doi:10.1111/evo.12623

Reed frog diversification in the Gulf of Guinea: Overseas dispersal, the progression rule, and in situ speciation

Rayna C. Bell,[1,2] Robert C. Drewes,[3] and Kelly R. Zamudio[1]

[1]*Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York 14853–2701*

[2]*E-mail: raynabell@berkeley.edu*

[3]*Herpetology California Academy of Sciences, San Francisco, California 94118*

Received August 22, 2014
Accepted January 3, 2015

Oceanic islands accumulate endemic species when new colonists diverge from source populations or by in situ diversification of resident island endemics. The relative importance of dispersal versus in situ speciation in generating diversity on islands varies with a number of archipelago characteristics including island size, age, and remoteness. Here, we characterize interisland dispersal and in situ speciation in frogs endemic to the Gulf of Guinea islands. Using mitochondrial sequence and genome-wide single-nucleotide polymorphism data, we demonstrate that dispersal proceeded from the younger island (São Tomé) to the older island (Príncipe) indicating that for organisms that disperse overseas on rafts, dispersal between islands may be determined by ocean currents and not island age. We find that dispersal between the islands is not ongoing, resulting in genotypically distinct but phenotypically similar lineages on the two islands. Finally, we demonstrate that in situ diversification on São Tomé Island likely proceeded in allopatry due to the geographic separation of breeding sites, resulting in phenotypically distinct species. We find evidence of hybridization between the species where their ranges are sympatric and the hybrid zone coincides with a transition from agricultural land to primary forest, indicating that anthropogenic development may have facilitated secondary contact between previously allopatric species.

KEY WORDS: **Hybridization, *Hyperolius*, phylogenomics, population genomics, RADseq, São Tomé and Príncipe.**

Oceanic islands accumulate endemic species via two key mechanisms: colonization by continental or adjacent island species that subsequently diverge from source populations or in situ diversification of resident island species (MacArthur and Wilson 1963; Heaney 2000; Whittaker et al. 2008). The relative importance of dispersal versus in situ speciation in generating diversity on islands varies predictably with a number of characteristics particular to each archipelago. The contribution of in situ diversification increases with island size (Losos and Schluter 2000; Parent et al. 2008; Rabosky and Glor 2010), age (Emerson and Oromí 2005), and remoteness (Gillespie and Roderick 2002), whereas dispersal dominates in archipelagos with numerous small and young islands that are near continental sources (Paulay 1994). Organismal traits also mediate the contributions of dispersal and in situ speciation to overall rates of diversification within an archipelago. Organisms with typically low dispersal abilities but a tendency for passive long-distance dispersal display high rates of interisland colonization and diversification (e.g., land snails; Chiba 1999; Parent et al. 2008), whereas those with limited vagility or potential for divergence in secondary sexual traits provide more opportunities for divergent ecological or sexual selection to drive speciation within an island (Paulay 1985; Mendelson and Shaw 2005). Most in situ diversification on islands proceeds via allopatric speciation and is therefore typically limited to larger islands (Coyne and Price 2000; Losos and Schluter 2000; Parent and Crespi 2006; Kisel and Barraclough 2010) because of increased opportunities for geographic isolation (Endler 1977; Rosenzweig 1995) and typically greater altitudinal variation and habitat diversity (Ricklefs and Lovette 1999; Ackerman et al. 2007; Losos and Parent 2009). In some cases, in situ diversification may proceed via sympatric speciation (i.e., with gene flow), particularly on smaller islands where there are

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

fewer opportunities for geographic isolation (Savolainen et al. 2006; Papadopulos et al. 2011). Here, we use population genomic approaches to characterize the relative roles of interisland dispersal and mechanisms of in situ speciation shaping diversification in reed frogs endemic to the Gulf of Guinea islands.

