

ENDEMIC, ENDANGERED, AND EVOLUTIONARILY SIGNIFICANT: CRYPTIC LINEAGES IN SEYCHELLES' FROGS (ANURA: SOOGLOSSIDAE).

SUPPORTING INFORMATION

PCR cycling conditions & sequence data

Sequences from two mitochondrial (mtDNA) and four nuclear (nuDNA) loci were amplified via standard polymerase-chain reaction (PCR) with total reaction volumes of 10-42 μ l. Due to difficulty obtaining adequate DNA yields from such small biological samples (toe-clips from frogs regularly less than 10 mm SVL) volumes of template DNA varied between some reactions. For a 25 μ l reaction, reaction volumes consisted of 10.5 μ l ddH₂O, 0.5 μ l each of forward and reverse primer (at a concentration of 25 pmol/ μ l), 12.5 μ l MyTaq HS Red mix™, and 1 μ l of template DNA. Details of primers used are shown in Table S1. Primer pairs developed for this study were generated using Primer-Blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast>). PCR cycling conditions were: denature at 95°C for 60 seconds (*16s*, *cytb*, *rag2*) or 94°C for 60 seconds (*pomc*, *rag1*, *rho*); followed by 35 (*16s*, *cytb*, *rag2*, *rho*) or 40 (*rag1*, *pomc*) cycles of denaturing at 95°C for 15 seconds (*16s*, *cytb*, *rag2*), or 94°C for 30 seconds (*pomc*, *rag1*, *rho*); annealing for 15 seconds at 53°C (*16s*, *cytb*), 59.5°C (*rag2*), or for 30 seconds at 56°C (*rag1*), 57°C (*pomc*), 60°C (*rho*); extending at 72°C for 10 seconds (*16s*, *cytb*), or 30 seconds (*pomc*, *rag1*, *rag2*, *rho*), with a final extension step of 72°C for 5 minutes. All *16s* samples were sequenced in both directions. Due to project constraints complimentary sequence data were not generated for all loci. Those obtained comprised the following: *cytb* = 17; *pomc* = 9; *rag1* = 2; *rag2* = 8; *rho* = 4. All sequences were cross-checked using the BLAST function in MEGA6 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013) and compared against sequences generated by this study. Ambiguous bases were coded accordingly.

Table S1.1 Primers used for PCR amplification and sequencing.

Gene fragment	Primer	Sequence (5' – 3')
<i>16s</i>	16sar-L ^a	CGC CTG TTT ATC AAA AAC AT
	16sbr-H ^a	CCG GTC TGA ACT CAG ATC ACG T
<i>cytb</i>	CBJ 10933 ^b	TAT GTT CTA CCA TGA GGA CAA ATA TC
	Cytb-c ^b	CTA CTG GTT GTC CTC CGA TTC ATG T
	CytbJL1f ^c	TAG ACC TCC CAA CCC CAT CC
	CytbJL1r ^c	GAG GTG TGT GTT AGT GGG GG
	CytbSGJL1f ^c	ACC GCT TTC GTA GGC TAT GT
	CytbSGJL1 ^c	GTG GAC GAA ATG ATA TTG CTC GT
<i>pomc</i>	POMCJLf ^c	GAC ATC GCC AAC TAT CCG GT
	POMCJLr ^c	AAG TGT TGT CCC CCG TGT TT
	POMCJL2f ^c	AAA CAC GGG GGA CAA CAC TT
	POMCJL2r ^c	CTT CTG AGT CGA CAC CAG GG
<i>rag1</i>	RAG1B ^d	ATG GGA GAT GTG AGT GAR AAR CA
	RAG1E ^d	TCC GCT GCA TTT CCR ATG TCR CA
<i>rag2</i>	RAG2 JG1-F ^c	TCG TCC TAC CAT GTT CAC CAA TGA GT
	RAG2 JG1-R ^c	TCC TGT CCA ATC AGG CAG TTC CA
	RAG2JLSG1f ^c	CCA GCA GTG ACC AGC ATC TT
	RAG2JLSG1r ^c	CGC TGT CTC TTG GAC TGG TT
	RAG2JLSG2r ^c	CCG ACA ATG AGG AAC TCG CT
<i>rho</i>	Rhod1A ^e	ACC ATG AAC GGA ACA GAA GGY CC
	Rhod1D ^e	GTA GCG AAG AAR CCT TCA AMG TA

