TITLE

Endemic, endangered, and evolutionarily significant: Cryptic lineages in Seychelles' frogs (Anura: Sooglossidae).

RUNNING TITLE

Cryptic diversity in the Sooglossidae

AUTHORS

Jim Labisko (corresponding author – jl693@kent.ac.uk)

Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, Canterbury, Kent. CT2 7NR. UK; Island Biodiversity and Conservation, P.O. Box 1348, Anse Royale, Mahé, Seychelles.

Richard A. Griffiths

Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, Canterbury, Kent. CT2 7NR. UK.

Lindsay Chong-Seng

Plant Conservation Action group, P.O. Box 392, Victoria, Mahé, Seychelles.

Nancy Bunbury

Seychelles Islands Foundation, La Ciotat Building, Mont Fleuri. P.O. Box 853, Victoria, Mahé, Seychelles; Centre for Ecology and Conservation, University of Exeter, Cornwall Campus, Penryn, TR10 9FE, UK.

Simon T. Maddock

School of Biology, Chemistry and Forensic Science, Faculty of Science and Engineering, University of Wolverhampton, Wulfruna Street, Wolverhampton. WV1 1LY. UK; Department of Life Sciences, The Natural History Museum, Cromwell Road, London. SW7 5BD. UK; Department of Genetics, Evolution and Environment, University College London, Gower

Street, London. WC1E 6BT. UK.; Island Biodiversity and Conservation, P.O. Box 1348, Anse Royale, Mahé, Seychelles.

Kay S. Bradfield

Perth Zoo, South Perth, WA 6151, Australia.

Michelle L. Taylor

Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, Canterbury, Kent. CT2 7NR. UK.

Jim J. Groombridge

Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, Canterbury, Kent. CT2 7NR. UK.

ABSTRACT

Cryptic diversity that corresponds with island origin has been previously reported in the endemic, geographically restricted sooglossid frogs of the Seychelles archipelago. The evolutionary pattern has not been fully explored, and given current amphibian declines and the increased extinction risk faced by island species, we sought to identify evolutionarily significant units (ESUs) to address conservation concerns for these highly threatened anurans. We obtained genetic data for two mitochondrial (mtDNA) and four nuclear (nuDNA) genes from all known populations of sooglossid frog (the islands of Mahé, Praslin, and Silhouette) to perform phylogenetic analyses and construct nuDNA haplotype networks. Bayesian and maximum likelihood analyses of mtDNA support monophyly and molecular differentiation of populations in all species that occur on multiple islands. Haplotype networks using statistical parsimony revealed multiple high-frequency haplotypes shared between islands and taxa, in addition to numerous geographically distinct (island-specific) haplotypes for each species. We consider each island-specific population of sooglossid frog as an ESU and advise conservation managers to do likewise. Furthermore, our results identify each island lineage as a candidate species, evidence for which is supported by Bayesian Poisson Tree Processes analyses of mtDNA, and independent analyses of mtDNA and nuDNA using the multispecies coalescent. Our findings add to the growing understanding of the biogeography and hidden diversity within this globally important region.

Keywords

Candidate species – cryptic diversity – evolutionarily significant unit – Indian Ocean – insular amphibians – islands – *Sechellophryne* – Seychelles – Sooglossidae – *Sooglossus*

INTRODUCTION

From the observations of Darwin (1859) and Wallace (1869) on the Galapagos and Malay archipelagos, to MacArthur and Wilson's (1967) seminal work on the theory of island biogeography, islands have played a significant role as model biological systems, progressing our understanding of evolutionary theory, ecological processes, and biogeography (Adsersen, 1995; Warren *et al.*, 2015; Santos *et al.*, 2016). The uniqueness of island endemic species is well documented, yet island biotas are particularly vulnerable to extinction, largely due to human-driven habitat change or introduced species (Paulay, 1994; Cronk, 1997; Whittaker & Fernández-Palacios, 2007). Understanding the evolutionary relationships of insular taxa, and addressing threats to endemic island lineages are therefore key components in mitigating the loss of global biodiversity (Robertson *et al.*, 2014).

The granitic Seychelles (most of the inner islands of the group; Fig. 1) form part of an isolated continental block with mixed faunal origins, including both overwater dispersed and ancient endemic clades (Ali, 2017; Ali, 2018) that reveal varying degrees of affinity to the Afrotropical and Indomalayan realms. Recent explorations of molecular phylogenetic relationships of Seychelles herpetofauna have identified a broad partitioning of two biogeographic units; a northern group (consisting of Praslin and surrounding islands) and a southern group (comprised of Mahé, Silhouette, and surrounding islands) (Fig. 1). This pattern of differentiation is documented in studies across a range of taxa, including the geckos *Ailuronyx* (Rocha *et al.*, 2016a), *Phelsuma* (Rocha *et al.*, 2013), *Urocotyledon* (Rocha *et al.*, 2011); and the skinks *Pamelaescincus* (Valente *et al.*, 2013) and *Trachylepis* (formerly *Mabuya*) (Rocha *et al.*, 2016b). However, within this north-south biogeographic pattern, further evidence of cryptic diversity is beginning to emerge in several taxa (e.g. Rocha *et al.*, 2016b). The discovery of a previously unknown population of sooglossid frogs on the island of Praslin—where the frogs had hitherto been unrecorded—and identification of this population as an evolutionarily significant unit (ESU) (Taylor *et al.*, 2012) provided the motivation to assess the genetic diversity of this family.

Sooglossid frogs

One of the world's most enigmatic and understudied frog families, the Sooglossidae (Noble, 1931) is one of only two amphibian families entirely restricted to an archipelago. Comprised of two genera, each with two species: *Sooglossus sechellensis* (Boettger, 1896) and *So. thomasseti* (Boulenger, 1909), and *Sechellophryne gardineri* (Boulenger, 1911) and *Se. pipilodryas* (Gerlach & Willi, 2002), each are recognised as Evolutionarily Distinct and Globally Endangered (EDGE) species and are placed in the Top 100 EDGE Amphibians (Isaac *et al.*, 2012; Zoological Society of London, 2015), and have been assessed for the IUCN Red List as either Critically Endangered (*So. thomasseti*, *Se. pipilodryas*) or Endangered (*So. sechellensis*, *Se. gardineri*) (IUCN SSC Amphibian Specialist Group, 2013a; IUCN SSC Amphibian Specialist Group, 2013b; IUCN SSC Amphibian Specialist Group, 2013c; IUCN SSC Amphibian Specialist Group, 2013d). Three of the four species occur on more than one island, with *So. thomasseti* and *Se. gardineri* found on Mahé and Silhouette (Nussbaum, 1984), and *So. sechellensis* found on Mahé, Silhouette, and Praslin (Nussbaum, 1984; Taylor *et al.*, 2012). The fourth species, *Se. pipilodryas*, is endemic to Silhouette (Gerlach & Willi, 2002).

Given the (i) importance of maintaining global and regional biological diversity, (ii) increased extinction risk faced by island species, (iii) unabated international crisis of amphibian declines, and (iv) global significance of the Sooglossidae as an evolutionarily distinct group, this unique family is in urgent need of research attention. A stronger knowledge base is also essential for conservation practitioners to make informed decisions and manage the sooglossid populations. Following recent accounts of geographic partitioning in Seychelles herpetofauna, and the identification of a novel, evolutionarily distinct population of sooglossid frogs on Praslin, we hypothesise that: (1) undocumented cryptic diversity exists across the three islands where these sooglossids occur, and (2) identification of such diversity will correspond with biogeographic (island) origin. To test these hypotheses, we reconstructed mitochondrial DNA phylogenies to explore the presence of divergent, cryptic lineages; generated nuclear DNA haplotype networks to reveal phylogeographic relationships; and performed species tree reconstructions using the multispecies coalescent. Our results enable the identification of ESUs for conservation purposes (Moritz, 1994) and further our understanding of the biogeography of the region.

MATERIALS AND METHODS

Study site

The inner islands of the Seychelles archipelago lie 4-5°S to 55-56°E in the western Indian Ocean, and sit upon the Seychelles Bank, a largely submerged microcontinent of some 129,500 km^2 (Davies & Francis, 1964). Its flat upper section spans ca. 44,000 km^2 and lies an average depth of 55 m below present sea level (bpsl), emerging from which are the granitic inner islands (Davies & Francis, 1964; Matthews & Davies, 1966; Ali, 2018) (Fig. 1). The granitic Seychelles are unique among oceanic islands, being composed of continental rock, and formed upon separation from India ~63 Ma (Collier *et al.*, 2008; Chatterjee *et al.*, 2013). Elevated, forested areas on the largest and highest islands of Mahé (154 km^2, 905 m elevation), Praslin (38 km^2, 367 m elevation) and Silhouette (20 km^2, 740 m elevation) are the only locations where sooglossid frogs are found.

Genetic sampling

Non-lethal tissue samples (toe-clips) were obtained from frogs representing each species and island population (Fig. 1; Table 1). We sequenced genes regularly utilised in amphibian phylogenetics that represented varying rates of molecular evolution. These comprised two mitochondrial DNA (mtDNA) fragments: *16S* rRNA (*16s*) and cytochrome b (*cytb*), plus fragments of four nuclear loci (nuDNA): proopiomelanocortin (*pomc*), recombination activating genes (*rag*) 1 and 2, and rhodopsin exon 1 (*rho*). Genomic DNA was extracted following manufacturer's guidelines using the Bioline Isolate Genomic DNA Kit. Sequences from all loci were amplified via standard polymerase-chain reaction (PCR). For primers and cycling conditions see Appendix S1; Table S1.1 in Supporting Information. Products from PCR were sequenced by Macrogen, Korea. We also utilised GenBank sequence data arising from Taylor *et al*. (2012) (*So. sechellensis 16s*), van der Meijden *et al.* (2007) (*Se. pipilodryas 16s*, *rag1*, *rag2*), and for outgroups used in phylogenetic analyses (Table S1.2). Novel sequence data generated by this study have been submitted to GenBank under accession numbers MK058722-70; MK058781-823; MK058825-979; MK058996-9390; MK072763-5.

Sequence alignment

Sequences were quality trimmed in SEQUENCHER v. 5.3 (Gene Codes Corporation, 2015) and cross-checked with chromatograms by eye in MEGA6 (Tamura *et al.*, 2013). MEGA6 was also employed to visually check (e.g. for stop codons and indels), edit, and align sequence data using default settings of the MUSCLE algorithm (Edgar, 2004). To remove any ambiguously aligned regions, sequence profiles were prepared via the GBLOCKS server v. 0.91b (Castresana, 2000; Talavera & Castresana, 2007). To preserve informative insertions and/or deletions, GBLOCKS parameters were set to allow gaps and less stringent flanking positions.

DATACONVERT 1.0 (Dyer *et al.*, n.d.), ALTER (Glez-Pena *et al.*, 2010) and FORMAT CONVERTOR (Los Alamos National Security LLC, 2005-2006) were employed to convert sequence profiles between required formats. SEQUENCEMATRIX v. 1.7.8 (Vaidya *et al.*, 2011) was used to concatenate mtDNA sequence profiles.

Mitochondrial phylogeny

Sooglossid mtDNA sequences were analysed using Bayesian inference (BI) (Huelsenbeck *et al.*, 2001) and maximum likelihood (ML) (Felsenstein, 1981) approaches. Partitioning schemes and models of nucleotide evolution were determined independently with PARTITIONFINDER v. 1.1.1 (Lanfear *et al.*, 2012) (Table S1.3). Branch lengths of alternative partitions were linked and all schemes evaluated using the Akaike information criterion. Bayesian analysis was performed in BEAST v. 2.3.2 (Bouckaert *et al.*, 2014) using two independent Markov chains of 100 million generations, sampling every 10,000 generations. BEAST input files were generated using BEAUTI v. 2.3.2 (Bouckaert *et al.*, 2014). Chain convergence and all parameters were checked using TRACER v. 1.6 (Rambaut *et al.*, 2014) to ensure adequate mixing and effective sample size (ESS) values ≥ 200. Initial runs were used to fine-tune final analyses, and we employed a relaxed lognormal clock as this approach may more accurately reflect lineage- and locus-specific heterogeneity in rates of molecular evolution (Drummond *et al.*, 2006; Lepage *et al.*, 2007; Heled & Drummond, 2010). As BEAST uses a molecular clock to estimate the root position, no outgroup taxa were used in BI analyses (Heled & Drummond, 2010). We assumed a stable environment for the Sooglossidae over recent geological time, and therefore applied a constant population for tree priors. However, given our inter- and intraspecific sampling we also performed phylogenetic reconstruction using the Yule model tree prior. Support for internal branches was evaluated using Bayesian posterior probabilities

(PP), with well-supported clades indicated by PP values ≥ 0.95. LOGCOMBINER V. 2.3.2 (Bouckaert *et al*., 2014) was used to combine tree files from the two independent runs, which was summarised as a single maximum clade credibility tree with mean PP values after a 10% burn-in using TREEANNOTATOR V. 1.8.2 (Drummond & Rambaut, 2007).

Maximum likelihood analyses were performed with RAXMLGUI V. 1.3.1 (Silvestro & Michalak, 2012; Stamatakis, 2014) using default settings with GTRGAMMA model parameters and 1,000 bootstrap replicates. Branch lengths were individually optimised for each partition. The Nasikabatrachidae have been hypothesised to be the closest extant relative of the Sooglossidae (Biju & Bossuyt, 2003; Frost *et al.*, 2006; Roelants *et al.*, 2007; Pyron & Wiens, 2011; Frazão *et al.*, 2015; Feng *et al.*, 2017) and used as an outgroup taxon in previous phylogenetic analyses of sooglossid frogs (van der Meijden *et al*., 2007; Taylor *et al*., 2012). However, GenBank derived *Nasikabatrachus sahyadrensis* sequence data rendered *Sooglossus* and *Sechellophryne* non-monophyletic in initial runs. Leiopelmatoidea (*Leiopelma*+*Ascaphus*) is widely accepted as the basal, sister lineage to all other extant anurans, and we therefore applied this taxon as an outgroup using GenBank sequence data arising from Irisarri *et al*. (2010) (*Leiopelma*) and Gissi *et al*. (2006) (*Ascaphus*). Support for internal branches was evaluated using bootstrap support (BS) values, with well-supported clades indicated by BS values ≥ 70. Bayesian and maximum likelihood phylogenies were visualised using FIGTREE v. 1.4.3 (Rambaut, 2016).

