
DOI

https://doi.org/10.1093/biolinnean/bly183

Link to record in KAR

https://kar.kent.ac.uk/71727/

Document Version

Author's Accepted Manuscript

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version. Users are advised to check http://kar.kent.ac.uk for the status of the paper. Users should always cite the published version of record.

Enquiries

For any further enquiries regarding the licence status of this document, please contact: researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at http://kar.kent.ac.uk/contact.html
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

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 (Adersen, 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 km² (Davies & Francis, 1964). Its flat upper section spans ca. 44,000 km² and lies an average depth of 55 m below present sea level (bpl), 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 km², 905 m elevation), Praslin (38 km², 367 m elevation) and Silhouette (20 km², 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.
DATA CONVERT 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. SEQUENCE MATRIX 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 PARTITION FINDER 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. Leiopelmoidea (*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 inference 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 evaluated using PP values. Trees were visualised using FIGTREE.

To infer population boundaries and aid the identification 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 $\geq$ 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. 297).
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 km², and been submerged to its present extent,
comprising an archipelago of 45 inner-islands covering ~247 km². 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


Gene Codes Corporation. 2015. Sequencher® version 5.3 sequence analysis software. Ann
Arbor, MI USA.


Hoelzer GA. 1997. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees


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.


**Rambaut A. 2016.** FigTree 1.4.3.


**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.


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


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


Table 1 Sequenced gene fragments and summary statistics for analysed loci from *Sooglossus* and *Sechelophryne* 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; $\pi$=nucleotide diversity.

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

<table>
<thead>
<tr>
<th></th>
<th>Sg</th>
<th>Sp</th>
<th>Ss</th>
<th>St</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg</td>
<td>0.0704</td>
<td>0.1435</td>
<td>0.1346</td>
<td></td>
</tr>
<tr>
<td>Sp</td>
<td>0.0672</td>
<td>0.1555</td>
<td>0.1447</td>
<td></td>
</tr>
<tr>
<td>Ss</td>
<td>0.1306</td>
<td>0.1404</td>
<td>0.0608</td>
<td></td>
</tr>
<tr>
<td>St</td>
<td>0.1232</td>
<td>0.1316</td>
<td>0.0584</td>
<td></td>
</tr>
</tbody>
</table>
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). Sechellophyne gardineri = Sg; Se. pipilodryas = Sp; Sooglossus sechellensis = Ss; So. thomasseti = St. M = Mahé, S = Silhouette, P = Praslin.

<table>
<thead>
<tr>
<th></th>
<th>Sg-M</th>
<th>Sg-S</th>
<th>Sp</th>
<th>Ss-M</th>
<th>Ss-P</th>
<th>Ss-S</th>
<th>St-M</th>
<th>St-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg-M</td>
<td>0.0373</td>
<td>0.0701</td>
<td>0.1491</td>
<td>0.1477</td>
<td>0.1479</td>
<td>0.1407</td>
<td>0.1243</td>
<td></td>
</tr>
<tr>
<td>Sg-S</td>
<td>0.0364</td>
<td>0.0708</td>
<td>0.1395</td>
<td>0.1339</td>
<td>0.1402</td>
<td>0.1383</td>
<td>0.1313</td>
<td></td>
</tr>
<tr>
<td>Sp</td>
<td>0.0669</td>
<td>0.0675</td>
<td>0.1546</td>
<td>0.1551</td>
<td>0.1599</td>
<td>0.1473</td>
<td>0.1410</td>
<td></td>
</tr>
<tr>
<td>Ss-M</td>
<td>0.1352</td>
<td>0.1273</td>
<td>0.1397</td>
<td>0.0465</td>
<td>0.0399</td>
<td>0.0577</td>
<td>0.0622</td>
<td></td>
</tr>
<tr>
<td>Ss-P</td>
<td>0.1341</td>
<td>0.1226</td>
<td>0.1401</td>
<td>0.0451</td>
<td>0.0205</td>
<td>0.0647</td>
<td>0.0623</td>
<td></td>
</tr>
<tr>
<td>Ss-S</td>
<td>0.1342</td>
<td>0.1279</td>
<td>0.1440</td>
<td>0.0389</td>
<td>0.0202</td>
<td>0.0556</td>
<td>0.0572</td>
<td></td>
</tr>
<tr>
<td>St-M</td>
<td>0.1283</td>
<td>0.1263</td>
<td>0.1337</td>
<td>0.0556</td>
<td>0.0620</td>
<td>0.0536</td>
<td>0.0209</td>
<td></td>
</tr>
<tr>
<td>St-S</td>
<td>0.1145</td>
<td>0.1205</td>
<td>0.1285</td>
<td>0.0597</td>
<td>0.0598</td>
<td>0.0551</td>
<td>0.0206</td>
<td></td>
</tr>
</tbody>
</table>
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 pomo 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 ddH2O, 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.