^a Palumbi *et al.*, (1991)^b Chiari *et al.*, (2004)^c Developed for this study^d Biju & Bossuyt, (2003)^e Bossuyt & Milinkovitch, (2000)

Table S1.2 GenBank derived and generated sequence data. Codes indicate GenBank accession numbers. Identical codes in adjacent columns for *Ascaphus truei* and *Leiopelma archeyi* represent sampling of independent sections of the mitochondrial genome of the same accessioned data. A dash (–) indicates sequence data not utilised or available/obtained. A star (*) indicates sequence data generated by this study.

Species	16s	cytb	pomc	rag1	rag2	rho
<i>Ascaphus truei</i>	AJ871087	AJ871087	–	–	–	–
<i>Leiopelma archeyi</i>	NC_014691	NC_014691	–	–	–	–
<i>Sechellophryne gardineri</i> (Mahé)	MK058722-36*	MK058749-52*	MK058753-9*	MK058761-70*	MK058781-95*	MK058806-16*
<i>Sechellophryne gardineri</i> (Silhouette)	MK058737-48*	–	MK058760*	MK072763-5*	MK058796-805*	MK058817-9*
<i>Sechellophryne pipilodryas</i> (Silhouette)	DQ872918 MK058820*	MK058821*	MK058822*	DQ872922	DQ872912 MK058823*	MK058825*
<i>Sooglossus sechellensis</i> (Mahé)	JF784361-3 JF784370-2 MK058826-95*	MK058964-79*	MK059019-28*	MK059049-67*	MK059108-66*	MK059238-55*
<i>Sooglossus sechellensis</i> (Praslin)	JF784364-8 JF784373-4 JF784376-83 MK058896-942*	MK058996-9006*	MK059029-39*	MK059068-92*	MK059167-217*	MK059256-74*
<i>Sooglossus sechellensis</i> (Silhouette)	MK058943-63*	MK059007-18*	MK059040-8*	MK059093-107*	MK059218-37*	MK059275-87*
<i>Sooglossus thomasseti</i> (Mahé)	MK059288-304*	MK059317-23*	MK059326-31*	MK059334-44*	MK059355-69*	MK059382-88*

Species	16s	cytb	pomc	rag1	rag2	rho
<i>Sooglossus thomasseti</i> (Silhouette)	MK059305-16*	MK059324-5*	MK059332-3*	MK059345-54*	MK059370- 81*	MK059389-90*

Table S1.3 Partitioning schemes and substitution models selected by PartitionFinder v1.1.1 (Lanfear *et al.*, 2012) using the AIC criterion for Bayesian (BEAST2/*BEAST) analyses. Codon positions in parentheses.

	Partitioning scheme	Substitution model
mtDNA	<i>16s</i> , <i>cytb</i> (1)	GTR+I+G
	<i>cytb</i> (2)	TrN+I
	<i>cytb</i> (3)	TrN+G
nuDNA	<i>pomc</i> (1-3)	TrN+I+G
	<i>rag1</i> (1-3)	TrN+I+G
	<i>rag2</i> (1-3)	TrN+I+G
	<i>rho</i> (1-3)	TrN+I+G

Table S1.4 Taxa used as composites in *BEAST analyses.

Species	Ref.	Locus	Composite
<i>Sechellophryne gardineri</i>	JMSG07	<i>rho</i>	JMSG09
<i>Sechellophryne pipilodryas</i>	DQ872922	<i>rag1</i>	JMSP01

Table S1.5 Species/population boundaries inferred from Bayesian Poisson Tree Processes (bPTP) analysis. The BEAST2 mtDNA phylogeny was used as the input tree. Posterior probabilities (PP) of maximum likelihood and Bayesian analyses were identical. Populations are listed in node order as per the phylogeny (Fig. 2 in the main text).