Multispecies coalescent and inferrence of population boundaries

To infer underlying species trees and support a robust phylogenetic insight, we performed reconstructions using the multispecies coalescent applied in the StarBEAST (*BEAST) package within BEAST v. 2.4.8 (Bouckaert *et al.*, 2014). Multiple samples per lineage are recommended to infer coalescent events, speciation and topology (Heled & Drummond 2010; Jockusch *et al*., 2014; Lambert *et al*., 2015), therefore where possible we utilised composite taxa to achieve coverage where only a single representative of a lineage was available. Data for composites was derived from individuals arising from the same taxon and population of origin, thereby meeting previously published criteria for composite taxa in amphibian studies (e.g. Alonso *et al*., 2012; Jockusch *et al*., 2014; Maia-Carvalho *et al*., 2014) (Table S1.4). The inclusion of variable loci such as mtDNA may exert disproportionate influence on other loci in *BEAST analysis (Jockusch *et al*., 2014). Accordingly, we carried out independent analyses of our mtDNA and nuDNA datasets. Partitioning schemes replicated that of our BEAST2 analyses (Table S1.3). Using a relaxed lognormal clock we ran two independent Markov chains of 100 million generations, sampling every 1,000 generations, and applied the ‘linear with constant root’ multispecies coalescent prior with the Yule model distribution of prior probability. Mitochondrial DNA shared the same tree partition, nuDNA tree partitions were locus specific. Checks on chain convergence and ESS values were performed as previously described. Clade support was evaulated using PP values. Trees were visualised using FIGTREE.

To infer popuation boundaries and aid the identifcation of ESUs we subjected our BEAST2 mtDNA phylogeny to Bayesian Poisson Tree Processes (bPTP) analysis implemented via the online bPTP service (http://species.h-its.org/ptp) (Zhang *et al*., 2013). The bPTP model applies two independent Poisson process classes (within- and among-species substitution events) under a coalescent model by assuming gene tree branch lengths to infer species/population boundaries. The bPTP analyses was run for 500 k Markov Chain Monte Carlo generations, with a thinning parameter of 100, and a burn-in of 0.1. Posterior probabilities of each node were assessed using maximum likelihood.

Genetic variation

MEGA6 was used to calculate nucleotide diversity, parsimony informative and variable sites, and obtain inter-and intra-specific genetic *p*-distances for mtDNA, with pair-wise deletion of missing sites. The PHASE algorithm (Stephens *et al.*, 2001; Stephens & Scheet, 2005) implemented in DNASP V. 5.10.1 (Librado & Rozas, 2009) was used to determine heterozygous positions and infer nuDNA haplotypes. Missing data can affect the success of haplotype phasing and detection of identical sequences (Salerno *et al.*, 2015), therefore short sequence reads were removed (*rag2*, *So. sechellensis*: Mahé = six, Praslin = two) and complete alignments for all nuDNA loci constructed. Random four-digit seeds generated by the TRUE RANDOM NUMBER SERVICE (Haahr, 2015) were applied to PHASE analyses which was run five times per locus with the highest pseudo-likelihood score used to select the best-fit model of haplotype estimation (Stephens & Donnelly, 2003). Heterozygous positions were deemed those achieving a score ≥ 0.7 (Harrigan *et al.*, 2008) and coded according to the International Union of Pure and Applied Chemistry. Remaining ambiguous positions were coded as ‘N’. To check for saturation across codon positions, the test of Xia (Xia *et al.*, 2003) was applied using DAMBE V. 5.5.29 (Xia, 2013). To check for evidence of recombination, the DATAMONKEY software suite (Pond & Frost, 2005; Delport *et al.*, 2010) was employed to select appropriate models and run analyses using the GARD application (Kosakovsky Pond *et al.*, 2006a; Kosakovsky Pond *et al.*, 2006b) under default settings. Haplotype networks were constructed using TCS V. 1.21 (Clement *et al.*, 2000) with a 95% connection limit and gaps treated as a fifth state. TCS networks were ‘beautified’ using TCSBU (Murias dos Santos *et al.*, 2016). Phased sequence data was used to infer haplotypes.

To detect evidence of historical population expansion or contraction in *So. sechellensis*, *So. thomasseti*, and *Se. gardineri* (*Se. pipilodryas* was excluded due to limited sampling), we applied neutrality tests and performed skyline plots. Tajima's D (Tajima, 1989), Fu's F_S (Fu, 1997), and the R_2 test statistic (Ramos-Onsins & Rozas, 2002) were run in DNASP v. 5.10.1 (Librado & Rozas, 2009) and applied to each locus individually. One thousand coalescent simulations were run for F_S and the R_2 test. A conventional P value of 0.05 was adopted for Tajima's D and R_2; Fu's F_S is interpreted as significant at $P < 0.02$. We performed Extended Bayesian Skyline plots (EBSP; Heled & Drummond, 2008) using unphased data for each island-specific population in BEAST v. 2.4.8 via the CIPRES Science Gateway (Miller *et al.*, 2010). The Jeffrey's (1/x) prior was applied to the data but to reduce over-parameterisation we adopted a strict clock and the HKY substitution model (Hasegawa *et al.*, 1985) to locus-specific partitions following the EBSP tutorial (http://www.beast2.org/tutorials). Chain length ranged from 50 to 300 million generations, sampling every 10,000 generations. Convergence, population size changes, and ESS values were assessed using TRACER, ESPB plots were visualised using R (R CORE TEAM, 2017).

Finally, to investigate patterns suggestive of isolation by distance across all multi-distributed sooglossid taxa, we performed Mantel tests with 999 permutations on independent (to reduce conflict from incomplete sampling) *16s* and *cytb* matrices using the VEGAN package (Oksanen *et al.*, 2017) in R (R CORE TEAM, 2017). The GEOGRAPHIC DISTANCE MATRIX GENERATOR v. 1.2.3 (Ersts, 2012) was used to generate pairwise distance matrices for geographic localities. Sequences without corresponding geographic data were omitted. Sampling localities are shown in Fig. 1.

RESULTS

Molecular phylogeny and genetic variation

Our final mtDNA sequence alignment of 56 sooglossids (*So. thomasseti* = 9, *So sechellensis* =

37, *Se. gardineri* = 9, *Se. pipilodryas* = 1) contained corresponding sequence data totalling 1,080 sites for 51 individuals. We were unable to obtain *cytb* sequence data for five Silhouette *Se. gardineri*, which constituted the majority (82%) of the total missing data of 3%. Rather than omit Silhouette *Se. gardineri* from our analyses (we are unaware of alternative *cytb* data for this taxon) we chose to maintain taxonomic coverage in all tree reconstructions.

Indels were present in the *16s* (*So. thomasseti* x1 double bp; *Se. gardineri* x2 single bp; *Se pipilodryas* x3 single bp) and *pomc* (*Sechellophryne* spp. x1 triple bp) sequence profiles. No evidence of saturation, or recombination events was detected in coding loci. Summary statistics of informative, uninformative, variable, and constant sites are shown in Table 1. Uncorrected and corrected genetic distances between taxa show values of 5.8%-14.0% and 6.1%-15.6% respectively (Table 2). Between population genetic distances are 2.0%-4.5% (uncorrected) and 2.1%-4.7% (corrected) for *So. sechellensis*; 2.1% for *So. thomasseti* (uncorrected and corrected); and 3.6% (uncorrected) and 3.7% (corrected) for *Se. gardineri* (Table 3).

Our Bayesian and maximum likelihood mtDNA reconstructions displayed highly concordant internal topologies (Fig. 2; Fig. S1.1), and recovered full support for the monophyly of *Sooglossus* and *Sechellophryne*. Island-specific populations of *Se*. *gardineri* and *So. thomasseti* are strongly supported. Geographic structuring in *So*. *sechellensis* receives strong support in BI analysis but moderate support in the ML tree, recovering a sister relationship between Mahé frogs and a clade comprising those from Silhouette and Praslin. A further distinction between Silhouette and Praslin populations receives strong support. Bayesian phylogenetic reconstructions applying the Yule tree prior reflect that of analyses using the constant population tree prior but provide reduced support for the monophyly of *Sooglossus* and independent island populations of *Se. gardineri* and *So. Sechellensis* (Fig.

S1.2).

Species trees and population boundaries

The multilocus species trees are broadly congruent with our mtDNA phylogenies (Fig. 2-3). The single topological disparity being internal relationships of *So. sechellensis* whereby the nuDNA species tree places Praslin frogs as sister to a clade comprised of those from Mahé and Silhouette. This contrasts with the mtDNA phylogeny and species trees which place Mahé frogs as sister to a Praslin and Silhouette clade. Clades and sub-clades are generally well supported except in the nuDNA tree where *Sooglossus* and *Sechellophryne* receive moderate support, and the sister taxon relationship between the Mahé and Silhouette populations of *Se. gardineri* is unresolved.

Ten well-supported entities are indicated from bPTP analyses, eight of which correspond with island populations of the multi-distributed sooglossid taxa shown in the mtDNA phylogeny (Fig. 2; Table S1.5). The remaining two entities represent members of an internal clade of *So. sechellensis* on Mahé; one a single sample from the southern-most population, the other comprised of one sample from the southern-most population and one from a more northerly locality (Fig. 1-2; Fig. S1.1-2; Table S1.5).

Nuclear DNA haplotypes

For each of the four nuclear loci, constructed networks show two or more high-frequency haplotypes in combination with multiple species- and population-specific haplotypes (Fig. 4-7). In the network for *pomc* (36 haplotypes; Fig. 4) two mutational steps separated both *So. sechellensis* and *So. thomasseti*, and the Mahé and Silhouette populations of *Se. gardineri*. One haplotype was shared between genera for *rag1* (37 haplotypes; Fig. 5) with seven mutational steps separating *So. sechellensis* and *So. thomasseti*. The *rag2* network (123 haplotypes; Fig. 6) shows five mutational steps separating *So. sechellensis* and *So. thomasseti*, and two mutational steps between one of the two *Se. pipilodryas* haplotypes and *Se. gardineri*. No genus specific characters were observed for *rho* (26 haplotypes; Fig. 7) where four haplotypes were shared between *Sooglossus* and *Sechellophryne*. Given the analytical thresholds we set, three networks (*pomc*, *rag1*, *rag2*) were divergent enough to differentiate (disconnect) genera and identify independent haplotypes for *Se. pipilodryas*. All loci displayed unique island-specific haplotypes for each multi-distributed species.

Population demography

Neutrality tests to understand population demographics in the Sooglossidae showed mostly negative values, indicating positive selection or recent population expansion (Table S1.6). However, statistically significant negative values are observed only in calculations of F_S, which may be less effective with small sample sizes (Ramos-Onsins & Rozas, 2002). Statistically significant positive values are evident in *16s* for all three species for Tajima's *D* but not F_S. Tajima's *D* is not as powerful as either F_S or the R_2 test statistic, and the R_2 test is considered to be more effective when applied to smaller sample sizes (Ramos-Onsins & Rozas, 2002; Ramirez-Soriano *et al.*, 2008). Significant positive values ($P < 0.05$) for a single locus in each species (*cytb*: *Se. gardineri*; *rag2*: *So. sechellensis*; *pomc*: *So. thomasseti*) were returned for the R_2 test. This suggests a lack of congruence that may be more indicative of differential levels of ancestral polymorphisms, selective pressures, and substitution rates across species and among loci, than statistically significant departures from neutrality.

In EBSP analyses the 95% highest posterior density (HPD) interval returned for Mahé and Silhouette populations of *So. sechellensis*, *So. thomasseti* and *Se. gardineri* included 0, therefore a constant population size for these taxa cannot be rejected (Table S1.7; Fig. S1.3-5). However, recent (within the last ~20 k years) population expansion appears to have occurred in *So. sechellensis* on Praslin (Fig. S1.3).

Isolation by distance

Matrices for our investigation of the effect of isolation by distance comprised 149 *So. sechellensis*, 29 *So. thomasseti*, and 26 *Se. gardineri* for *16s*, and 39 *So. sechellensis* and 9 *So. thomasseti* for *cytb*. Mantel tests indicated significant correlation between genetic and geographic distances in all species for both loci (*16s*: *So. sechellensis*, $r = 0.8253$, $P < 0.001$; *So. thomasseti*, $r = 0.9895$, $P < 0.001$; *Se gardineri*, $r = 0.9642$, $P < 0.001$; *cytb*: *So. sechellensis*, $r = 0.6995$, $P < 0.001$; *So. thomasseti*, $r = 0.9755$, $P < 0.05$).

DISCUSSION

Sooglossid phylogeny and genetic differentiation

Our analyses provide the first multi-gene phylogeny to use island-specific sampling to reveal intraspecific relationships within this endemic family. The mtDNA phylogeny supports our first hypothesis—that cryptic sooglossid diversity exists across the three islands where these frogs occur—and confirms the evolutionary distinctiveness of multiple geographically restricted sooglossid populations (Fig. 2). Our second hypothesis—that cryptic diversity corresponds with biogeographic (island) origin—is supported by independent evolutionary histories for the multi-distributed *Sooglossus* and *Sechellophryne* spp. in the mtDNA phylogeny, with distinct populations of *So. sechellensis* on Mahé, Silhouette, and Praslin, and *So. thomasseti* and *Se. gardineri* on Mahé and Silhouette (Fig. 2).

Mean uncorrected genetic distances among taxa for *16s* clearly reflect the greater differences expected between genera (range: 12.32%-14.04%; Table 2). Within genera, the Jukes-Cantor (JC) corrected *p*-distances between *So. sechellensis* and *So. thomasseti* (6.1%), and *Se. gardineri* and *Se. pipilodryas* (7.0%) exceed the values previously reported by van der Meijden *et al*. (2007) (4.4% in *Sooglossus* and 5.7% in *Sechellophryne*) for sequence data of comparable length. However, van der Meijden *et al*. (2007) sampled considerably fewer than 20 individuals in each case (*Sooglossus*: n = 7; *Sechellophryne*: n = 2); a limitation associated with an increased probability of underestimation of nucleotide diversity (Luo *et al.*, 2015). Estimations of genetic distance resulting from increased sampling are therefore more likely to represent the true population mean (Luo *et al*., 2015).

van der Meijden *et al*. (2007) also identified a JC corrected *p*-distance of 3.0% between the Mahé and Silhouette populations of *So. thomasseti* but did so from four samples; two from Mahé, two from Silhouette. We report a JC corrected *p*-distance of 2.1% from a pool of 29 individuals (Table 3) originating from four sites on the island of Mahé, and two sites on Silhouette. The spatial representation of our sampling, and greater sample size is therefore more likely to reflect a value closer to the true mean. Taylor *et al*. (2012) found uncorrected *16s p*-distances of 4.1%-6.1% between the Mahé, Silhouette, and Praslin populations of *So. sechellensis* from a total sample size of 26. We incorporate all but two of the *16s* sequences arising from Taylor *et al*. (2012) (these two omissions are Praslin samples placed within the Mahé clade in their study which are likely to be the result of laboratory contamination, as subsequent *cytb* analysis reflects their geographic origin; J. Labisko, unpubl. data) and report genetic distances of 2.1%-4.7% from 159 samples (Table 3).