The Gulf of Guinea archipelago is located on the Cameroon Volcanic Line 40 to 350 kilometers from the western coast of Central Africa and comprises one land-bridge island (Bioko) and three oceanic islands (São Tomé, Príncipe, and Annobón). The oceanic islands have remained isolated from continental Africa throughout their history, yet because they are relatively old, ranging from approximately 5 (Annobón) to 13 (São Tomé) to 30 (Príncipe) million years (Myr), they have accumulated hundreds of endemic species (Jones 1994). Due to the high taxonomic diversity of island endemics, and close proximity of the oceanic islands to coastal Africa, dispersal from the mainland to the islands has been proposed as a key mechanism shaping patterns of diversity in the archipelago (Jones 1994; Measey et al. 2007). Furthermore, the islands share a number of sister species across taxonomic groups, indicating that interisland dispersal *within* the island chain may have been an important mechanism generating diversity in the archipelago (Jesus et al. 2009; Melo et al. 2011; Miller et al. 2012; Bell et al. 2015). Although the oceanic islands are small, ranging in size from approximately 18 (Annobón) to 136 (Príncipe) to 850 (São Tomé) km^2, some lineages may have diversified rapidly within a single island to fill divergent ecological niches (Melo et al. 2011). However, mechanisms driving in situ diversification, as well as the relative contributions of dispersal versus in situ diversification in shaping total diversity, remain poorly characterized in the Gulf of Guinea archipelago.

Among the islands' endemic vertebrates, reed frogs (genus *Hyperolius*) are thought to be one of the only lineages that diversified within a single island and also dispersed between islands in the archipelago (Jones 1994; Bell et al. 2015). Therefore, this lineage provides an opportunity to jointly investigate mechanisms driving in situ diversification as well as the frequency and demographic consequences of interisland dispersal. The São Tomé Giant Reed frog (*H. thomensis*) is found only in forest habitats above 1000 m elevation on São Tomé and its sister taxon *H. molleri* is broadly distributed on two islands, occurring up to 1400 m elevation on São Tomé and up to the summit on Príncipe (900 m). *Hyperolius thomensis* and *H. molleri* are considered distinct species based on differences in body size, coloration, and breeding ecology (Drewes and Wilkinson 2004), but individuals exhibiting intermediate phenotypes are found between 1000 and 1400 m where the species' ranges overlap on São Tomé (R. C. Bell and R. C.Drewes, unpubl. data). The distribution of these intermediate forms implies some level of gene flow between the two species; either throughout their evolutionary history (i.e., divergence in sympatry), or more recently as a consequence of range expansions in one or both species (i.e., divergence in allopatry with secondary contact). In contrast, although the presence of *H. molleri* on both São Tomé and Príncipe indicates that successful dispersal between the islands occurred at least once, these populations are reciprocally monophyletic at mitochondrial loci (mtDNA; Bell et al. 2015). Therefore, we expect that dispersal events between the islands are relatively uncommon and that populations of *H. molleri* on Príncipe diverged allopatrically from those on São Tomé.

A recent multilocus phylogeographic study of the island *Hyperolius* and their mainland sister taxon (*H. cinnamomeoventris*) indicated that *Hyperolius* dispersed from West-Central Africa approximately 8.9–3.4 million years ago (M.Y.B.P.) and subsequently diversified within the archipelago (Bell et al. 2015). In most archipelagos, dispersal and colonization proceed from older to younger islands, following the "progression rule" (Wagner and Funk 1995; Roderick and Gillespie 1998; Juan et al. 2000), but patterns of mtDNA divergence among the island lineages suggested initial colonization of São Tomé (the younger of the two islands), in situ diversification on São Tomé resulting in *H. thomensis* and *H. molleri*, followed by dispersal of *H. molleri* to Príncipe. Further inferences about the frequency of interisland dispersal and the context of in situ diversification in the earlier study were limited due to small sample sizes and because the island lineages were undifferentiated at the slowly evolving nuclear coding genes (nuDNA; Bell et al. 2015). Here, we combine mitochondrial sequence and genome-wide single-nucleotide polymorphism (SNP) data with population level sampling of the island species to (1) determine whether dispersal and colonization within the archipelago is an exception to the progression rule, (2) quantify the extent of interisland dispersal and ensuing diversification, and (3) characterize the temporal and geographic extent of gene flow between sister species in a case of in situ diversification.