Species/population	Island	Sample reference	PP
<i>Sechellophryne pipilodryas</i>	Silhouette	JMSP01	1.00
<i>Sechellophryne gardineri</i>	Mahé	CDSG01, MBSG04, LRSG03, MBSG02	0.97
<i>Sechellophryne gardineri</i>	Silhouette	DGSG01, JMSG01, JMSG05, JMSG07, JMSG10	0.75
<i>Sooglossus thomasseti</i>	Silhouette	GBST01, JMST06	0.99
<i>Sooglossus thomasseti</i>	Mahé	CDST02, CRST01, MCST02, LMST01, MSST01, CDST01, MBST01	0.98
<i>Sooglossus sechellensis</i>	Mahé	LRSS14	0.95
<i>Sooglossus sechellensis</i>	Mahé	LRSS01, MSSS02	0.95
<i>Sooglossus sechellensis</i>	Mahé	RSS01, LRSS02, LMSS01, SFSS02, CRSS14, CSS01, MSSS01, MCSS10, MBSS07, CRSS02, CRSS01, MBSS01	0.97
<i>Sooglossus sechellensis</i>	Praslin	VMSP16, CMSP01, CMSP07, CMSP02, FPSP03, FAT2, ZSP01, ZSP04, ZSP03, ZSP08	0.98
<i>Sooglossus sechellensis</i>	Silhouette	JMSS05, GBSS01, JMSS11, JMSS08, GBSS07, JMSS01, JMSS04, JMSS03, JMSS07, GBSS10, JMSS06, JMSS09	0.98

Table S1.6 Population demographic tests for the Sooglossidae. Positive values of Tajima's D and Fu's F_S indicate stable population structure, balancing selection or recent population decrease; negative values indicate positive selection, or suggest evidence of recent population expansion. Tajima's D and R_2 are interpreted as significant at $P < 0.05$, Fu's F_S at $P < 0.02$.

	<i>Sooglossus sechellensis</i>			<i>Sooglossus thomasseti</i>			<i>Sechellophryne gardineri</i>		
	Tajima's D	Fu's F_S	R_2	Tajima's D	Fu's F_S	R_2	Tajima's D	Fu's F_S	R_2
<i>16s</i>	2.36870	3.278	0.1621	3.10581	12.422	0.2512	2.24857	3.284	0.2073
	$P < 0.05$	$P > 0.02$	$P > 0.05$	$P < 0.01$	$P > 0.02$	$P > 0.05$	$P < 0.05$	$P > 0.02$	$P > 0.05$
<i>cytb</i>	1.71837	-1.421	0.1910	0.44661	3.394	0.1967	-0.55827	-0.361	0.1061
	$P > 0.05$	$P > 0.02$	$P > 0.05$	$P > 0.05$	$P > 0.02$	$P > 0.05$	$P > 0.05$	$P > 0.02$	$P < 0.05$
<i>pomc</i>	-1.40180	-8.378	0.0560	-1.21781	-1.557	0.0963	-0.64112	-10.089	0.1300
	$P > 0.05$	$P < 0.02$	$P > 0.05$	$P > 0.05$	$P > 0.02$	$P < 0.01$	$P > 0.05$	$P < 0.02$	$P > 0.05$
<i>rag1</i>	-1.21313	-7.542	0.0534	0.21337	0.346	0.1401	-0.13712	-0.421	0.1331
	$P > 0.05$	$P < 0.02$	$P > 0.05$	$P > 0.05$	$P > 0.02$	$P > 0.05$	$P > 0.05$	$P > 0.02$	$P > 0.05$
<i>rag2</i>	-1.77022	-82.555	0.0335	-0.12593	-1.420	0.1044	-0.65881	-8.246	0.0910
	$P < 0.05$	$P < 0.02$	$P < 0.05$	$P > 0.05$	$P > 0.02$	$P > 0.05$	$P > 0.05$	$P = 0.02$	$P > 0.05$
<i>rho</i>	-1.37952	-1.467	0.1000	0.65931	-0.801	0.1846	0.89497	-0.346	0.1582
	$P > 0.05$	$P > 0.02$	$P > 0.05$	$P > 0.05$	$P > 0.05$	$P > 0.05$	$P > 0.05$	$P > 0.02$	$P > 0.05$

Table S1.7 Extended Bayesian Skyline Plot (EBSP) results for sooglossid populations. Results are the 95% highest posterior density (HPD) interval for population size changes from all loci in a combined analyses. Constant population size cannot be rejected if the 95% HPD interval includes 0. Plus sign (+) indicates population expansion. Low sample sizes can lead to unreliable EBSP results (Heller & Siegismund, 2013) and consistent ESS values were not obtained for the Silhouette population of *Se. gardineri* until we removed underrepresented loci (*pomc*, *rag1*, *rho*).