Species trees and population boundaries

The multilocus species trees are highly congruent with our mtDNA phylogenies but clade support differs between the mtDNA and nuDNA analyses (Fig. 3). The lower levels of support displayed may reflect statistical inaccuracy from missing (*cytb*) sequence data as well as the inherent differential qualities of the loci we sampled. While the specific status of *Sooglossus* and *Sechellophryne* taxa are not in question, further exploration of the data incorporating additional loci may elucidate the strength of relationship between the two isolated populations of *Se. gardineri*. Overall, and in spite of topological disparity between two island lineages of *So. sechellensis*, the multispecies coalescent and bPTP model independently provide further support for the monophyly of multiple island-specific lineages of sooglossid frogs. For *So. sechellensis*, bPTP results also indicate additional intraspecific structure within the Mahé population.

Nuclear variation

There is an increasing body of evidence reporting discordant patterns between mtDNA and nuDNA markers in animal systems (Toews & Brelsford, 2012). Discordance between molecular markers may be especially pronounced in amphibians (Hoelzer, 1997; Monsen & Blouin, 2003), and nuclear genes are frequently recognised for their conflicting results in genealogical estimations in amphibian studies (e.g. Fisher-Reid & Wiens, 2011; Eto & Matsui, 2014). Our analyses identified multiple haplotypes shared between island populations, species, and genera. Nevertheless, geographic structuring of the Sooglossidae is visibly evident in the nuclear loci we sampled, showing numerous unique haplotypes across all multi-distributed taxa and a commonality between our mtDNA and nuDNA datasets. While these nuDNA patterns may indicate a level of diversity within each population that differentiates it from congeners on other islands, the data are also likely to reflect incomplete sampling and incomplete lineage sorting; the latter especially so considering maternal line of inheritance and smaller effective population size of mtDNA in comparison to the diploid, bi-parental nature of nuDNA.

Biogeographic and conservation implications

Due to their intolerance of salt water, trans-oceanic dispersal is assumed to be an infrequent method of range expansion for amphibians (Duellman & Trueb, 1986; Green *et al.*, 1988; de Queiroz, 2005). The presence of endemic amphibians on oceanic islands may therefore be considered unusual, yet rafting is increasingly cited as an explanation for transoceanic dispersal of frogs (Vences *et al.*, 2003; Heinicke *et al.*, 2007; Maddock *et al.*, 2014; Bell *et al.*, 2015a; Bell *et al.*, 2015b), and even caecilians (Measey *et al.*, 2006). However, in each case, pioneering dispersers have mainland congeners. Aside from their sister taxon relationship with the Nasikabatrachidae of India's Western Ghats, from which they may have diverged 66-131 Ma—prior to the geographic separation of India and Seychelles (Biju & Bossuyt, 2003; Roelants *et al*., 2007; Ruane *et al.*, 2011; Pyron, 2014; Frazão *et al*., 2015; Feng *et al*., 2017)—the Sooglossidae have no recent relatives. The level of evolutionary distinctiveness displayed by these frogs, undoubtedly a result of their lengthy isolation, is clearly evidence of their historic and continuing presence on the archipelago.

Following its separation from India, the inner Seychelles region has formed both a continuous landmass of some 129,500 km^2, and been submerged to its present extent, comprising an archipelago of 45 inner-islands covering ~247 km^2. Had the Seychelles Bank ever been completely submerged, this would be strongly reflected in the composition of its fauna and flora, with an expectation of greater similarity to that of Africa and/or Asia (Nussbaum, 1984). The region has been subject to eustatic fluctuations, climatic variability, and vicariant events, which have played an influential role in the distribution of its biota.

Recently identified phylogeographic patterns within the archipelago's endemic herpetofauna have revealed a variety of geographic correlations: skinks and geckos broadly differentiate into northern (Praslin and surrounding islands) and southern (Mahé, Silhouette, and surrounding islands) groups (Rocha *et al.*, 2010; Rocha *et al.*, 2011; Rocha *et al.*, 2013; Valente *et al.*, 2013; Rocha *et al.*, 2016a; Rocha *et al.*, 2016b); while for the non-sooglossid anurans, a distinct lack of variability is shown in the multi-distributed treefrog *Tachycnemis seychellensis* (Maddock *et al.*, 2014), conflicting with observed structuring in Seychelles endemic caecilians (Adamson *et al.*, 2016; Maddock *et al.*, 2016; Maddock *et al.*, 2017). Our mtDNA analyses appear to confirm the relationship posited by Taylor *et al*. (2012), namely that Silhouette and Praslin populations of *So. sechellensis* comprise a clade sister to that of frogs from Mahé. Yet our nuDNA species tree presents a topological contrast by inferring Praslin frogs as sister to a Silhouette and Mahé clade; harmonious with the north-south split identified in other Seychelles herpetofauna. This disparity raises the question as to what these conflicting biogeographic patterns in the data may reflect.

Since the Late Pleistocene, regional instability caused by either hydro-isostatic uplift of the Seychelles Bank or volcanic subsidence (Montaggioni & Hoang, 1988) and substantial low sea-level stands (Colonna *et al.*, 1996; Camoin *et al.*, 2004) have likely generated irregular cycles of biogeographic isolation and reconnection across the Seychelles. Bathymetric data indicate a sea-level drop of ~60 m bpsl would effectively link the granitic islands (Rocha *et al.*, 2013; Ali, 2018), providing the opportunity for dispersal and connection/reconnection of previously disparate populations. Incongruence between and among the phylogeographic patterns exhibited by Seychelles' herpetofauna are undoubtedly the result of a number of contributory factors, including the inherent ecology and dispersal ability of each taxon. Although these and other aspects are yet to be fully explored, Maddock *et al*. (2014) found low levels of genetic variation in *T. seychellensis* concluding, *inter-alia*, that relatively recent admixture during low sea-level stands may explain this observation. The treefrogs are regularly encountered in appropriate habitat at lower elevations down to sea-level, and may on occasion raft across the ocean as a means of dispersal, as their ancestors are believed to have done from Madagascar (Vences *et al*., 2003; Maddock *et al*., 2014). The terrestrial Sooglossidae (although *Sechellophryne* spp. may be observed in low-level vegetation; Gerlach & Willi, 2002; J. Labisko, pers. obs.) are generally restricted to high elevation moist forest, such that lower (and dryer) elevations combined with even a limited oceanic distance between suitable habitats, may act as a considerable barrier to dispersal. However, the islands of Mahé and Silhouette share high elevation peaks, similar forest habitat, and are currently separated by less than 20 km, and given a significant drop in sea level, the opportunity for dispersal between the two would inevitably increase. Mahé and Silhouette frogs may therefore be expected to share greater similarities than either do with those from Praslin—an island which is lower, drier, 37 km distant from Mahé, and 51 km from Silhouette—and the locus-specific nuDNA gene trees produced in our multispecies coalescent analyses display a largely congruent topology, suggesting these loci are representative of true relationships across the nuclear genome.

Clear geographic patterns of discordance between mtDNA and nuDNA are likely to exclude incomplete lineage sorting as an underlying explanation, and may instead indicate biogeographic discordance (Funk & Omland, 2003; Toews & Brelsford, 2012). Extended periods of isolation combined with previous range contact are an intrinsic factor in most taxa that display patterns of this nature, during which high frequency mutations accumulate and are followed by interbreeding in hybrid zones upon range reconnection, generating divergent patterns in the mtDNA and nuDNA genomes (Toews & Brelsford, 2012). The cycles of emergence and submergence of the Seychelles Bank are unknown but the patterns of genetic differentiation and population demography we report could be attributed to infrequent stable environmental conditions of adequate duration that would arise as a result of significant but sporadic eustatic fluctuations, and should not be discounted as a mechanism to explain the patterns observed in our data. It is noteworthy that one population—*So. sechellensis* from Praslin—appears to have recently expanded (Fig. S1.3). That no other sooglossids are found on Praslin suggests that (i) *So. sechellensis* is the only sooglossid to have occurred here, or (ii) other members of the Sooglossidae have since died out, perhaps as a result of the climactic effects and loss of terrestrial habitat following deglaciation and the rise in sea levels from the Late Pleistocene to Early Holocene (Dutton *et al*., 2015; Woodroffe *et al*., 2015). In either scenario, the Praslin frogs have seemingly been successful in exploiting available habitat on this island in the absence of other sooglossids.

Reciprocal monophyly in mtDNA together with significant divergence in nuDNA loci have long been criteria for defining evolutionarily significant units (Moritz, 1994). Our results meet these criteria, showing numerous unique, geographically specific haplotypes in nuclear loci, and reciprocal monophyly in mtDNA for sooglossid populations, additional analyses of which indicates significant effects of isolation by distance. Sooglossid lineages that reflect island origin are defined across all multi-distributed species: *So. sechellensis* on Mahé, Silhouette, and Praslin, and *So. thomasseti* and *Se. gardineri* on Mahé and Silhouette (Fig. 2-3). Furthermore, and in accordance with the criteria ascribed by Vieites *et al*. (2009), we consider these lineages as unconfirmed candidate species. Given the limitations of species delimitation methods in distinguishing structure from population isolation versus species boundaries (Sukumaran & Knowles, 2017; Leaché *et al*., 2018) (evidenced in our study by the identification of intraspecific structure within the Mahé population of *So. sechellensis*; Fig. 2;

Table S1.5), a continuing formal taxonomic appraisal for the Sooglossidae, combining multiple lines of evidence to corroborate hypotheses of distinct lineages is underway and will be presented elsewhere.

Our investigation of an understudied insular taxon, endemic to the Seychelles archipelago, adds to the developing biogeographic picture of this unique region. Patterns of cryptic diversity in Seychelles' amphibians have only recently begun to be explored, yet already appear to be highly prevalent and complex. Prior to our study, four sooglossid species were recognised across the three islands upon which they occur, with one population—the only sooglossid found on the island of Praslin, *So. sechellensis*—determined as fitting the criteria of an additional ESU (Taylor *et al*., 2012). The cryptic diversity we have uncovered denotes a total of eight independent island lineages that should be managed accordingly. Such management action should include regular long-term population and habitat assessments, support of the genetic integrity of each ESU by carrying out no inter-island translocations, and the establishment of regular screening activities for invasive pathogens including *Batrachochytrium dendrobatidis*, *B. salamandrivorans*, and *Ranavirus*—notably, the Seychelles is one of only two global regions where pathogenic chytrid is yet to be detected (Labisko *et al*, 2015; Lips, 2016). The identification of distinct, island-specific populations of these frogs warrants continued investigation of their intraspecific relationships, and further insights are likely to reveal additional factors important for their future conservation.

ACKNOWLEDGEMENTS

We express our gratitude for support provided by the Darwin Initiative (grant 19-002 to RAG and JJG); Durrell Institute of Conservation and Ecology; The Natural History Museum, London; Seychelles Islands Foundation; Seychelles National Parks Authority; The Systematics

Association and The Linnean Society (independent Systematics Research Fund awards to JL and STM); University College London; and the University of Kent. We thank the Seychelles Bureau of Standards for permission to carry out fieldwork; Seychelles Department of Environment for permission to collect and export samples; Matthieu La Buschagne for access to Coco de Mer Hotel land on Praslin; Island Conservation Society for field assistance on Silhouette; Islands Development Company for permission and hosting on Silhouette; and Rachel Bristol for organisational and field assistance. We also thank The Mohammad bin Zayed Species Conservation Fund for their continuing support of JL (Project 172515128) and STM (Project 162513749). Wilna Accouche, Katy Beaver, Darryl Birch, Georgia French, David Gower, Philip Haupt, Marc Jean-Baptiste, Christopher Kaiser-Bunbury, Pete Haverson, James Mougal, Marcus Pierre, Nathachia Pierre, Dainise Quatre, Anna Reuleaux, Heather Richards, Mark Wilkinson, and many other NGO staff, researchers, and Seychellois provided in- and ex-situ support, for which we are especially grateful. We thank Katy Beaver, Jeff Streicher, Ben Tapley, and three anonymous reviewers for helpful comments on previous drafts of this manuscript.

REFERENCES

Adamson EAS, Saha A, Maddock ST, Nussbaum RA, Gower DJ, Streicher JW. 2016. Microsatellite discovery in an insular amphibian (*Grandisonia alternans*) with comments on cross-species utility and the accuracy of locus identification from unassembled Illumina data. *Conservation Genetics Resources* **8:** 541-551.

Adsersen H. 1995. Research on islands: Classic, recent, and prospective approaches. In: Vitousek PM, Loope LL and Adsersen H, eds. *Islands*. Berlin, Heidelberg: Springer Berlin Heidelberg. 7-21.

Ali JR. 2017. Islands as biological substrates: classification of the biological assemblage components and the physical island types. *Journal of Biogeography* **44:** 984-994.

Ali JR. 2018. Islands as biological substrates: Continental. *Journal of Biogeography* **45:** 1003-1018.

Alonso R, Crawford AJ, Bermingham E. 2012. Molecular phylogeny of an endemic radiation of Cuban toads (Bufonidae: *Peltophryne*) based on mitochondrial and nuclear genes. *Journal of Biogeography* **39:** 434-451.

Bell RC, Drewes RC, Channing A, Gvoždík V, Kielgast J, Lötters S, Stuart BL, Zamudio KR, Emerson B. 2015a. Overseas dispersal of *Hyperolius* reed frogs from Central Africa to the oceanic islands of São Tomé and Príncipe. *Journal of Biogeography* **42:** 65-75.

Bell RC, Drewes RC, Zamudio KR. 2015b. Reed frog diversification in the Gulf of Guinea: overseas dispersal, the progression rule, and in situ speciation. *Evolution* **69:** 904-915.

Biju SD, Bossuyt F. 2003. New frog family from India reveals an ancient biogeographical link with the Seychelles. *Nature* **425:** 711-714.

Boettger O. 1896. Neue Kriechthiere (Scelotes, *Arthroleptis*) von den Seychellen. *Zoologischer Anzeiger* **19:** 349.

Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10:** e1003537.

Boulenger GA. 1909. No. XVI.-A list of the freshwater fishes, batrachians, and reptiles obtained by Mr. J. Stanley Gardiner's expedition to the Indian Ocean. *Transactions of the Linnean Society of London. 2nd Series: Zoology* **12:** 291-300.