Methods

SAMPLING DETAILS

Hyperolius molleri is distributed on both islands and although it is currently considered a single species, we refer to the genetically distinct populations as the São Tomé and Príncipe lineages of *H. molleri* for clarity. Between 2001 and 2013, we collected 97 samples from 20 localities of *H. molleri* throughout its range on the islands of São Tomé and Príncipe, 20 samples from two localities of *H. thomensis* on São Tomé, and six samples from four localities of *H. cinnamomeoventris* from Gabon in continental Central Africa (Fig. 1). For sites between 1000 and 1400 m elevation on São Tomé where *H. molleri* and *H. thomensis* are sympatric and potentially hybridizing, we preliminarily classified individuals according to diagnostic differences in body size between the

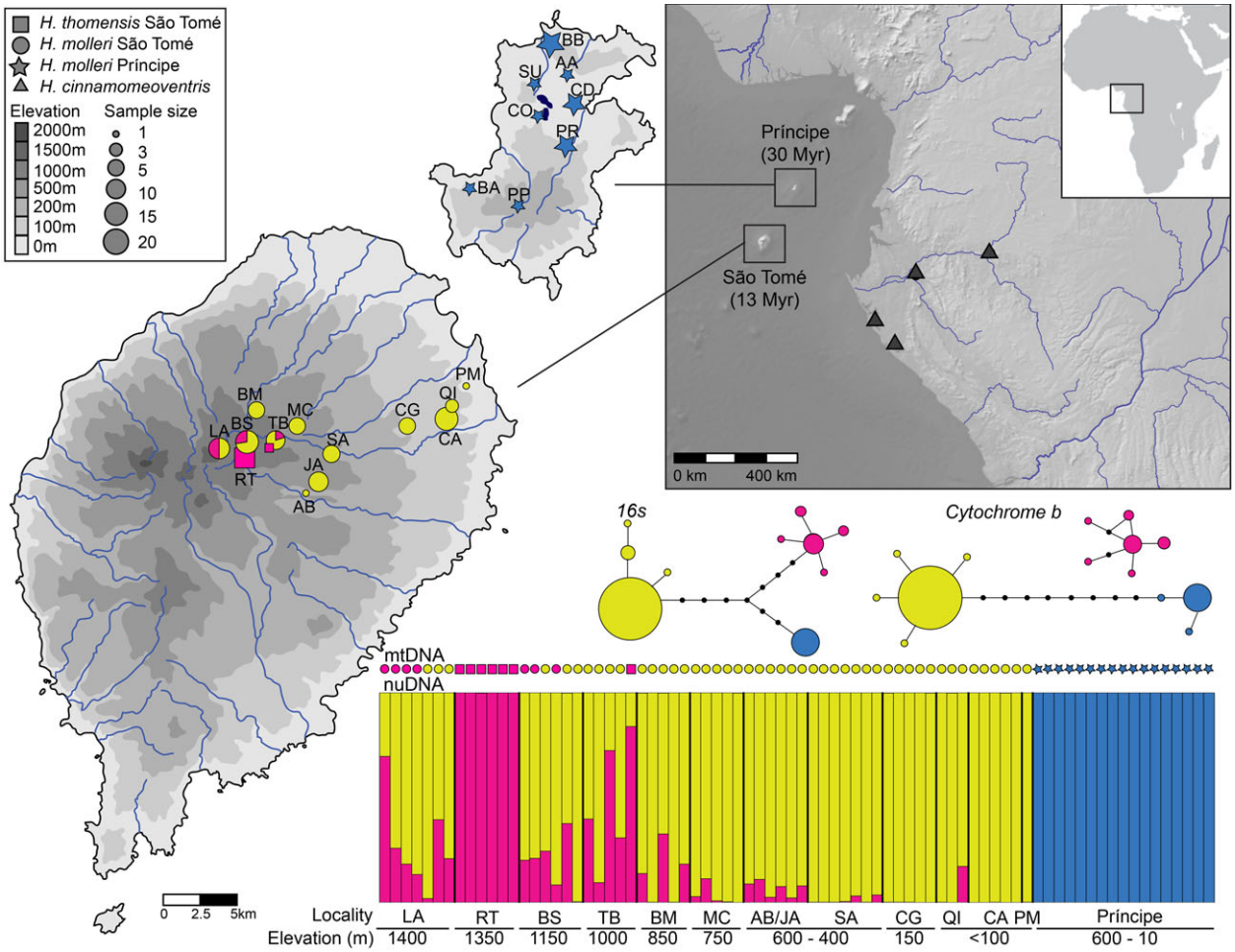

Figure 1. Sampling localities on the islands of São Tomé and Príncipe (*Hyperolius thomensis*, *H. molleri*) and in Gabon, Central Africa (*H. cinnamomeoventris*). Sampling localities are scaled according to sample size and colored according to the mitochondrial haplotype groups represented in the population (*H. thomensis*, *H. molleri* São Tomé, or *H. molleri* Príncipe). Parsimony networks of *16s* and *cytochrome b* mitochondrial haplotypes are scaled according to sample size and colors correspond to the three main haplotype groups (*H. thomensis*, *H. molleri* São Tomé, or *H. molleri* Príncipe). For *cytochrome b*, the pink haplotypes are disconnected from the rest of the network because it requires more than 10 steps to connect this group to the remaining haplotypes. Mitochondrial haplotype group and individual assignment probabilities from the STRUCTURE analysis of 3644 SNP genotypes are