Island	<i>Sooglossus sechellensis</i>			<i>Sooglossus thomasseti</i>		<i>Sechellophryne gardineri</i>	
	Mahé	Praslin	Silhouette	Mahé	Silhouette	Mahé	Silhouette
Chain length	3×10^8	2×10^8	5×10^7	1×10^8	7.5×10^7	5×10^7	5×10^7
EBSP	[0, 3]	[1, 3] ⁺	[0, 3]	[0, 3]	[0, 3]	[0, 3]	[0, 2]

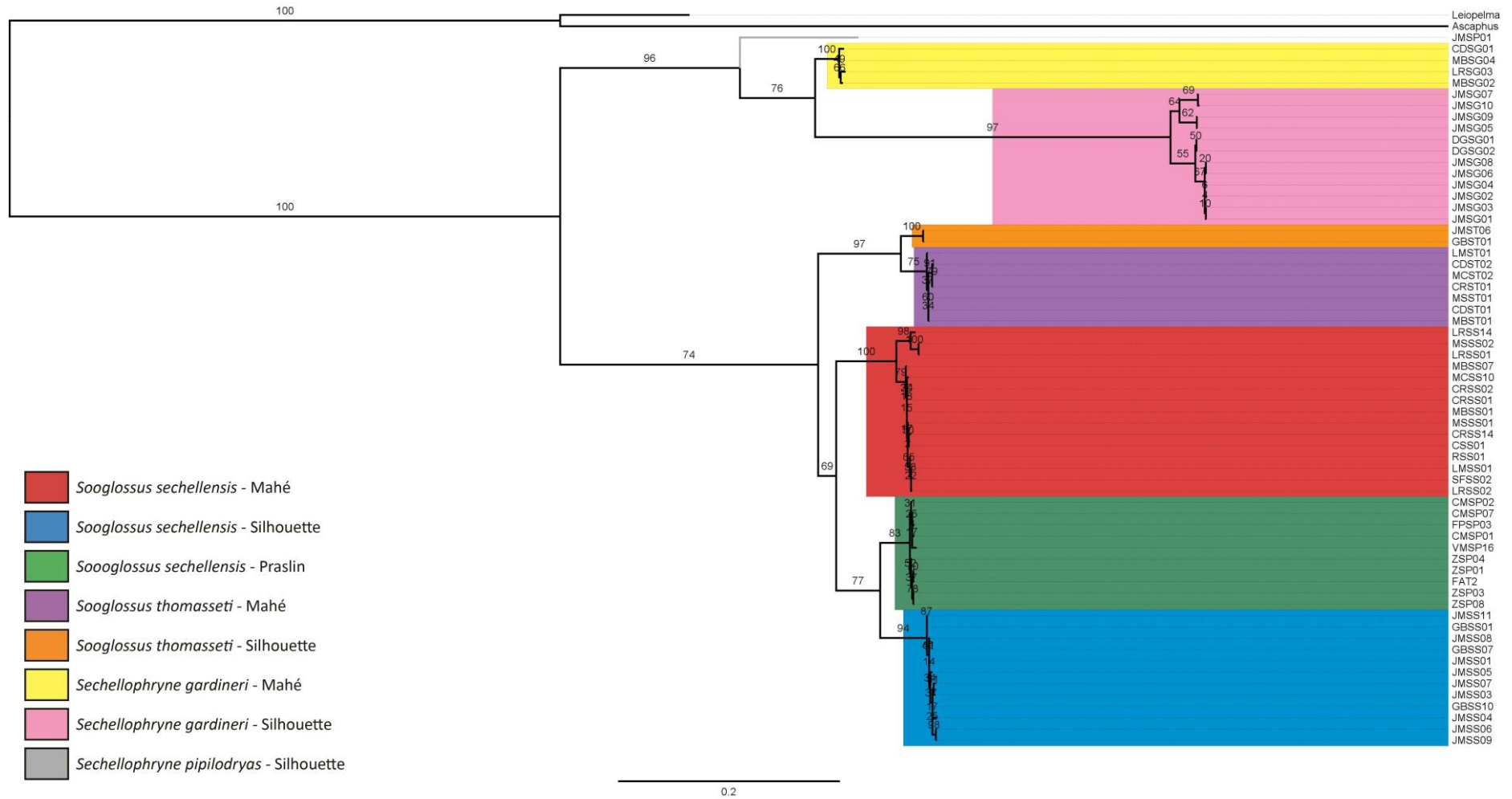


Figure S1.1 Maximum likelihood inferred mitochondrial DNA phylogeny of the Sooglossidae. Leiopelema (Leiopelema+Ascaphus) rooted outgroup. Branch support is indicated by maximum likelihood bootstrap (BS) values. Scale bar indicates substitutions per site.