Boulenger GA. 1911. No. XVII.-List of the batrachians and reptiles obtained by Prof. Stanley Gardiner on his second expedition to the Seychelles and Aldabra. *Transactions of the Linnean Society of London. 2nd Series: Zoology* **14:** 375-378.

Camoin GF, Montaggioni LF, Braithwaite CJR. 2004. Late glacial to post glacial sea levels in the western Indian Ocean. *Marine Geology* **206:** 119-146.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17:** 540-552.

Chatterjee S, Goswami A, Scotese CR. 2013. The longest voyage: Tectonic, magmatic, and paleoclimatic evolution of the Indian plate during its northward flight from Gondwana to Asia. *Gondwana Research* **23:** 238-267.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9:** 1657-1659.

Collier JS, Sansom V, Ishizuka O, Taylor RN, Minshull TA, Whitmarsh RB. 2008. Age of Seychelles–India break-up. *Earth and Planetary Science Letters* **272:** 264-277.

Colonna M, Casanova J, Dullo W-C, Camoin G. 1996. Sea-level changes and δ18O record for the past 34,000 yr from Mayotte Reef, Indian Ocean. *Quaternary Research* **46:** 335-339.

Cronk QCB. 1997. Islands: stability, diversity, conservation. *Biodiversity and Conservation* **6:**

477-493.

Darwin C. 1859. *On the origin of species*. Murray: London.

Davies D, Francis TJG. 1964. The crustal structure of the Seychelles bank. *Deep Sea Research and Oceanographic Abstracts* **11:** 921-927.

de Queiroz A. 2005. The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology and Evolution* **20:** 68-73.

Delport W, Poon AF, Frost SD, Kosakovsky Pond SL. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* **26:** 2455-2457.

Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4:** e88.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 214.

Duellman WE, Trueb L. 1986. *Biology of amphibians*. Johns Hopkins University Press.

Dutton A, Webster JM, Zwartz D, Lambeck K, Wohlfarth B. 2015. Tropical tales of polar ice: evidence of Last Interglacial polar ice sheet retreat recorded by fossil reefs of the granitic Seychelles islands. *Quaternary Science Reviews* **107:** 182-196.

Dyer M, Sailsbery J, McClellan D. n.d. DataConvert–biological data file conversion. Available at http://www.mybiosoftware.com/dataconvert-1-0-converts-proteindna-data-formats.html

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32:** 1792-1797.

Ersts P. 2012. Geographic Distance Matrix Generator version 1.23: American Museum of Natural History.

Eto K, Matsui M. 2014. Cytonuclear discordance and historical demography of two brown frogs, *Rana tagoi* and *R. sakuraii* (Amphibia: Ranidae). *Molecular Phylogenetics and Evolution* **79:** 231-239.

Felsenstein J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* **17:** 368-376.

Feng YJ, Blackburn DC, Liang D, Hillis DM, Wake DB, Cannatella DC, Zhang P. 2017. Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous-Paleogene boundary. *Proceedings of the National Academy of Sciences of the United States of America* **114:** E5864-E5870.

Fisher-Reid MC, Wiens J. 2011. What are the consequences of combining nuclear and mitochondrial data for phylogenetic analysis? Lessons from *Plethodon* salamanders and 13 other vertebrate clades. *BMC Evolutionary Biology* **11:** 300.

Frazão A, da Silva HR, Russo CA. 2015. The Gondwana breakup and the history of the Atlantic and Indian Oceans unveils two new clades for early neobatrachian diversification. *PLoS One* **10:** e0143926.

Frost DR, Grant T, Faivovich J, Bain RH, Haas A, Haddad CFB, De Sá RO, Channing A, Wilkinson M, Donnellan SC, Raxworthy CJ, Campbell JA, Blotto BL, Moler P, Drewes RC, Nussbaum RA, Lynch JD, Green DM, Wheeler WC. 2006. The Amphibian Tree of Life. *Bulletin of the American Museum of Natural History* **297:** 1-291.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147:** 915-925.

Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* **34:** 397-423.

Gene Codes Corporation. 2015. Sequencher® version 5.3 sequence analysis software. Ann

Arbor, MI USA.

Gerlach J, Willi J. 2002. A new species of frog, genus *Sooglossus* (Anura, Sooglossidae) from Silhouette Island, Seychelles. *Amphibia-Reptilia* **23:** 445-458.

Gissi C, San Mauro D, Pesole G, Zardoya R. 2006. Mitochondrial phylogeny of Anura (Amphibia): a case study of congruent phylogenetic reconstruction using amino acid and nucleotide characters. *Gene* **366:** 228-237.

Glez-Pena D, Gomez-Blanco D, Reboiro-Jato M, Fdez-Riverola F, Posada D. 2010. ALTER: program-oriented conversion of DNA and protein alignments. *Nucleic Acids Research* **38:** W14-18.

Green DM, Nussbaum RA, Datong Y. 1988. Genetic divergence and heterozygosity among frogs of the family Sooglossidae. *Herpetologica* **44:** 113-119.

Haahr M. 2015. True Random Number Service: http://www.random.org.

Harrigan RJ, Mazza ME, Sorenson MD. 2008. Computation vs. cloning: evaluation of two methods for haplotype determination. *Molecular Ecology Resources* **8:** 1239-1248.

Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22:** 160-174.

Heinicke MP, Duellman WE, Hedges SB. 2007. Major Caribbean and Central American frog faunas originated by ancient oceanic dispersal. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 10092-10097.

Heled J, Drummond AJ. 2008. Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology* **8:** 289.

Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27:** 570-580.

Hoelzer GA. 1997. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees revisited. *Evolution* **51:** 622-626.

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **294:** 2310-2314.

Irisarri I, San Mauro D, Green DM, Zardoya R. 2010. The complete mitochondrial genome of the relict frog *Leiopelma archeyi*: insights into the root of the frog Tree of Life. *Mitochondrial DNA* **21:** 173-182.

Isaac NJ, Redding DW, Meredith HM, Safi K. 2012. Phylogenetically-informed priorities for amphibian conservation. *PLoS One* **7:** e43912.

IUCN SSC Amphibian Specialist Group. 2013a. *Sechellophryne gardineri The IUCN Red List of Threatened Species 2013*.

IUCN SSC Amphibian Specialist Group. 2013b. *Sechellophryne pipilodryas The IUCN Red List of Threatened Species 2013*.

IUCN SSC Amphibian Specialist Group. 2013c. *Sooglossus sechellensis The IUCN Red List of Threatened Species 2013*.

IUCN SSC Amphibian Specialist Group. 2013d. *Sooglossus thomasseti The IUCN Red List of Threatened Species 2013*.

Jockusch EL, Martínez-Solano I, Timpe EK. 2015. The effects of inference method, population sampling, and gene sampling on species tree inferences: An empirical study in slender salamanders (Plethodontidae: Batrachoseps). *Systematic Biology* **64:** 66-83.

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. 2006a. Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology and Evolution* **23:** 1891-1901.

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. 2006b. GARD: a genetic algorithm for recombination detection. *Bioinformatics* **22:** 3096-3098.

Labisko J, Maddock ST, Taylor ML, Chong-Seng L, Gower DJ, Wynne FJ, Wombwell E, Morel C, French GCA, Bunbury N, Bradfield KS. 2015. Chytrid fungus (*Batrachochytrium dendrobatidis*) undetected in the two orders of Seychelles amphibians. *Herpetological Review* **46:** 41-45.

Lambert SM, Reeder TW, Wiens JJ. 2015. When do species-tree and concatenated estimates disagree? An empirical analysis with higher-level scincid lizard phylogeny. *Molecular Phylogenetics and Evolution* **82 Pt A:** 146-155.

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.

Leaché AD, Zhu T, Rannala B, Yang Z. 2018. The spectre of too many species. *Systematic Biology*.

Lepage T, Bryant D, Philippe H, Lartillot N. 2007. A general comparison of relaxed molecular clock models. *Molecular Biology and Evolution* **24:** 2669-2680.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:** 1451-1452.

Lips KR. 2016. Overview of chytrid emergence and impacts on amphibians. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **371**.

Los Alamos National Security LLC. 2005-2006. Format Converter: U.S. Department of Energy's National Nuclear Security Administration: https://www.hiv.lanl.gov/content/sequence/FORMAT_CONVERSION/form.html

Luo A, Lan H, Ling C, Zhang A, Shi L, Ho SY, Zhu C. 2015. A simulation study of sample size for DNA barcoding. *Ecology and Evolution* **5:** 5869-5879.

MacArthur RH, Wilson EO. 1967. *Theory of island biogeography*. Princeton University Press:

Princeton.

Maddock ST, Briscoe AG, Wilkinson M, Waeschenbach A, San Mauro D, Day JJ, Littlewood DT, Foster PG, Nussbaum RA, Gower DJ. 2016. Next-generation mitogenomics: A comparison of approaches applied to caecilian amphibian phylogeny. *PLoS One* **11:** e0156757.

Maddock ST, Day JJ, Nussbaum RA, Wilkinson M, Gower DJ. 2014. Evolutionary origins and genetic variation of the Seychelles treefrog, *Tachycnemis seychellensis* (Dumeril and Bibron, 1841) (Amphibia: Anura: Hyperoliidae). *Molecular Phylogenetics and Evolution* **75:** 194-201.

Maddock ST, Wilkinson M, Nussbaum RA, Gower DJ. 2017. A new species of small and highly abbreviated caecilian (Gymnophiona: Indotyphlidae) from the Seychelles island of Praslin, and a recharacterization of *Hypogeophis brevis* Boulenger, 1911. *Zootaxa* **4329:** 301.

Maia-Carvalho B, Goncalves H, Ferrand N, Martinez-Solano I. 2014. Multilocus assessment of phylogenetic relationships in *Alytes* (Anura, Alytidae). *Molecular Phylogenetics and Evolution* **79:** 270-278.

Matthews DH, Davies D. 1966. Geophysical studies of the Seychelles Bank. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **259:** 227-239.

Measey GJ, Vences M, Drewes RC, Chiari Y, Melo M, Bourles B. 2006. Freshwater paths across the ocean: molecular phylogeny of the frog *Ptychadena newtoni* gives insights into amphibian colonization of oceanic islands. *Journal of Biogeography* **34:** 7-20.

Miller MA, Pfeiffer W, Schwartz T. 2010. Proceedings of the Gateway Computing Environments Workshop (GCE) *Creating the CIPRES science gateway for inference of* *large phylogenetic trees*: IEEE New Orleans. 1-8.

Monsen KJ, Blouin MS. 2003. Genetic structure in a montane ranid frog: restricted gene flow and nuclear-mitochondrial discordance. *Molecular Ecology* **12:** 3275-3286.

Montaggioni LF, Hoang CT. 1988. The last interglacial high sea level in the granitic Seychelles, Indian ocean. *Palaeogeography, Palaeoclimatology, Palaeoecology* **64:** 79-91.

Moritz C. 1994. Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology and Evolution* **9:** 373-375.

Murias dos Santos A, Cabezas MP, Tavares AI, Xavier R, Branco M. 2016. tcsBU: a tool to extend TCS network layout and visualization. *Bioinformatics* **32:** 627-628.

Noble GK. 1931. *The biology of the Amphibia*. McGraw-Hill Book Company, Inc.: New York and London.

Nussbaum RA. 1984. Amphibians of the Seychelles. In: Stoddart DR, ed. *Biogeography and ecology of the Seychelles Islands*. The Hague; Boston: Hingham, Massachusetts, USA: W. Junk. 379-415.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, M. Henry H. Stevens, Szoecs E, Wagner H. 2017. vegan: Community Ecology Package. 2.4-4. ed.

Paulay G. 1994. Biodiversity on oceanic Islands: Its origin and extinction. *American Zoologist* **34:** 134-144.

Pond SL, Frost SD. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* **21:** 2531-2533.

Pyron RA. 2014. Biogeographic analysis reveals ancient continental vicariance and recent oceanic dispersal in amphibians. *Systematic Biology* **63:** 779-797.

Pyron RA, Wiens JJ. 2011. A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Molecular Phylogenetics and Evolution* **61:** 543-583.

R Core Team. 2017. R: A language and environment for statistical computing. Vienna, Austria.: R Foundation for Statistical Computing.

Rambaut A. 2016. FigTree 1.4.3.

Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Tracer v1.6.

Ramirez-Soriano A, Ramos-Onsins SE, Rozas J, Calafell F, Navarro A. 2008. Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. *Genetics* **179:** 555-567.

Ramos-Onsins SE, Rozas J. 2002. Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19:** 2092-2100.

Robertson JM, Langin KM, Sillett TS, Morrison SA, Ghalambor CK, Funk WC. 2014. Identifying evolutionarily significant units and prioritizing populations for management on islands. *Monographs of the Western North American Naturalist* **7:** 397-411.

Rocha S, Harris D, Posada D. 2011. Cryptic diversity within the endemic prehensile-tailed gecko *Urocotyledon inexpectata* across the Seychelles Islands: patterns of phylogeographical structure and isolation at the multilocus level. *Biological Journal of the Linnean Society* **104:** 177-191.

Rocha S, Harris DJ, Carretero M. 2010. Genetic diversity and phylogenetic relationships of *Mabuya* spp. (Squamata: Scincidae) from western Indian Ocean islands. *Amphibia-Reptilia* **31:** 375-385.

Rocha S, Perera A, Bunbury N, Kaiser-Bunbury CN, Harris DJ. 2016a. Speciation history and species-delimitation within the Seychelles Bronze geckos, *Ailuronyx* spp.: molecular and morphological evidence. *Biological Journal of the Linnean Society*.

Rocha S, Perera A, Silva A, Posada D, Harris DJ. 2016b. Evolutionary history of *Trachylepis* skinks in the Seychelles islands: introgressive hybridization, morphological evolution and geographic structure. *Biological Journal of the Linnean Society* **119:** 15-36.

Rocha S, Posada D, Harris DJ. 2013. Phylogeography and diversification history of the day-gecko genus *Phelsuma* in the Seychelles islands. *BMC Evolutionary Biology* **13:** 3.

Roelants K, Gower DJ, Wilkinson M, Loader SP, Biju SD, Guillaume K, Moriau L, Bossuyt F. 2007. Global patterns of diversification in the history of modern amphibians. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 887-892.

Ruane S, Pyron RA, Burbrink FT. 2011. Phylogenetic relationships of the Cretaceous frog *Beelzebufo* from Madagascar and the placement of fossil constraints based on temporal and phylogenetic evidence. *Journal of Evolutionary Biology* **24:** 274-285.