<table>
<thead>
<tr>
<th>Gene fragment</th>
<th>Primer</th>
<th>Sequence (5’ – 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16s</td>
<td>16s A-L&lt;sup&gt;a&lt;/sup&gt;</td>
<td>CGC CTG TTT ATC AAA AAC AT</td>
</tr>
<tr>
<td></td>
<td>16s B-H&lt;sup&gt;a&lt;/sup&gt;</td>
<td>CCG GTC TGA ACT CAG ATC ACG T</td>
</tr>
<tr>
<td>cytb</td>
<td>CBJ 10933&lt;sup&gt;b&lt;/sup&gt;</td>
<td>TAT GTT CTA CCA TGA GGA CAA ATA TC</td>
</tr>
<tr>
<td></td>
<td>Cytb-c&lt;sup&gt;b&lt;/sup&gt;</td>
<td>CTA CTG GTT GTC CTC CGA TTC ATG T</td>
</tr>
<tr>
<td></td>
<td>CytbJL1f&lt;sup&gt;c&lt;/sup&gt;</td>
<td>TAG ACC TCC CAA CCC CAT CC</td>
</tr>
<tr>
<td></td>
<td>CytbJL1r&lt;sup&gt;c&lt;/sup&gt;</td>
<td>GAG GTG TGT GTT AGT GGG GG</td>
</tr>
<tr>
<td></td>
<td>CytbSGJL1f&lt;sup&gt;c&lt;/sup&gt;</td>
<td>ACC GCT TTC GTA GGC TAT GT</td>
</tr>
<tr>
<td></td>
<td>CytbSGJL1c&lt;sup&gt;c&lt;/sup&gt;</td>
<td>GTG GAC GAA ATG ATA TTG CTC GT</td>
</tr>
<tr>
<td>pomc</td>
<td>POMCJLf&lt;sup&gt;c&lt;/sup&gt;</td>
<td>GAC ATC GCC AAC TAT CCG GT</td>
</tr>
<tr>
<td></td>
<td>POMCJLr&lt;sup&gt;c&lt;/sup&gt;</td>
<td>AAG TGT TGT CCC CCG TGT TT</td>
</tr>
<tr>
<td></td>
<td>POMCJL2f&lt;sup&gt;c&lt;/sup&gt;</td>
<td>AAA CAC GGG GGA CAA CAC TT</td>
</tr>
<tr>
<td></td>
<td>POMCJL2r&lt;sup&gt;c&lt;/sup&gt;</td>
<td>CTT CTG AGT CGA CAC CAG GG</td>
</tr>
<tr>
<td>rag1</td>
<td>RAG1B&lt;sup&gt;d&lt;/sup&gt;</td>
<td>ATG GGA GAT GTG AGT GAR AAR CA</td>
</tr>
<tr>
<td></td>
<td>RAG1E&lt;sup&gt;d&lt;/sup&gt;</td>
<td>TCC GCT GCA TTT CCR ATG TCR CA</td>
</tr>
<tr>
<td>rag2</td>
<td>RAG2 JG1-F&lt;sup&gt;c&lt;/sup&gt;</td>
<td>TCG TCC TAC CAT GTT CAC CAA TGA GT</td>
</tr>
<tr>
<td></td>
<td>RAG2 JG1-R&lt;sup&gt;c&lt;/sup&gt;</td>
<td>TCC TGT CCA ATC AGG CAG TTC CA</td>
</tr>
<tr>
<td></td>
<td>RAG2JLSG1f&lt;sup&gt;c&lt;/sup&gt;</td>
<td>CCA GCA GTG ACC AGC ATC TT</td>
</tr>
<tr>
<td></td>
<td>RAG2JLSG1r&lt;sup&gt;c&lt;/sup&gt;</td>
<td>CGC TGT CTC TTG GAC TGG TT</td>
</tr>
<tr>
<td></td>
<td>RAG2JLSG2r&lt;sup&gt;c&lt;/sup&gt;</td>
<td>CCG ACA ATG AGG AAC TCG CT</td>
</tr>
<tr>
<td>rho</td>
<td>Rhod1A&lt;sup&gt;e&lt;/sup&gt;</td>
<td>ACC ATG AAC GGA ACA GAA GGY CC</td>
</tr>
<tr>
<td></td>
<td>Rhod1D&lt;sup&gt;e&lt;/sup&gt;</td>
<td>GTA GCG AAG AAR CCT TCA AMG TA</td>
</tr>
</tbody>
</table>