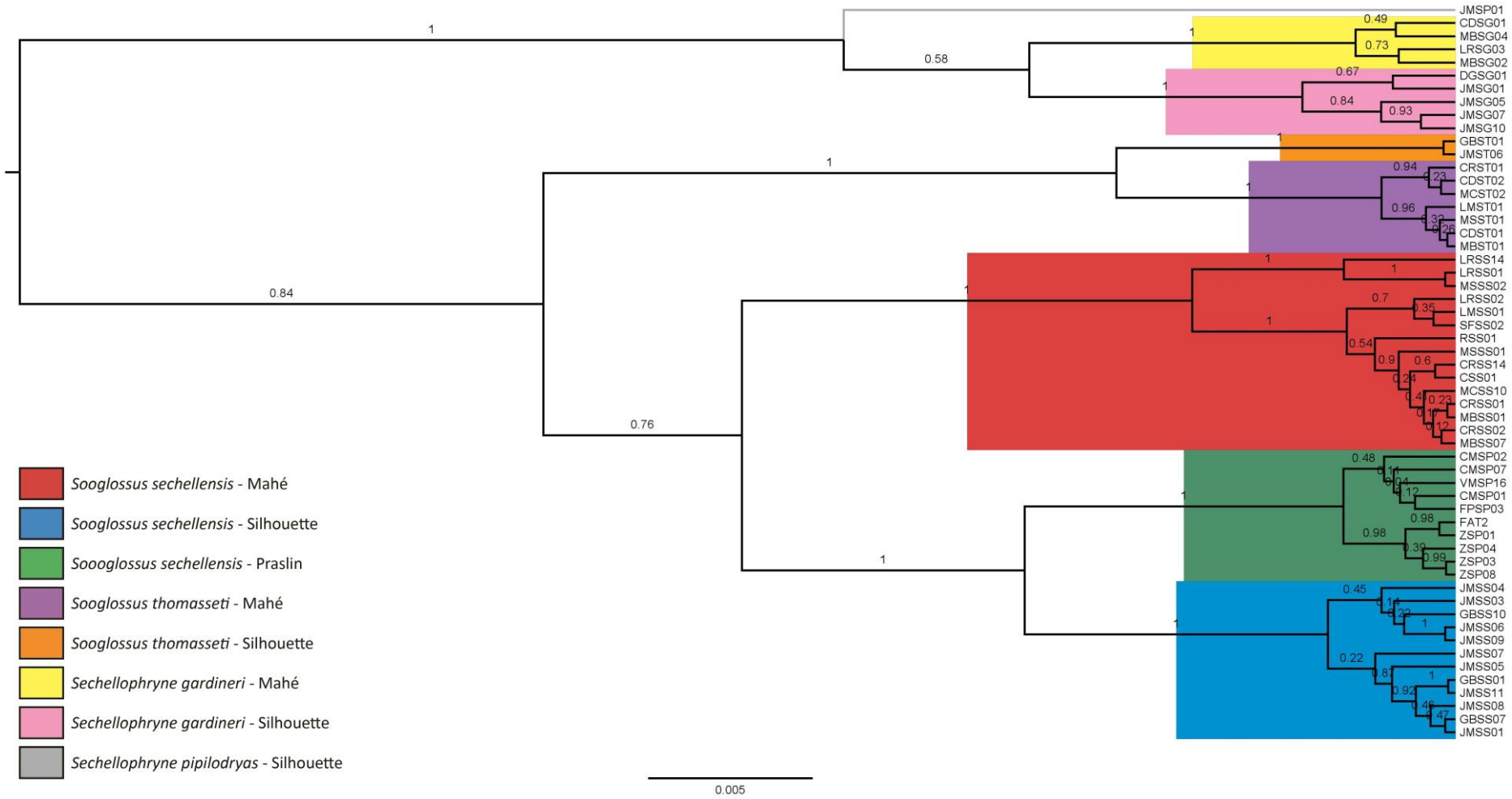


Figure S1.2 Bayesian inferred mitochondrial DNA phylogeny of the Sooglossidae using the Yule tree prior in BEAST2. Branch support is indicated by Bayesian Posterior Probabilities (PP). Scale bar indicates substitutions per site.