Salerno PE, Senaris JC, Rojas-Runjaic FJ, Cannatella DC. 2015. Recent evolutionary history of Lost World endemics: population genetics, species delimitation, and phylogeography of sky-island treefrogs. *Molecular Phylogenetics and Evolution* **82 Pt A:** 314-323.

Santos AMC, Field R, Ricklefs RE. 2016. New directions in island biogeography. *Global Ecology and Biogeography* **25:** 751-768.

Silvestro D, Michalak I. 2012. raxmlGUI: a graphical front-end for RAxML. *Organisms Diversity & Evolution* **12:** 335-337.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312-1313.

Stephens M, Donnelly P. 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics* **73:** 1162-1169.

Stephens M, Scheet P. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *American Journal of Human Genetics* **76:** 449-462.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* **68:** 978-989.

Sukumaran J, Knowles LL. 2017. Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences* **114:** 1607-1612.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123:** 585-595.

Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* **56:** 564-577.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30:** 2725-2729.

Taylor ML, Bunbury N, Chong-Seng L, Doak N, Kundu S, Griffiths RA, Groombridge JJ. 2012. Evidence for evolutionary distinctiveness of a newly discovered population of sooglossid frogs on Praslin Island, Seychelles. *Conservation Genetics* **13:** 557-566.

Toews DP, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* **21:** 3907-3930.

Vaidya G, Lohman DJ, Meier R. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* **27:** 171-180.

Valente J, Rocha S, Harris DJ. 2013. Differentiation within the endemic burrowing skink

Pamelaescincus gardineri, across the Seychelles islands, assessed by mitochondrial and nuclear markers. *African Journal of Herpetology* **63:** 25-33.

van der Meijden A, Boistel R, Gerlach J, Ohler A, Vences M, Meyer A. 2007. Molecular phylogenetic evidence for paraphyly of the genus *Sooglossus*, with the description of a new genus of Seychellean frogs. *Biological Journal of the Linnean Society* **91:** 347 - 359.

Vences M, Vieites DR, Glaw F, Brinkmann H, Kosuch J, Veith M, Meyer A. 2003. Multiple overseas dispersal in amphibians. *Proceedings of the Royal Society B: Biological Sciences* **270:** 2435-2442.

Vieites DR, Wollenberg KC, Andreone F, Kohler J, Glaw F, Vences M. 2009. Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 8267-8272.

Wallace AR. 1869. *The Malay Archipelago: the land of the orang-utan and the bird of paradise; a narrative of travel, with studies of man and nature*. Courier Corporation.

Warren BH, Simberloff D, Ricklefs RE, Aguilee R, Condamine FL, Gravel D, Morlon H, Mouquet N, Rosindell J, Casquet J, Conti E, Cornuault J, Fernandez-Palacios JM, Hengl T, Norder SJ, Rijsdijk KF, Sanmartin I, Strasberg D, Triantis KA, Valente LM, Whittaker RJ, Gillespie RG, Emerson BC, Thebaud C. 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters* **18:** 200-217.

Whittaker RJ, Fernández-Palacios JM. 2007. *Island biogeography: Ecology, evolution, and conservation*. OUP Oxford.

Woodroffe SA, Long AJ, Milne GA, Bryant CL, Thomas AL. 2015. New constraints on late

Holocene eustatic sea-level changes from Mahé, Seychelles. *Quaternary Science Reviews* **115:** 1-16.

Xia X. 2013. DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. *Molecular Biology and Evolution* **30:** 1720-1728.

Xia X, Xie Z, Salemi M, Chen L, Wang Y. 2003. An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution* **26:** 1-7.

Zhang J, Kapli P, Pavlidis P, Stamatakis A. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* **29:** 2869-2876.

Zoological Society of London. 2015. EDGE of Existence Programme. https://www.edgeofexistence.org/species/species-category/amphibians/ Accessed 14th October 2018.

Table 1 Sequenced gene fragments and summary statistics for analysed loci from *Sooglossus* and *Sechellophryne* spp. tissue samples. Mitochondrial DNA = *16S* rRNA (*16s*), cytochrome b (*cytb*); nuclear DNA = proopiomelanocortin (*pomc*), recombination activating genes (*rag*) 1 and 2, rhodopsin exon 1 (*rho*). Data incorporates *16s* sequence data obtained from GenBank (superscript denotes no. of GenBank samples included in total). Dash (-) indicates sequence data not obtained. *N*=sample size; bp=base pairs; Pi=parsimony informative sites; V=variable sites; π=nucleotide diversity.

Species	**Island**	***16s***	***cytb***	***pomc***	***rag1***	***rag2***	***rho***
So. sechellensis	Mahé	$76^{(6)}$	16	10	19	59	18
	Silhouette	21	12	9	15	20	13
	Praslin	$62^{(15)}$	11	11	25	51	19
So. thomasseti	Mahé	17	7	6	11	15	7
	Silhouette	12	2	2	10	12	2
Se. gardineri	Mahé	15	4	7	10	15	11
	Silhouette	12	-	1	3	10	3
Se. pipilodryas	Silhouette	$2^{(1)}$	1	1	$1^{(1)}$	$2^{(1)}$	1
	N	217	53	47	94	184	74
	bp	532	549	348	383	521	279
	Pi	111	154	31	22	57	13
	V	119	184	42	32	76	15
	π	0.0557	0.1113	0.0312	0.0169	0.0232	0.0043

Table 2 Between taxa *16s* distance matrix for the Sooglossidae. Lower diagonal: uncorrected *p*-distance; upper diagonal: corrected Jukes-Cantor *p*-distance (Jukes & Cantor, 1969). *Sechellophryne gardineri* = Sg; *Se. pipilodryas* = Sp; *Sooglossus sechellensis* = Ss; *So. thomasseti* = St.

	Sg	**Sp**	**Ss**	**St**
Sg		0.0704	0.1435	0.1346
Sp	0.0672		0.1555	0.1447
Ss	0.1306	0.1404		0.0608
St	0.1232	0.1316	0.0584	

Table 3 Between population *16s* *p*-distance distance matrix for the Sooglossidae. Lower diagonal: uncorrected *p*-distance; upper diagonal: corrected Jukes-Cantor *p*-distance (Jukes & Cantor, 1969). *Sechellophryne gardineri* = Sg; *Se. pipilodryas* = Sp; *Sooglossus sechellensis* = Ss; *So. thomasseti* = St. M = Mahé, S = Silhouette, P = Praslin.

	Sg-M	**Sg-S**	**Sp**	**Ss-M**	**Ss-P**	**Ss-S**	**St-M**	**St-S**
Sg-M		0.0373	0.0701	0.1491	0.1477	0.1479	0.1407	0.1243
Sg-S	0.0364		0.0708	0.1395	0.1339	0.1402	0.1383	0.1313
Sp	0.0669	0.0675		0.1546	0.1551	0.1599	0.1473	0.1410
Ss-M	0.1352	0.1273	0.1397		0.0465	0.0399	0.0577	0.0622
Ss-P	0.1341	0.1226	0.1401	0.0451		0.0205	0.0647	0.0623
Ss-S	0.1342	0.1279	0.1440	0.0389	0.0202		0.0556	0.0572
St-M	0.1283	0.1263	0.1337	0.0556	0.0620	0.0536		0.0209
St-S	0.1145	0.1205	0.1285	0.0597	0.0598	0.0551	0.0206	

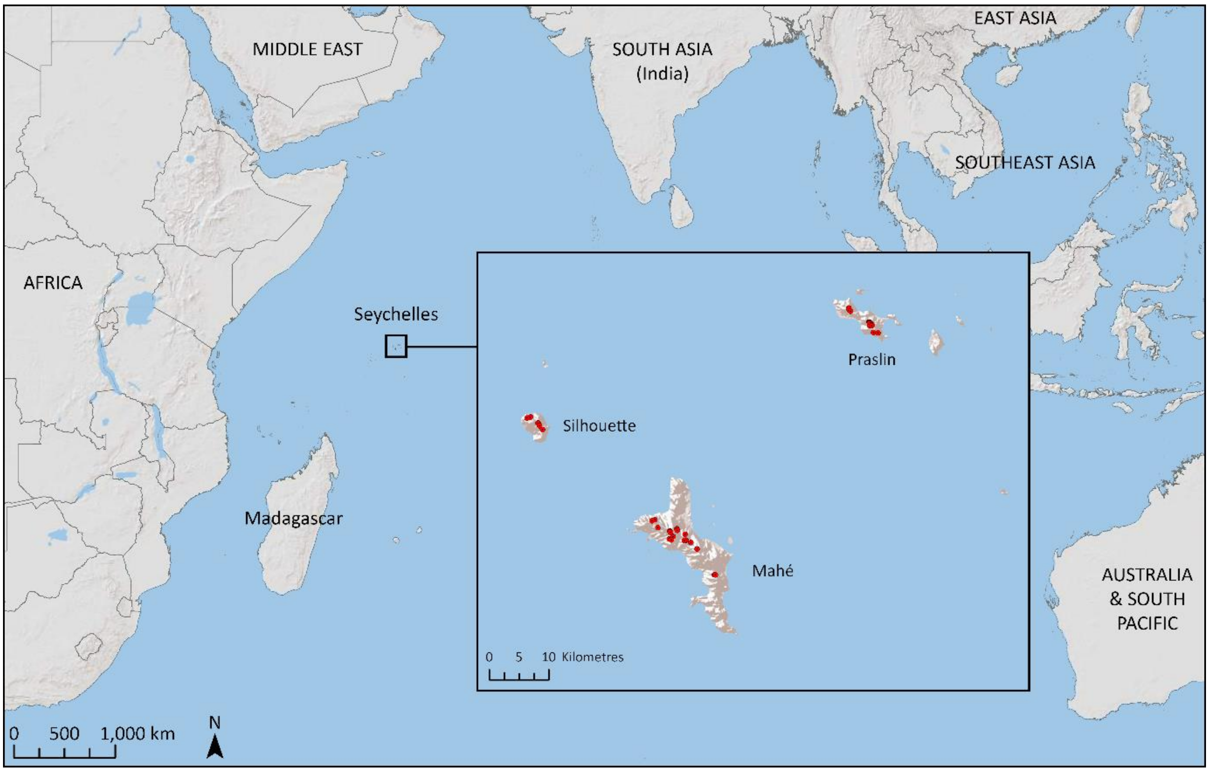

Figure 1 Seychelles archipelago and the surrounding geographic regions of the Indian Ocean, inset with the inner islands of Mahé, Praslin, and Silhouette—the only locations where the Sooglossidae (Noble, 1931) are found. *Sooglossus sechellensis* (Boettger, 1896), *So. thomasseti* (Boulenger, 1909), and *Sechellophryne gardineri* (Boulenger, 1911) are sympatric on Mahé and Silhouette, with the addition of *Se. pipilodryas* (Gerlach & Willi, 2002) on Silhouette. *Sooglossus sechellensis* is the only sooglossid to occur on Praslin. Red circles indicate sampling localities and associated geographic data used to test for the effects of isolation by distance.

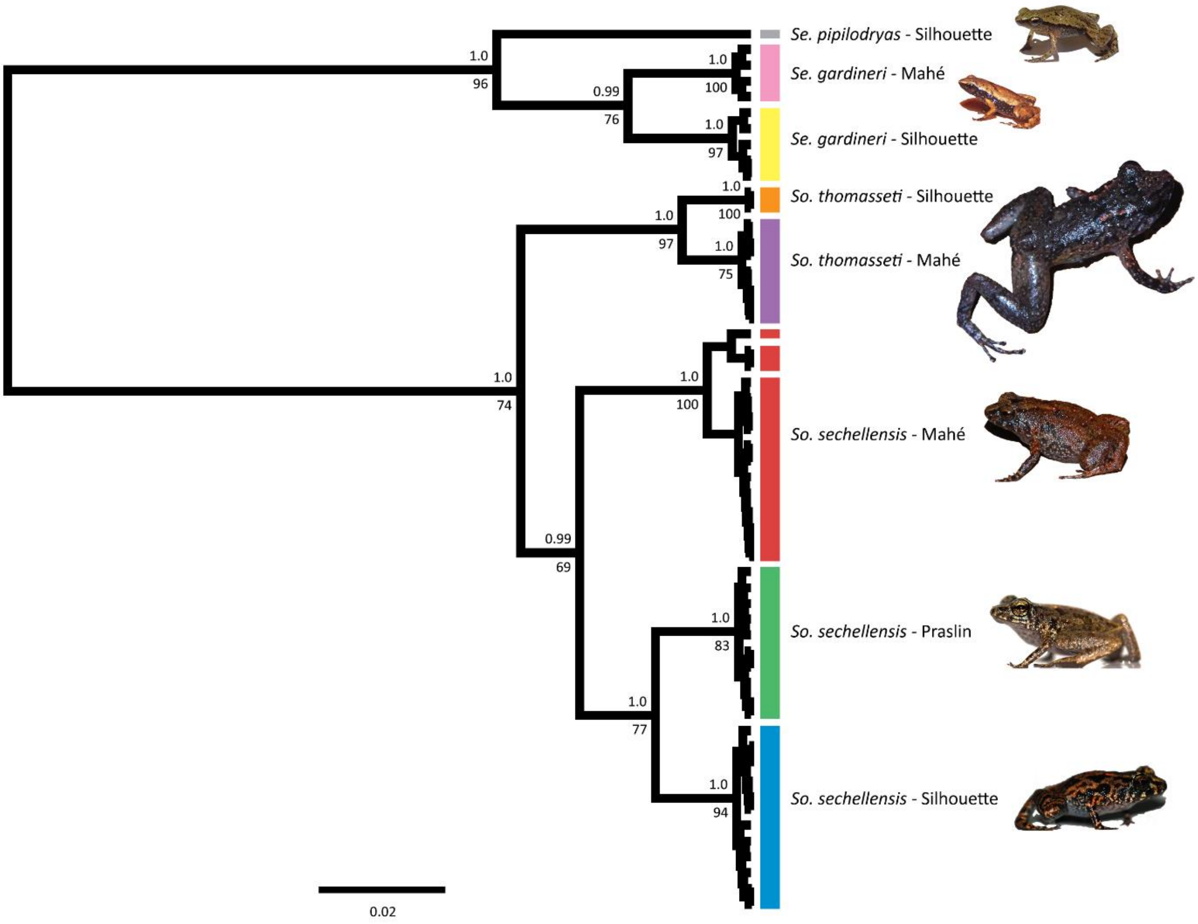

Figure 2 Bayesian inferred mitochondrial DNA phylogeny of Seychelles Sooglossidae. Support values are shown as Bayesian posterior probabilities (PP: above branches) and maximum likelihood bootstrap values (BS: below branches). Scale bar indicates substitutions per site. Vertical coloured bars adjacent to branch tips correspond to the ten population/species boundaries returned by the maximum likelihood partition in bPTP analysis. Colour coding identifies the island lineage of each species: *Sooglossus thomasseti* (Mahé – orange; Silhouette – purple) is the largest sooglossid; followed by *So. sechellensis* (Mahé – red; Praslin – green; Silhouette – blue), which is also the most widely geographically distributed; then *Sechellophryne pipilodryas* (Silhouette – grey); and the smallest in the family, *Sechellophryne gardineri* (Mahé – pink; Silhouette – yellow).