depicted for $K = 3$. AA, Airport Army Depot; AB, Abade; BA, Baia das Agulhas; BB, Road to Bom Bom; BM, Bem Posta; BS, Bom Sucesso; CA, Caxueira; CD, Chada Água Doutor; CG, Caxão Grande; CO, Conceição; JA, Java; LA, Lagoa Amélia; MC, Monte Café; PM, Praia Melão; PP, Pico de Príncipe; PR, Rio Papagaio; QI, Quisinda; RT, Radio Tower; SA, Santy; SU, Road to Sundy; TB, Terra Batata.

two species (*H. thomensis* male snout-vent-length [SVL] > 35 mm, female SVL > 40 mm; *H. molleri* male SVL < 30 mm female SVL < 33 mm; Schiøtz 1999; Drewes and Wilkinson 2004). Tissue samples (liver) were preserved in 95% ethanol or RNAlater for subsequent DNA extraction and genetic analyses. Preserved specimens are accessioned in the Cornell University Museum of Vertebrates (CUMV) and the California Academy of Sciences (CAS; Table S1 in Supporting Information).

MITOCHONDRIAL DIVERSITY AND DIVERGENCE

We extracted total genomic DNA using a DNeasy Blood & Tissue Kit (Qiagen Inc., Valencia, CA) and used polymerase chain reaction (PCR) to amplify and sequence two mitochondrial fragments for each individual (*cytochrome-b* and *16s*) using published primers (Table S2). PCRs were carried out in a final volume of 20 μl containing: 20 ng template DNA, 1× Buffer, 0.2 μM of each primer, 0.4 mM dNTP mix, and 0.125 units of *Taq* DNA polymerase (Roche Diagnostics, Indianapolis, IN). Amplifications were carried out with an initial denaturation for 5 min at 94°C, followed by 35 cycles (60 s denaturation at 94°C, 60 s annealing at 42–50°C [Table S2], 60 s extension at 72°C), and a final extension at 72°C for 5 min. PCR products were visualized on an agarose gel, purified using ExoSAP-IT (USB Corp., Cleveland, OH), and sequenced using a BigDye Terminator Cycle Sequencing