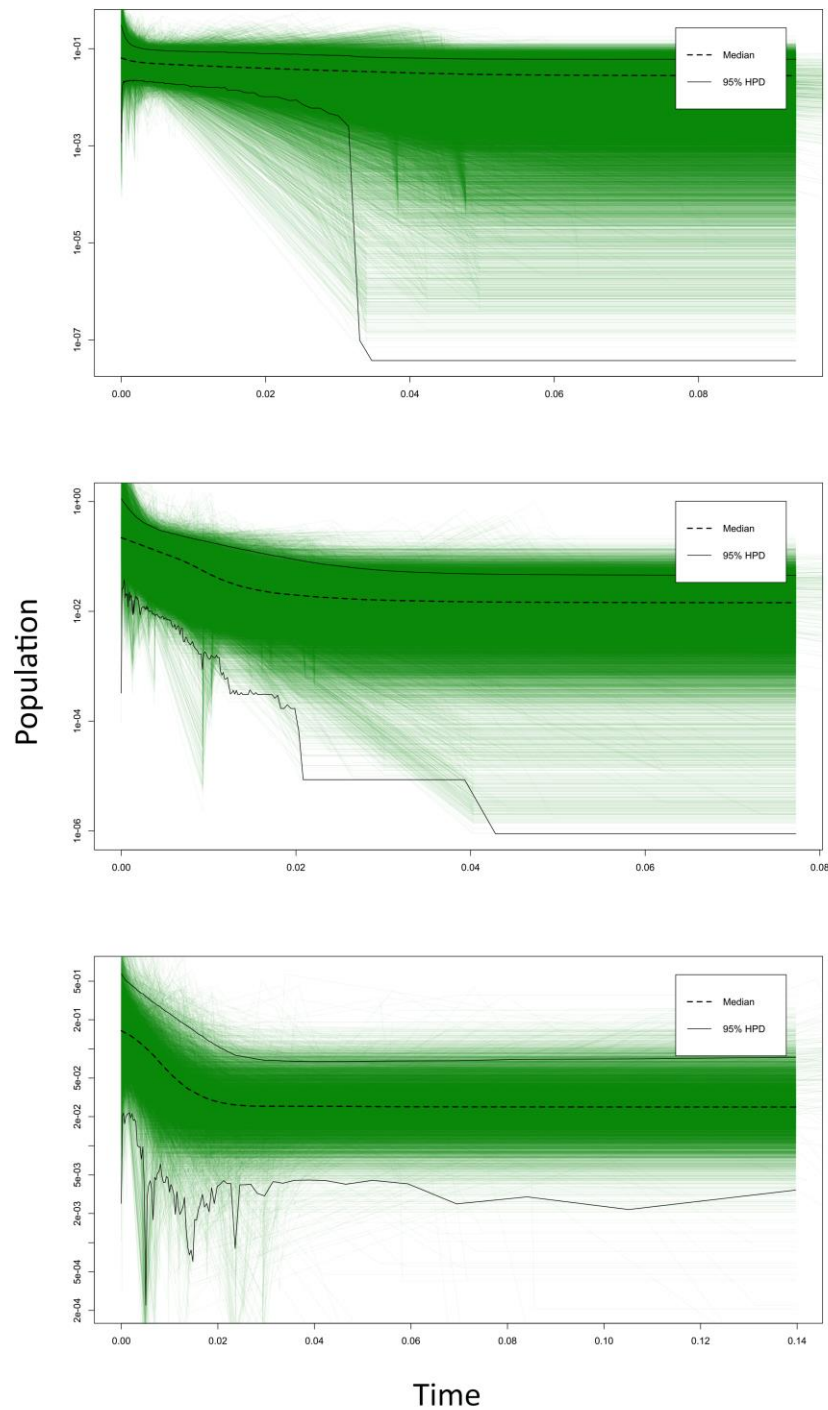


Figure S1.3 Extended Bayesian Skyline Plots of population size through time for *Sooglossus sechellensis*. The full view of the posterior all of the samples that are summarised by the median and 95% HPD interval are shown for the Mahé (top), Praslin (centre), and Silhouette (bottom) populations. The Praslin frogs are the only sooglossid population to reject a constant population size. EBS analyses comprised all six loci. Time on x-axis in millions of years. Population on the y-axis represents the effective population size assuming a generation time of 1 year.