<sup>a</sup> Palumbi <i>et al.</i>, (1991)

<sup>b</sup> Chiari <i>et al.</i>, (2004)

<sup>c</sup> Developed for this study

<sup>d</sup> Biju & Bossuyt, (2003)

<sup>e</sup> 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.

<table>
<thead>
<tr>
<th>Species</th>
<th>16s</th>
<th>cytb</th>
<th>rag1</th>
<th>rag2</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Ascaphus truei</em></td>
<td>AJ871087</td>
<td>AJ871087</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Leiopelma archeyi</em></td>
<td>NC_014691</td>
<td>NC_014691</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sechellophryne pipilodryas</em></td>
<td>DQ872918</td>
<td>-</td>
<td>DQ872922</td>
<td>DQ872912</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784361</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784362</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784363</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784364</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784365</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784366</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784367</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784368</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784370</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784371</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784372</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784373</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784374</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784376</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784377</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784378</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784379</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784380</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784381</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784382</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>JF784383</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Partitioning scheme</th>
<th>Substitution model</th>
</tr>
</thead>
<tbody>
<tr>
<td>mtDNA</td>
<td></td>
</tr>
<tr>
<td>16s, cytb (1)</td>
<td>GTR+I+G</td>
</tr>
<tr>
<td>cytb (2)</td>
<td>TrN+I</td>
</tr>
<tr>
<td>cytb (3)</td>
<td>TrN+G</td>
</tr>
<tr>
<td>nuDNA</td>
<td></td>
</tr>
<tr>
<td>pomc (1-3)</td>
<td>TrN+I+G</td>
</tr>
<tr>
<td>rag1 (1-3)</td>
<td>TrN+I+G</td>
</tr>
<tr>
<td>rag2 (1-3)</td>
<td>TrN+I+G</td>
</tr>
<tr>
<td>rho (1-3)</td>
<td>TrN+I+G</td>
</tr>
</tbody>
</table>
Table S1.4 Taxa used as composites in *BEAST analyses.

<table>
<thead>
<tr>
<th>Species</th>
<th>Ref.</th>
<th>Locus</th>
<th>Composite</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Sechellophryne gardineri</em></td>
<td>JMSG07</td>
<td>rho</td>
<td>JMSG09</td>
</tr>
<tr>
<td><em>Sechellophryne pipilodryas</em></td>
<td>DQ872922</td>
<td>rag1</td>
<td>JMSP01</td>
</tr>
</tbody>
</table>
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).

