: 0.045 (89.7%)	B: 1 F: 1 (3.9%)	B: 0.045 F: 11.41 (6.2%)	B:1 F:11.4 (0.3%)	BrS_Null	0.36	1	0.546
Mammalia	BrS_Null	55	-11238.69	2.16	B: 0.036 F: 0.036 (94.4%)	B: 1 F: 1 (5.6%)	B: 0.036 F: 1 (0%)	B:1 F: 1 (0%)	n/a	-	-	-
	BrS_Alt	56	-11238.69	2.16	B: 0.036 F: 0.036 (94.4%)	B: 1 F: 1 (5.6%)	B: 0.036 F: 1 (0%)	B:1 F: 1 (0%)	BrS_Null	0	1	1
Theria	BrS_Null	55	-10552.41	2.36	B: 0.044 F: 0.044 (82%)	B: 1 F: 1 (3.9%)	B: 0.044 F: 1 (13.5%)	B: 1 F: 1 (0.6%)	n/a	-	-	-
	BrS_Alt	56	-10543.17	2.39	B: 0.044 F: 0.044 (93.1%)	B: 1 F: 1 (4.4%)	B: 0.044 F: 999 (2.4%)	B:1 F: 999 (0.1%)	BrS_Null	18.48	1	**0.000**

Table S6. List of codon sites found to be under positive selection along the therian branch by Bayes Empirical Bayes (BEB) analysis using the branch-site model of PAML.

Codon Site	Posterior Probability
13	0.997
37	0.967
225	0.993
290	0.922
345	1
346	1
348	0.972