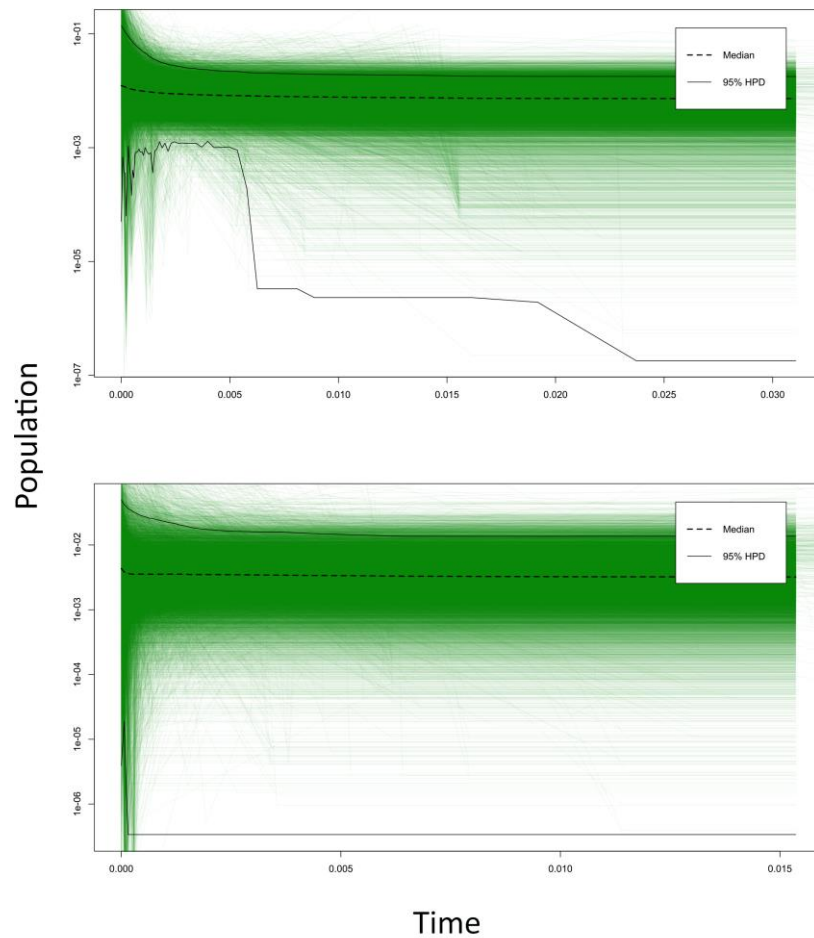


Figure S1.4 Extended Bayesian Skyline Plots of population size through time for *Sooglossus thomasseti*. The full view of the posterior all of the samples that are summarised by the median and 95% HPD interval are shown for the Mahé (top) and Silhouette (bottom) populations. EBSP analyses comprised all six loci. Time on x-axis in millions of years. Population on the y-axis represents the effective population size assuming a generation time of 1 year.