Kit version 3.1 (Applied Biosystems, Foster City, CA) on an ABI Automated 3730xl Genetic Analyzer (Applied Biosystems). DNA sequences were edited using SEQUENCHER 5.0.1 (Gene Codes Corp., Ann Arbor, MI) and accessioned in GenBank (*16s*: KP137113-KP137228; *cytochrome-b*: KJ865997-KJ865998, KJ866004-KJ866011, KP137229-KP137316).

Sequences were aligned using CLUSTAL X version 2.0.10 (Larkin et al. 2007) and we used TCS version 1.21 (Clement et al. 2000) to create haplotype networks for each locus. We used ARLEQUIN version 3.1 (Excoffier et al. 2005) to calculate nucleotide diversity based on the number of segregating sites (θs) and based on pairwise sequence comparisons ($\theta\pi$), uncorrected and net sequence divergence (*Dxy* and *Da* using the Tamura–Nei model; Tamura and Nei 1993), and F_{ST} for the three island lineages (*H. thomensis*, *H. molleri* from São Tomé, and *H. molleri* from Príncipe).

SNP DATASET COLLECTION

We used the double-digest RADseq laboratory protocol (ddRADseq; Peterson et al. 2012) to collect genome-wide SNP data from a representative subset of *H. molleri* (17 from Príncipe and 54 from São Tomé) and *H. thomensis* (seven from São Tomé) as well as the six samples of *H. cinnamomeoventris* from continental Central Africa (Fig. 1). For each sample, we digested 250–1000 ng of freshly extracted DNA with the restriction enzymes S*bf1* and *MspI* (New England Biolabs, Ipswich, MA), which have 8 base pairs (bp; 5′-CCTGCAGG-3′) and 4 bp (5′-CCGG-3′) recognition sites, respectively. DNA digests were purified with Agencourt AMPure beads prior to ligating barcoded Illumina adaptors to the fragments. We pooled equimolar quantities of each sample prior to size selection using a Blue Pippin Prep (fragment size range 430–530) and PCR amplified the libraries with 12 cycles using proofreading *Taq* and Illumina's indexed primers (all of which differed by at least 2 bp to reduce de-multiplexing errors). To check the quality of our libraries, we quantified the concentration of the pooled samples using Qubit Fluorometric Quantitation (Invitrogen, Carlsbad, CA) and confirmed the fragment sizes in our libraries on an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA). We sequenced two pooled libraries of 44 samples each on two lanes of Illumina HiSeq 2000 (100 bp, single end) at the Cornell University Genomics Facility. Raw sequencing reads are available on the NCBI Short Reads Archive (BioProject ID PRJNA268025).

We processed Illumina data with the STACKS pipeline version 1.13 (Catchen et al. 2011, 2013), which identifies putative loci and infers haplotypes for each individual. To create putative loci and detect SNPs at each locus, we implemented the ustacks program (Hohenlohe et al. 2010), which uses a maximum likelihood framework to group reads into loci that differ by a threshold of two mismatches with a minimum depth of coverage of five reads. Using cstacks, we generated a catalogue of consensus loci by merging unique loci across all individuals with a mismatch threshold of two differences allowed between sample tags. Finally, we resolved haplotypes for each individual for each locus in the catalogue using sstacks and removed putative loci with more than twice the standard deviation of coverage depth to filter out potentially paralogous loci. To check for consistency of results between library preparations, we replicated two samples (*H. thomensis* CAS251635 and *H. molleri* CAS233703) in each library and processed the reads through the STACKS pipeline as described above. We assessed repeatability of SNP calls for each sample by comparing haplotype assignments for loci recovered in both of the technical replicates.