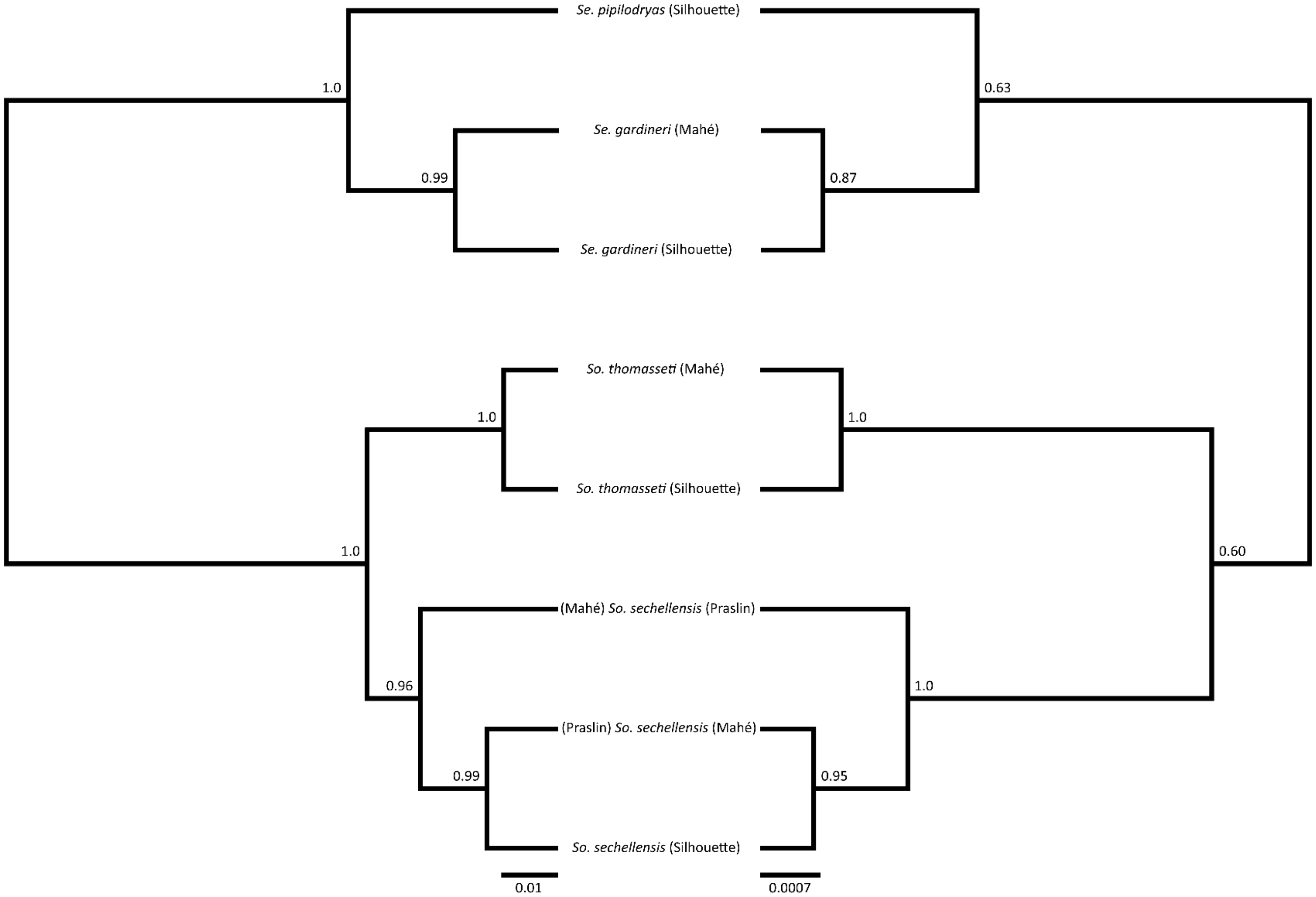

Figure 3 *BEAST generated mitochondrial (left) and nuclear (right) DNA species trees for the Sooglossidae. Branch numbers show PP support. The single topological disparity identifies Mahé *So. sechellensis* in the mtDNA species tree as sister to a clade comprised of those from Silhouette and Praslin, whereas in the nuDNA tree Silhouette and Mahé frogs form a clade sister to those from Praslin. Scale bar indicates substitutions per site.

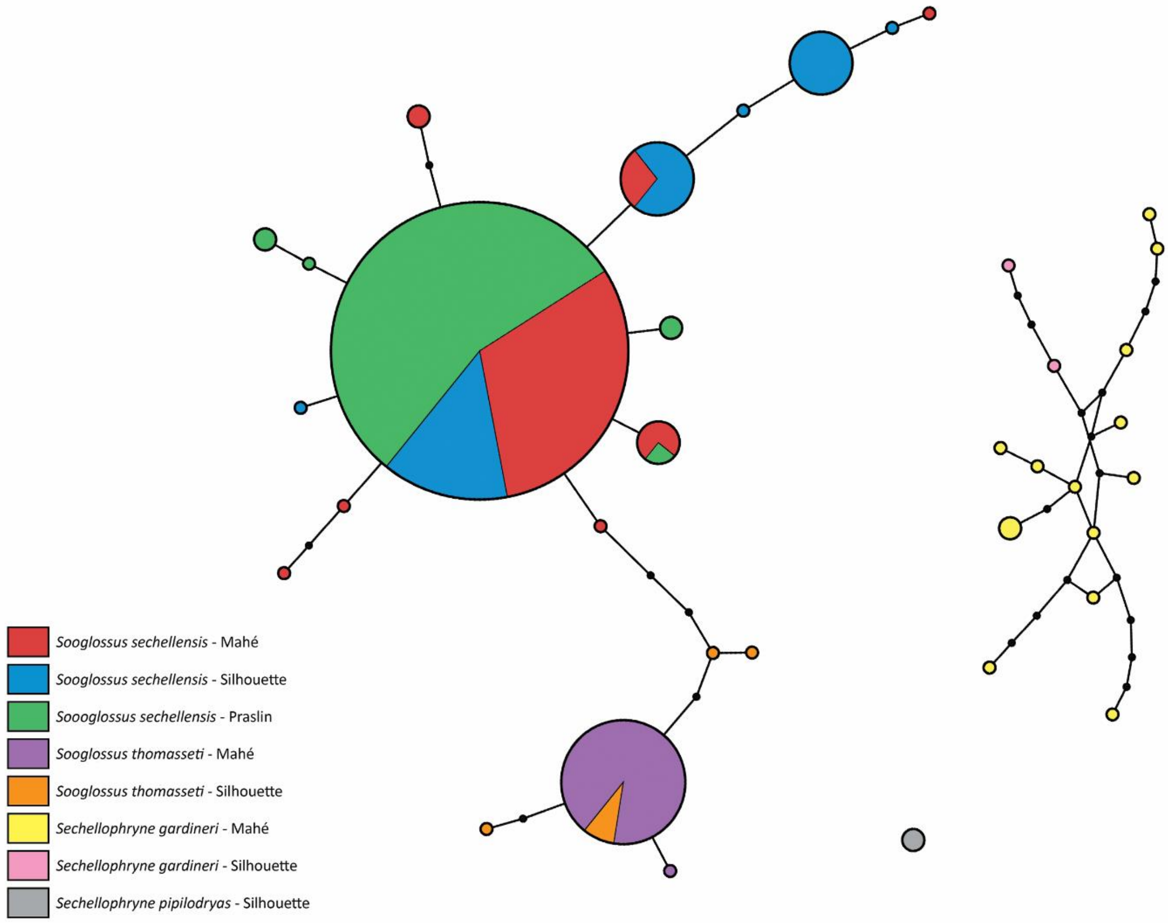

Figure 4 Nuclear *pomc* DNA haplotype network for the Sooglossidae. Thirty-six haplotypes are present. Circle size is proportional to the frequency with which the haplotype was observed, i.e. larger circles represent high-frequency, shared haplotypes, smaller circles represent low-frequency/rare haplotypes. Closed black circles indicate mutational steps. Colours represent island populations (see legend/Fig. 2).

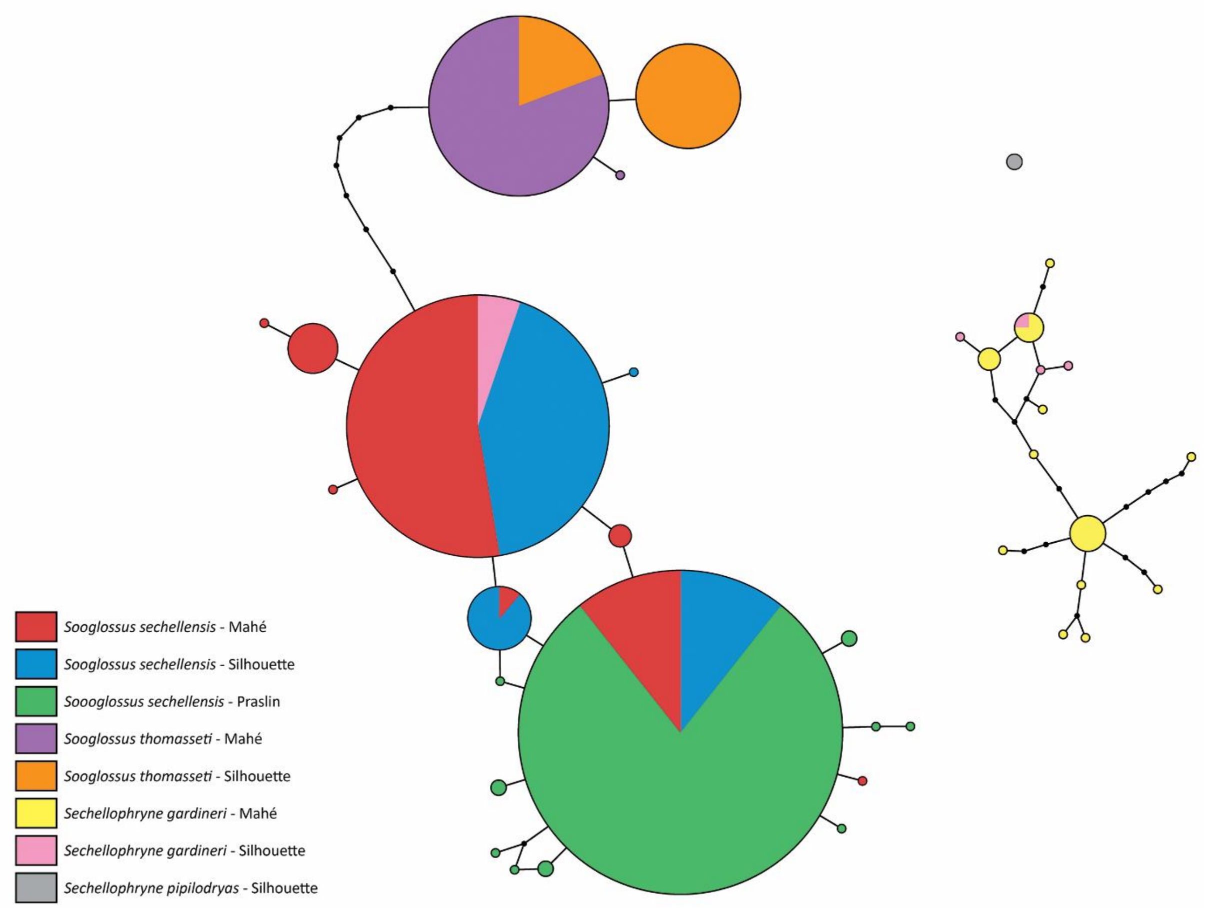

Figure 5 Nuclear *rag1* DNA haplotype network for the Sooglossidae. Thirty-seven haplotypes are present. Circle size is proportional to the frequency with which the haplotype was observed, i.e. larger circles represent high-frequency, shared haplotypes, smaller circles represent low-frequency/rare haplotypes. Closed black circles indicate mutational steps. Colours represent island populations; colour coding follows that of previous figures.