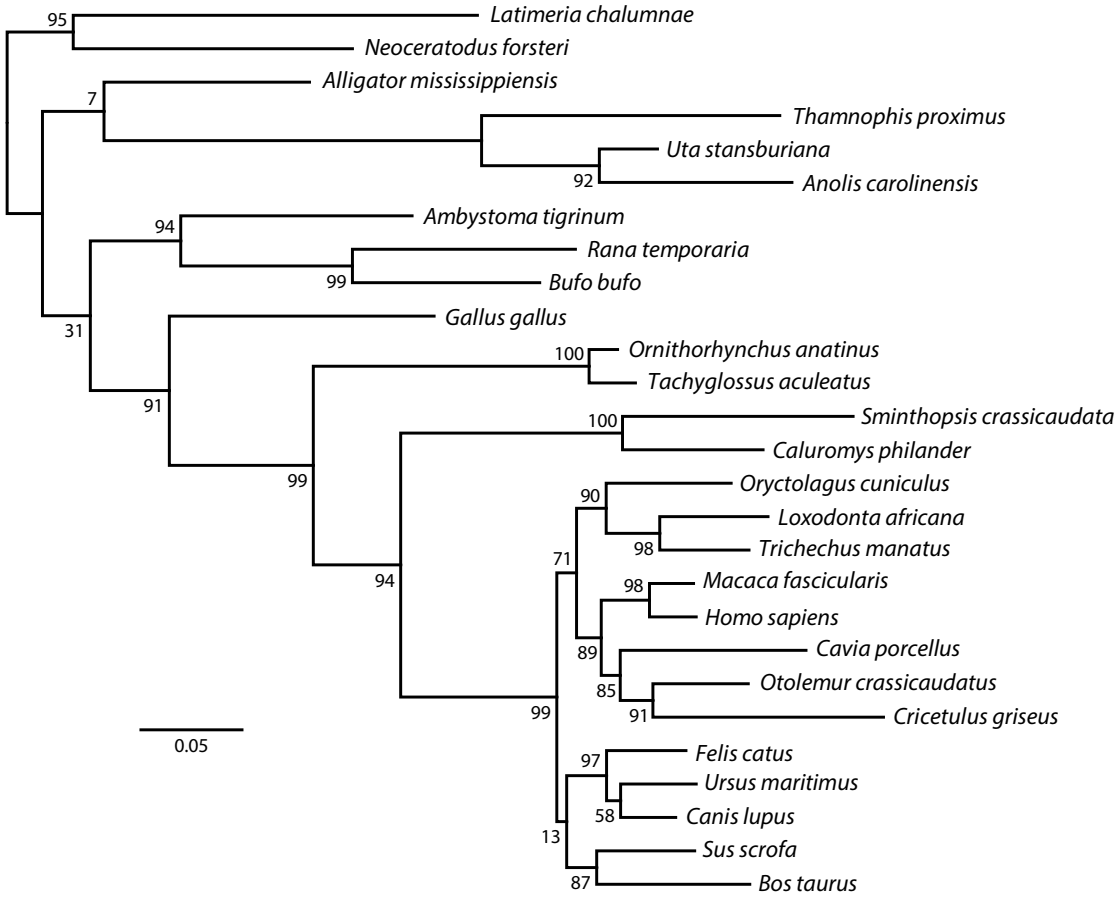

Figure S1. Rhodopsin gene tree inferred by maximum likelihood (PhyML) with aLRT support values.

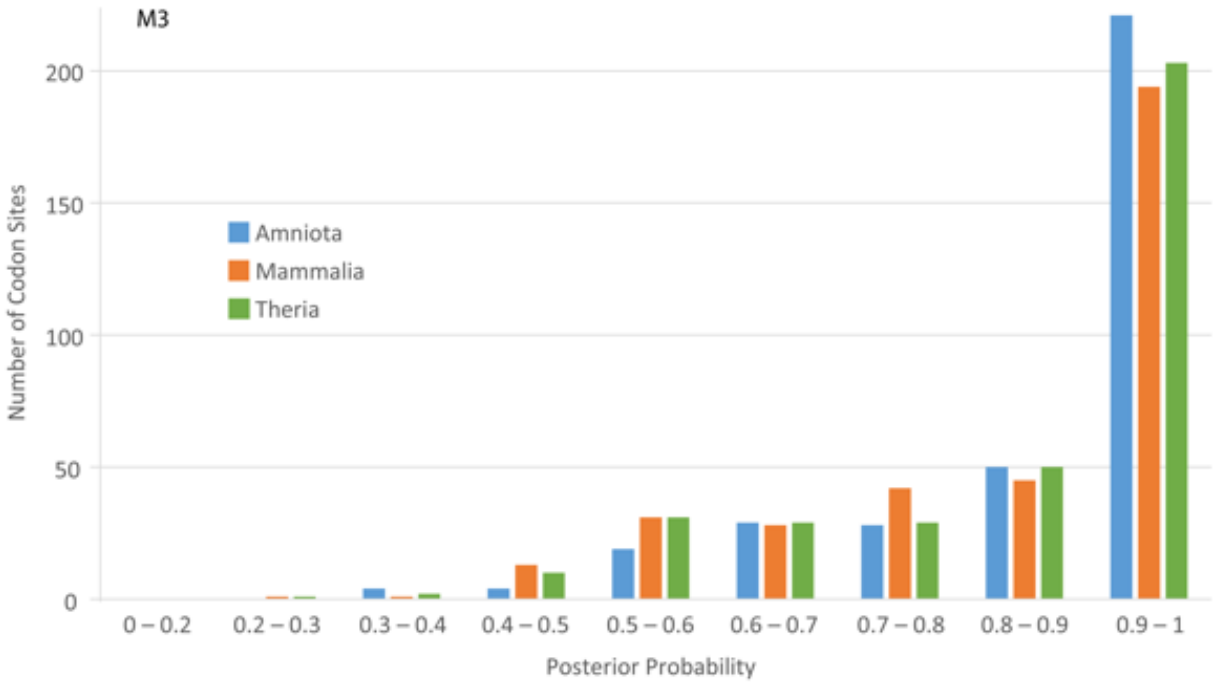

Figure S2. Histogramm of distribution of posterior probabilities for codon sites reconstructed under the M3 model for the amniote, mammalia, and therian ancestral nodes. See also Table S4 for a list of amino sites reconstructed with low posterior probilities for each of the three ancestral proteins.