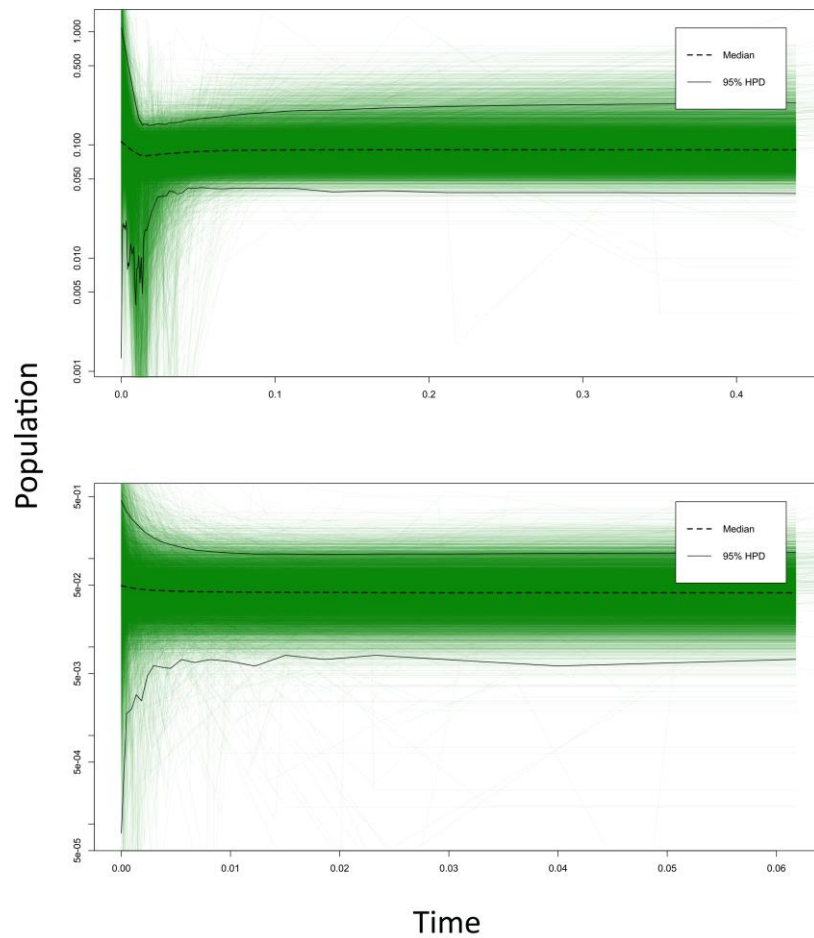


Figure S1.5 Extended Bayesian Skyline Plots of population size through time for *Sechellophryne gardineri*. The full view of the posterior all of the samples that are summarised by the median and 95% HPD interval are shown for the Mahé (top) and Silhouette (bottom) populations. EBSP analyses of the Mahé population comprised all six loci. Analyses of the Silhouette population comprised two loci (*16s*, *rag2*). Time on x-axis in millions of years. Population on the y-axis represents the effective population size assuming a generation time of 1 year.

References

- Biju SD, Bossuyt F. 2003.** New frog family from India reveals an ancient biogeographical link with the Seychelles. *Nature* **425**: 711-714.
- Bossuyt F, Milinkovitch MC. 2000.** Convergent adaptive radiations in Madagascan and Asian ranid frogs reveal covariation between larval and adult traits. *Proceedings of the National Academy of Sciences* **97**: 6585-6590.
- Chiari Y, Vences M, Vieites DR, Rabemananjara F, Bora P, Ramilijaona Ravoahangimalala O, Meyer A. 2004.** New evidence for parallel evolution of colour patterns in Malagasy poison frogs (*Mantella*). *Molecular Ecology* **13**: 3763-3774.
- Heller R, Chikhi L, Siegismund HR. 2013.** The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. *PLoS One* **8**: e62992.
- Jukes TH, Cantor CR. 1969.** Evolution of protein molecules. In: Munro HN, ed. *Mammalian protein metabolism*. New York.: Academic Press. 21-132.
- Lanfear R, Calcott B, Ho SY, Guindon S. 2012.** Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* **29**: 1695-1701.
- Palumbi S, Martin A, Romano S, McMillan WO, Stice L, Grabowski G. 1991.** *The simple fool's guide to PCR*. Dept. of Zoology and Kewalo Marine Laboratory, University of Hawaii: Honolulu, HI.
- Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S. 2013.** MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**: 2725-2729.