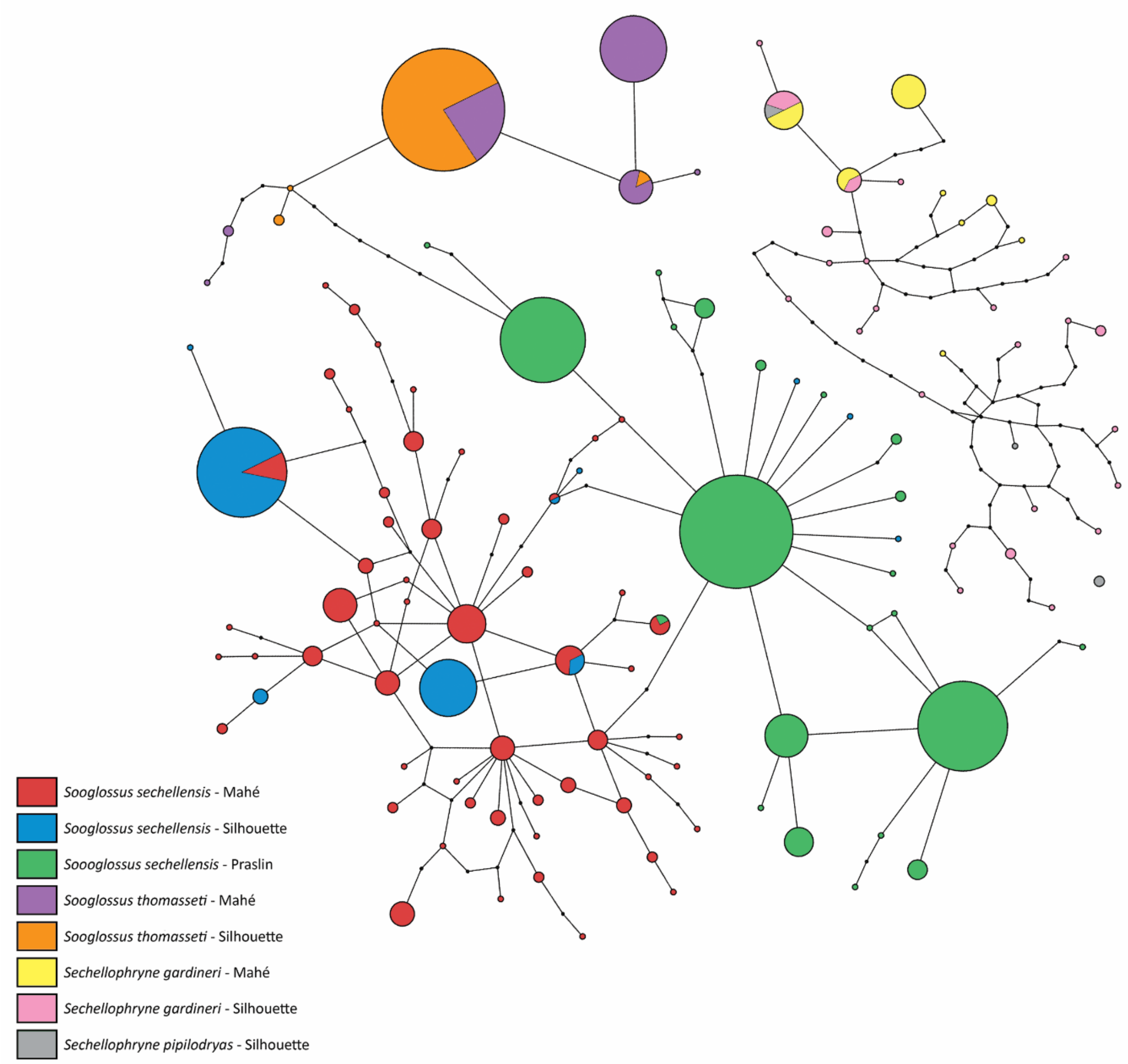

Figure 6 Nuclear *rag2* DNA haplotype network for the Sooglossidae. One-hundred and twenty-three haplotypes are present. Circle size is proportional to the frequency with which the haplotype was observed, i.e. larger circles represent high-frequency, shared haplotypes, smaller circles represent low-frequency/rare haplotypes. Closed black circles indicate mutational steps. Colours represent island populations; colour coding follows that of previous figures.

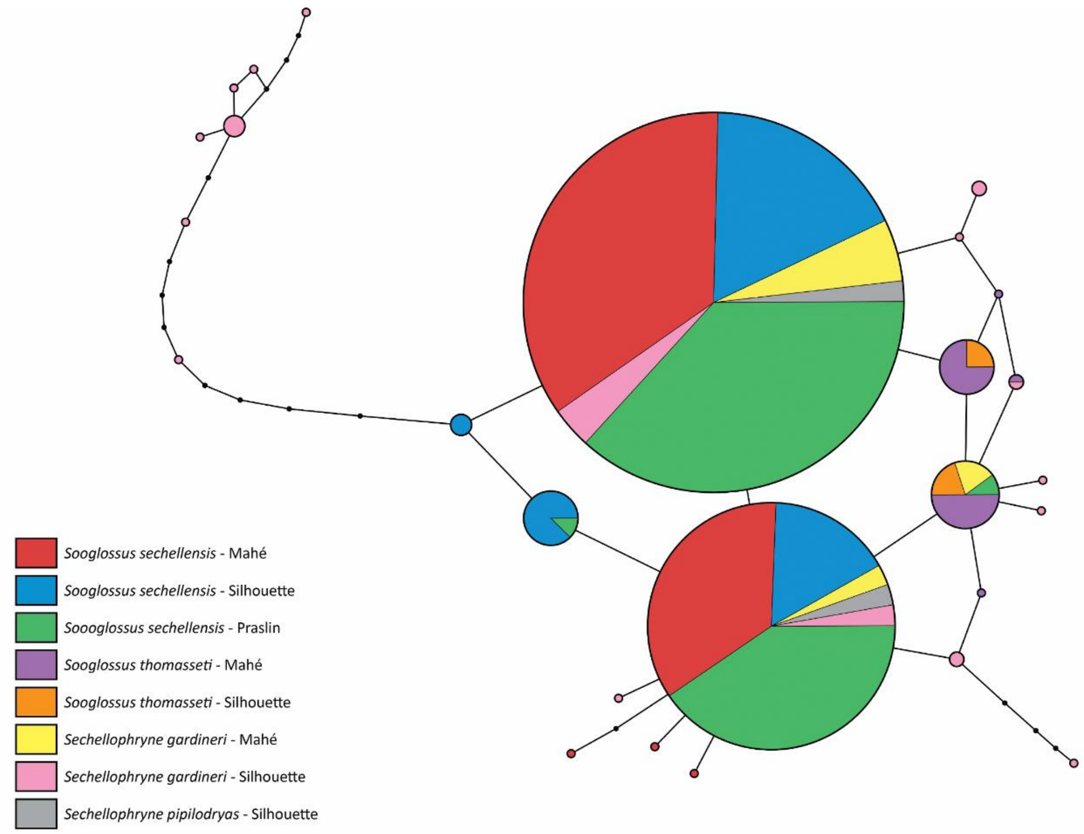

Figure 7 Nuclear *rho* DNA haplotype network for the Sooglossidae. Twenty-six haplotypes are present. Circle size is proportional to the frequency with which the haplotype was observed, i.e. larger circles represent high-frequency, shared haplotypes, smaller circles represent low-frequency/rare haplotypes. Closed black circles indicate mutational steps. Colours represent island populations; colour coding follows that of previous figures.

SUPPORTING INFORMATION

Appendix S1

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 µl. 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 µl reaction, reaction volumes consisted of 10.5 µl ddH_2O, 0.5 µl each of forward and reverse primer (at a concentration of 25 pmol/µl), 12.5 µl MyTaq HS Red mix™, and 1 µl 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')
16s	*16s* A-L[a]	CGC CTG TTT ATC AAA AAC AT
	16s B-H[a]	CCG GTC TGA ACT CAG ATC ACG T
cytb	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
pomc	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
rag1	RAG1B[d]	ATG GGA GAT GTG AGT GAR AAR CA
	RAG1E[d]	TCC GCT GCA TTT CCR ATG TCR CA
rag2	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
rho	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 sequence data used in this study. 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.

Species	***16s***	***cytb***	***rag1***	***rag2***
Ascaphus truei	AJ871087	AJ871087	-	-
Leiopelma archeyi	NC_014691	NC_014691	-	-
Sechellophryne pipilodryas	DQ872918	-	DQ872922	DQ872912
Sooglossus sechellensis	JF784361	-	-	-
Sooglossus sechellensis	JF784362	-	-	-
Sooglossus sechellensis	JF784363	-	-	-
Sooglossus sechellensis	JF784364	-	-	-
Sooglossus sechellensis	JF784365	-	-	-
Sooglossus sechellensis	JF784366	-	-	-
Sooglossus sechellensis	JF784367	-	-	-
Sooglossus sechellensis	JF784368	-	-	-
Sooglossus sechellensis	JF784370	-	-	-
Sooglossus sechellensis	JF784371	-	-	-
Sooglossus sechellensis	JF784372	-	-	-
Sooglossus sechellensis	JF784373	-	-	-
Sooglossus sechellensis	JF784374	-	-	-
Sooglossus sechellensis	JF784376	-	-	-
Sooglossus sechellensis	JF784377	-	-	-
Sooglossus sechellensis	JF784378	-	-	-
Sooglossus sechellensis	JF784379	-	-	-
Sooglossus sechellensis	JF784380	-	-	-
Sooglossus sechellensis	JF784381	-	-	-
Sooglossus sechellensis	JF784382	-	-	-
Sooglossus sechellensis	JF784383	-	-	-

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	*16s*, *cytb* (1)	GTR+I+G
	cytb (2)	TrN+I
	cytb (3)	TrN+G
nuDNA	*pomc* (1-3)	TrN+I+G
	rag1 (1-3)	TrN+I+G
	rag2 (1-3)	TrN+I+G
	rho (1-3)	TrN+I+G

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

Species	**Ref.**	**Locus**	**Composite**
Sechellophryne gardineri	JMSG07	*rho*	JMSG09
Sechellophryne pipilodryas	DQ872922	*rag1*	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
Sechellophryne pipilodryas	Silhouette	JMSP01	1.00
Sechellophryne gardineri	Mahé	CDSG01, MBSG04, LRSG03, MBSG02	0.97
Sechellophryne gardineri	Silhouette	DGSG01, JMSG01, JMSG05, JMSG07, JMSG10	0.75
Sooglossus thomasseti	Silhouette	GBST01, JMST06	0.99
Sooglossus thomasseti	Mahé	CDST02, CRST01, MCST02, LMST01, MSST01, CDST01, MBST01	0.98
Sooglossus sechellensis	Mahé	LRSS14	0.95
Sooglossus sechellensis	Mahé	LRSS01, MSSS02	0.95
Sooglossus sechellensis	Mahé	RSS01, LRSS02, LMSS01, SFSS02, CRSS14, CSS01, MSSS01, MCSS10, MBSS07, CRSS02, CRSS01, MBSS01	0.97
Sooglossus sechellensis	Praslin	VMSP16, CMSP01, CMSP07, CMSP02, FPSP03, FAT2, ZSP01, ZSP04, ZSP03, ZSP08	0.98
Sooglossus sechellensis	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 $P < 0.02$.

	Sooglossus sechellensis			***Sooglossus thomasseti***			***Sechellophryne gardineri***		
	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
16s	2.36870	3.278	0.1621	3.10581	12.422	0.2512	2.24857	3.284	0.2073
	$P < 0.05$	$P > 0.02$	$P > 0.05$	$P < 0.01$	$P > 0.02$	$P > 0.05$	$P < 0.05$	$P > 0.02$	$P > 0.05$
cytb	1.71837	-1.421	0.1910	0.44661	3.394	0.1967	-0.55827	-0.361	0.1061
	$P > 0.05$	$P > 0.02$	$P > 0.05$	$P > 0.05$	$P > 0.02$	$P > 0.05$	$P > 0.05$	$P > 0.02$	$P < 0.05$
pomc	-1.40180	-8.378	0.0560	-1.21781	-1.557	0.0963	-0.64112	-10.089	0.1300
	$P > 0.05$	$P < 0.02$	$P > 0.05$	$P > 0.05$	$P > 0.02$	$P < 0.01$	$P > 0.05$	$P < 0.02$	$P > 0.05$
rag1	-1.21313	-7.542	0.0534	0.21337	0.346	0.1401	-0.13712	-0.421	0.1331
	$P > 0.05$	$P < 0.02$	$P > 0.05$	$P > 0.05$	$P > 0.02$	$P > 0.05$	$P > 0.05$	$P > 0.02$	$P > 0.05$
rag2	-1.77022	-82.555	0.0335	-0.12593	-1.420	0.1044	-0.65881	-8.246	0.0910
	$P < 0.05$	$P < 0.02$	$P < 0.05$	$P > 0.05$	$P > 0.02$	$P > 0.05$	$P > 0.05$	$P = 0.02$	$P > 0.05$
rho	-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*).

	Sooglossus sechellensis			***Sooglossus thomasseti***		***Sechellophryne gardineri***	
Island	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]

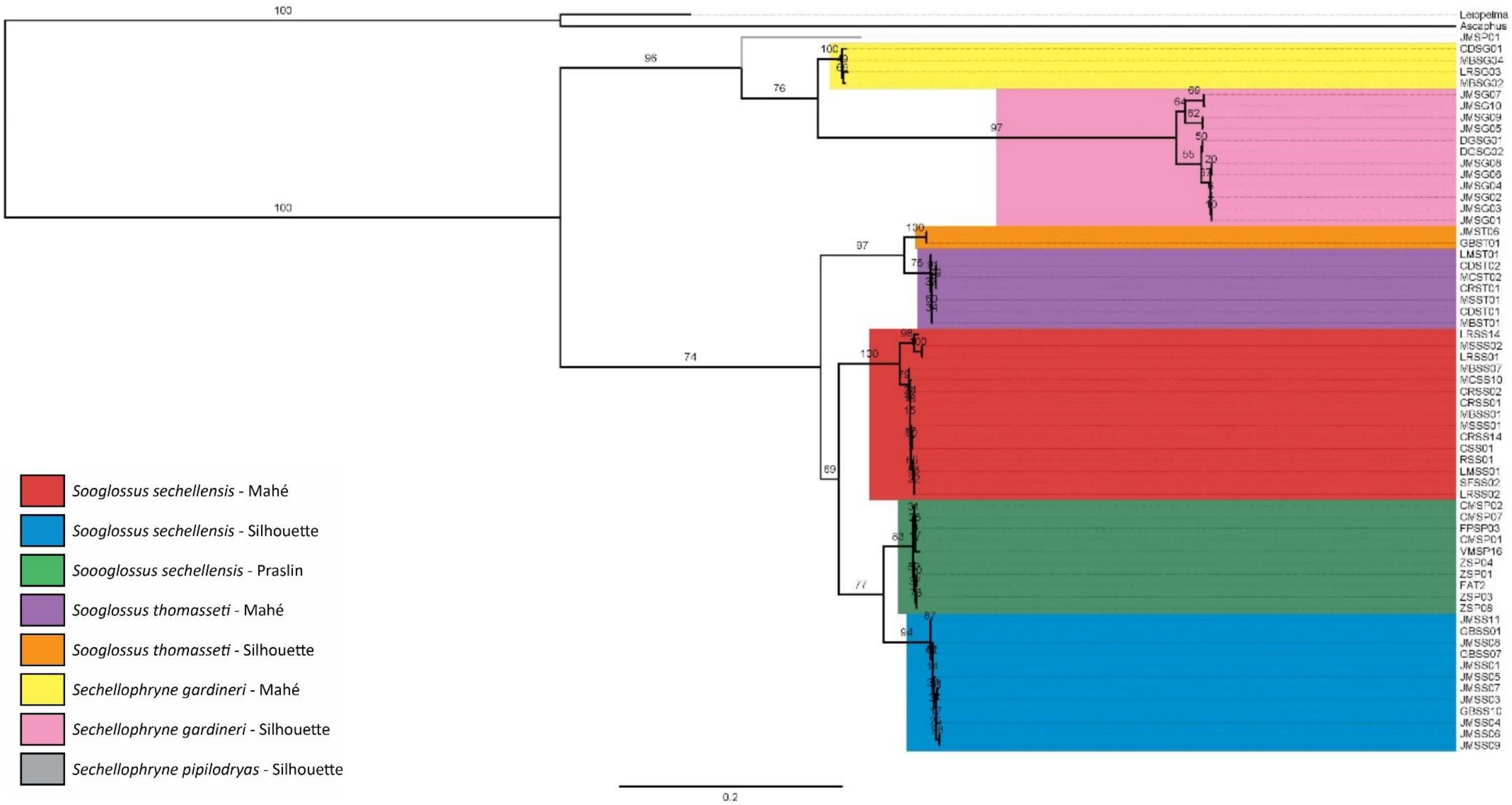

Figure S1.1 Maximum likelihood inferred mitochondrial DNA phylogeny of the Sooglossidae. Leiopelmatoidea (*Leiopelma*+*Ascaphus*) rooted outgroup. Branch support is indicated by maximum likelihood bootstrap (BS) values. Scale bar indicates substitutions per site.