<table>
<thead>
<tr>
<th>Species/population</th>
<th>Island</th>
<th>Sample reference</th>
<th>PP</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Sechellophryne pipilodryas</em></td>
<td>Silhouette</td>
<td>JMSP01</td>
<td>1.00</td>
</tr>
<tr>
<td><em>Sechellophryne gardineri</em></td>
<td>Mahé</td>
<td>CDSG01, MBSG04, LRSG03, MBSG02</td>
<td>0.97</td>
</tr>
<tr>
<td><em>Sechellophryne gardineri</em></td>
<td>Silhouette</td>
<td>DGSG01, JMSG01, JMSG05, JMSG07, JMSG10</td>
<td>0.75</td>
</tr>
<tr>
<td><em>Sooglossus thomasseti</em></td>
<td>Silhouette</td>
<td>GBST01, JMST06</td>
<td>0.99</td>
</tr>
<tr>
<td><em>Sooglossus thomasseti</em></td>
<td>Mahé</td>
<td>CDST02, CRST01, MCST02, LMST01, MSST01, CDST01, MBST01</td>
<td>0.98</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>Mahé</td>
<td>LRSS14</td>
<td>0.95</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>Mahé</td>
<td>LRSS01, MSSS02</td>
<td>0.95</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>Mahé</td>
<td>RSS01, LRSS02, LMSS01, SFSS02, CRSS14, CSS01, MSSS01, MCSS10, MBSS07, CRSS02, CRSS01, MBSS01</td>
<td>0.97</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>Praslin</td>
<td>VMSP16, CMSP01, CMSP07, CMSP02, FPSP03, FAT2, ZSP01, ZSP04, ZSP03, ZSP08</td>
<td>0.98</td>
</tr>
<tr>
<td><em>Sooglossus sechellensis</em></td>
<td>Silhouette</td>
<td>JMSS05, GBSS01, JMSS11, JMSS08, GBSS07, JMSS01, JMSS04, JMSS03, JMSS07, GBSS10, JMSS06, JMSS09</td>
<td>0.98</td>
</tr>
</tbody>
</table>
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 $P < 0.02$.

<table>
<thead>
<tr>
<th></th>
<th>Sooglossus sechellensis</th>
<th>Sooglossus thomasseti</th>
<th>Sechellophryne gardineri</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tajima’s D</td>
<td>Fu’s $F_S$</td>
<td>$R_2$</td>
</tr>
<tr>
<td>16s</td>
<td>2.36870</td>
<td>3.278</td>
<td>0.1621</td>
</tr>
<tr>
<td></td>
<td>$P &lt; 0.05$</td>
<td>$P &gt; 0.02$</td>
<td>$P &gt; 0.05$</td>
</tr>
<tr>
<td>cytb</td>
<td>1.71837</td>
<td>-1.421</td>
<td>0.1910</td>
</tr>
<tr>
<td></td>
<td>$P &gt; 0.05$</td>
<td>$P &gt; 0.02$</td>
<td>$P &gt; 0.05$</td>
</tr>
<tr>
<td>pomc</td>
<td>-1.40180</td>
<td>-8.378</td>
<td>0.0560</td>
</tr>
<tr>
<td></td>
<td>$P &gt; 0.05$</td>
<td>$P &lt; 0.02$</td>
<td>$P &gt; 0.05$</td>
</tr>
<tr>
<td>rag1</td>
<td>-1.21313</td>
<td>-7.542</td>
<td>0.0534</td>
</tr>
<tr>
<td></td>
<td>$P &gt; 0.05$</td>
<td>$P &lt; 0.02$</td>
<td>$P &gt; 0.05$</td>
</tr>
<tr>
<td>rag2</td>
<td>-1.77022</td>
<td>-82.555</td>
<td>0.0335</td>
</tr>
<tr>
<td></td>
<td>$P &lt; 0.05$</td>
<td>$P &lt; 0.02$</td>
<td>$P &gt; 0.05$</td>
</tr>
<tr>
<td>rho</td>
<td>-1.37952</td>
<td>-1.467</td>
<td>0.1000</td>
</tr>
<tr>
<td></td>
<td>$P &gt; 0.05$</td>
<td>$P &gt; 0.02$</td>
<td>$P &gt; 0.05$</td>
</tr>
</tbody>
</table>
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*).

<table>
<thead>
<tr>
<th>Island</th>
<th><em>Sooglossus sechellensis</em></th>
<th><em>Sooglossus thomasseti</em></th>
<th><em>Sechellophryne gardineri</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mahé</td>
<td>Praslin</td>
<td>Silhouette</td>
</tr>
<tr>
<td>Chain length</td>
<td>$3 \times 10^8$</td>
<td>$2 \times 10^8$</td>
<td>$5 \times 10^7$</td>
</tr>
<tr>
<td>EBSP</td>
<td>[0, 3]</td>
<td>[1, 3]$^+$</td>
<td>[0, 3]</td>
</tr>
</tbody>
</table>
Figure S1.1 Maximum likelihood inferred mitochondrial DNA phylogeny of the Sooglossidae. Leiopelmatidae (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