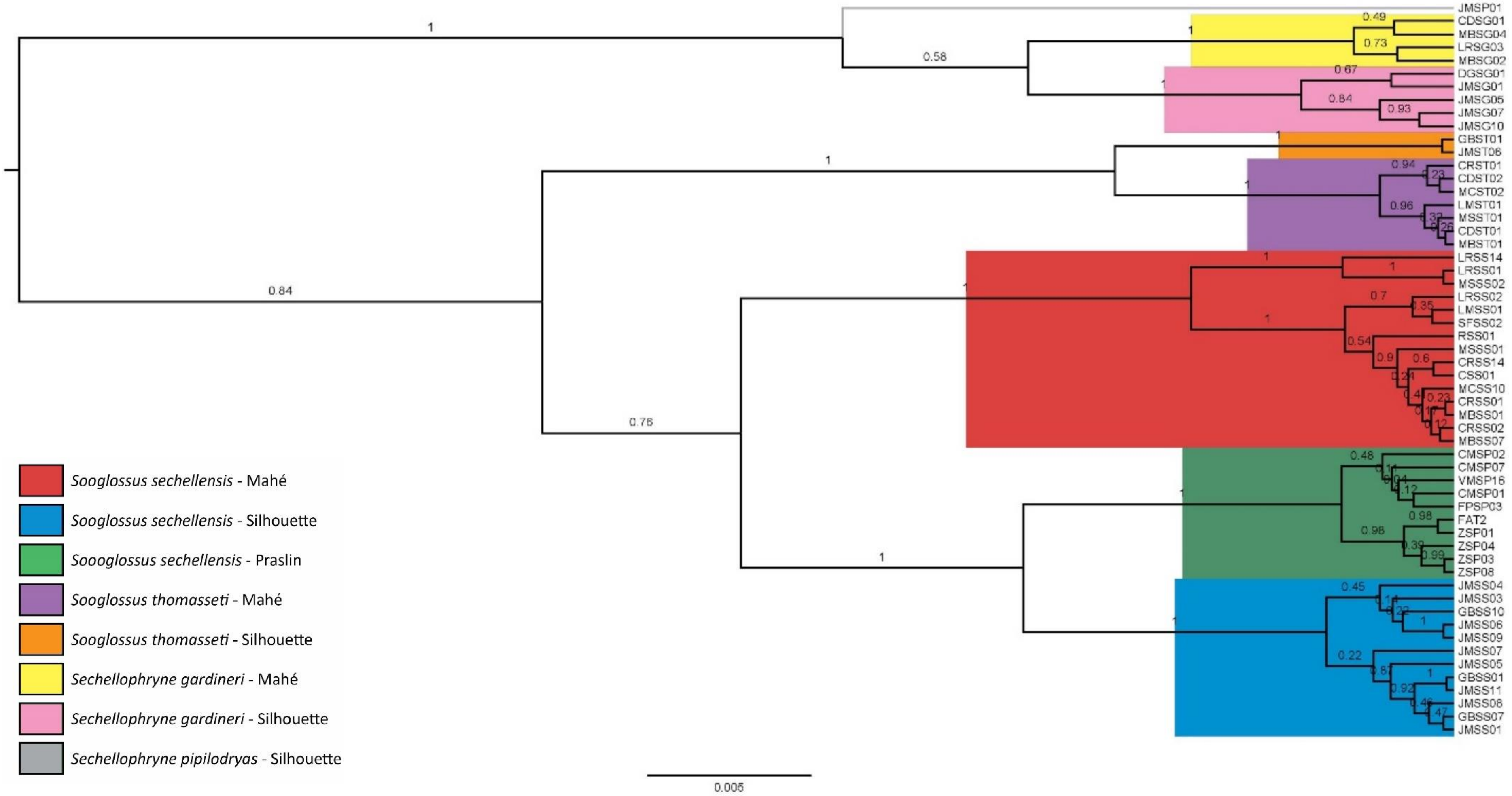

Figure S1.2 Bayesian inferred mitochondrial DNA phylogeny of the Sooglossidae using the Yule tree prior in BEAST2. Branch support is indicated by Bayesian Posterior Probabilities (PP). Scale bar indicates substitutions per site.

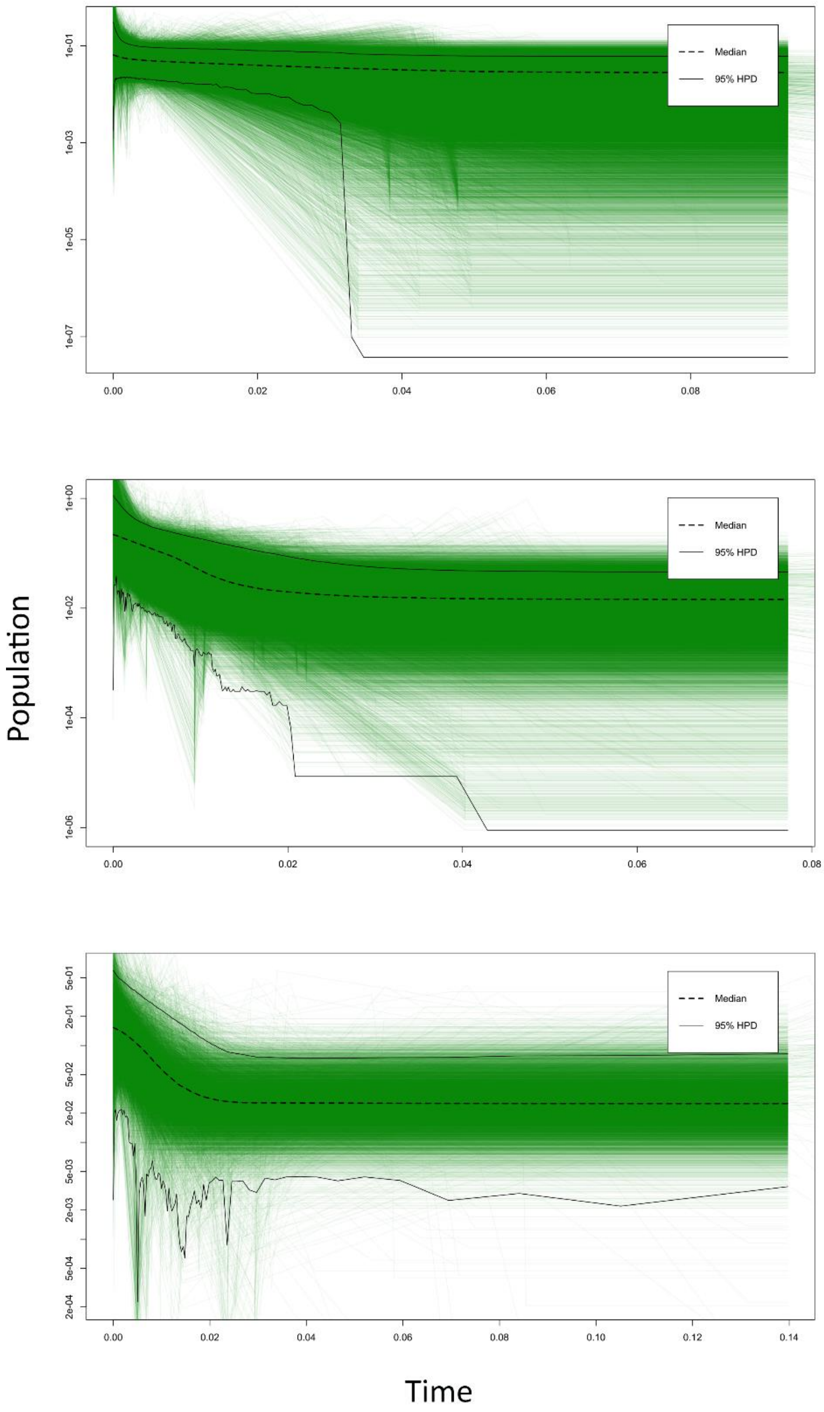

Figure S1.3 Extended Bayesian Skyline Plots of population size through time for *Sooglossus sechellensis*. The full view of the posterior all of the samples that are summarised by the median and 95% HPD interval are shown for the Mahé (top), Praslin (centre), and Silhouette (bottom) populations. The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

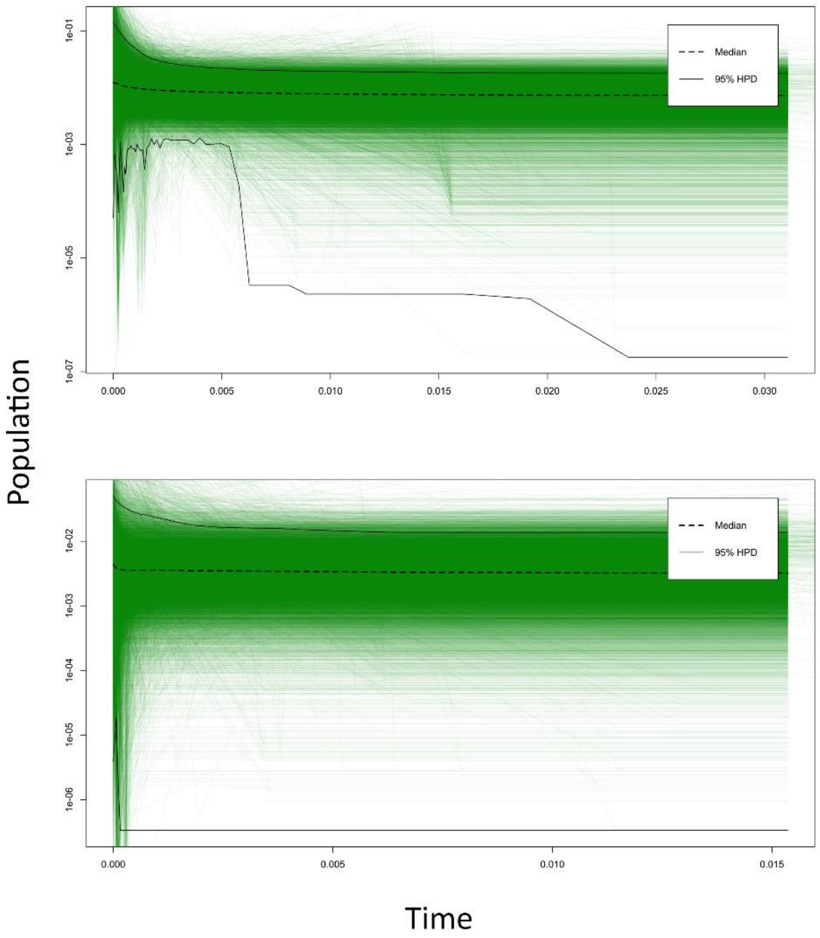

Figure S1.4 Extended Bayesian Skyline Plots of population size through time for *Sooglossus thomasseti*. The full view of the posterior all of the samples that are summarised by the median and 95% HPD interval are shown for the Mahé (top) and Silhouette (bottom) populations. EBSP analyses comprised all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

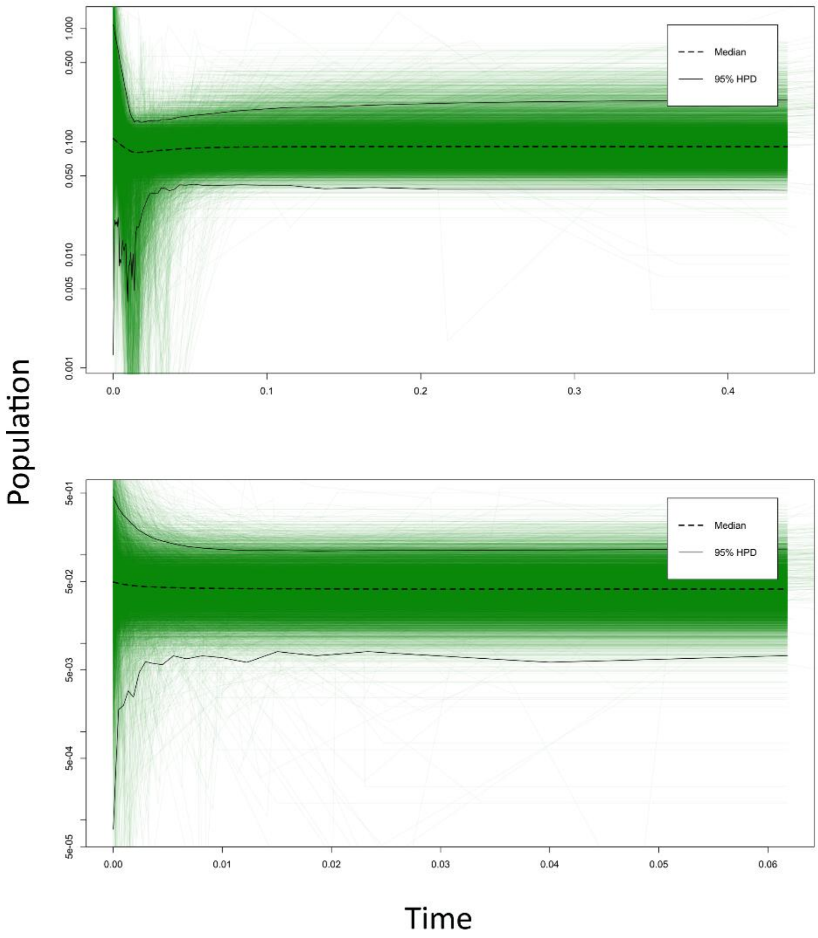

Figure S1.5 Extended Bayesian Skyline Plots of population size through time for *Sechellophryne gardineri*. The full view of the posterior all of the samples that are summarised by the median and 95% HPD interval are shown for the Mahé (top) and Silhouette (bottom) populations. EBSP analyses of the Mahé population comprised all six loci. Analyses of the Silhouette population comprised two loci (*16s*, *rag2*). Time on x-axis in millions of years. Population size on y-axis in millions of years assuming a generation time of one 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, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30:** 2725-2729.