Mitochondrial and nuclear loci differ in their patterns of inheritance and effective population sizes; therefore, inferences of population genetic structure, historical population demography, and gene flow based on these two classes of loci are not always concordant (Birky et al. 1989; Ballard and Whitlock 2004). RADseq methods generate SNP data for both types of markers, therefore, to differentiate between mitochondrial versus nuclear SNPs in our dataset we blasted all loci recovered in STACKS to the NCBI Vertebrate Nucleotide Database (www.blast.ncbi.nlm.nih.gov) and removed all loci that mapped to mitochondrial sequences from subsequent analyses. To generate output files for downstream analyses, we used the STACKS program populations and modified the files using custom perl scripts. STACKS may be prone to false positives in SNP calling in regions with low sequence quality (i.e., the tail ends of reads). To test whether SNPs in our dataset follow the expectation that SNPs should be called with equal frequency at every position along the STACK or locus, we plotted the frequency of SNPs called by the STACKS program with respect to the position of the SNP within the locus (Fig. S1). We found that SNP calls are evenly distributed across the entire length of the loci with the exception of the last 5 bp, which exhibit more SNP calls than expected. For each locus that met our filtering criteria for downstream analyses (see below), we selected the first SNP in the stack and excluded any loci in which the only SNP occurs in the last 5 bp where false positives are more likely.

POPULATION STRUCTURE AND PHYLOGENETIC RELATIONSHIPS OF ISLAND ENDEMICS

We used the program STRUCTURE version 2.3.4 (Pritchard et al. 2000; Falush et al. 2003) to determine the number of genetic demes and degree of admixture among demes present in our samples of *H. molleri* and *H. thomensis* from São Tomé and Príncipe. For this analysis, we included loci from our RADseq dataset that were present in all three lineages and present in at least 75% of individuals in a lineage (maximum 25% missing data per SNP within a lineage; Appendix S1 and Fig. S2). We

used 3644 SNPs from our RADseq dataset (we filtered data to include only the first SNP in each RAD locus), implemented the admixture model with correlated allele frequencies among populations and performed 10 runs at each value of K (ranging from one to four), with a burn-in of 1,000,000 steps and MCMC length of 5,000,000 steps. This dataset is archived in Dryad (http://dx.doi.org/10.5061/dryad.7s7s7). We plotted log-likelihood scores for the range of K-values (Evanno et al. 2005) to determine the most likely number of genetic clusters in the dataset and used STRUCTURE HARVESTER (Earl and vonHoldt 2011) to combine individual assignment probabilities across replicate runs. To confirm that our results were robust to missing data, we conducted preliminary clustering analyses with varying levels of missing data within lineages (10–25%) and recovered the same number of genetic clusters and that individual assignment probabilities only differed by 0.5% on average (range 0–4.3%).

To estimate a species tree from the subset of SNPs represented in the three island lineages and the mainland sister taxon ($n = 467$), we used the Bayesian program SNAPP version 1.1.1 (Bryant et al. 2012). SNAPP estimates the species tree from unlinked bi-allelic SNPs and makes the assumption of no gene flow between lineages; therefore, we filtered our dataset to include a single bi-allelic SNP from each RAD locus present in all four lineages and selected individuals of *H. molleri* (five each from São Tomé and Príncipe) and *H. thomensis* (five from São Tomé) with no evidence of admixed ancestry in the STRUCTURE analysis. We conducted preliminary analyses with varying levels of missing data (20–60%) and found no qualitative differences in topology and branch lengths between the analyses; therefore, our final dataset includes loci present in at least two individuals within each lineage (maximum 60% missing data within a lineage). We used BEAUti to generate the input file with default settings for most parameters, ran the analysis for two replicate runs of 5,000,000 MCMC steps, and assessed convergence using TRACER (Rambaut et al. 2013). We selected a gamma distribution for the θ prior and selected two sets of alpha and beta parameters that reflect either small (2, 2000; mean $\theta = 0.001$) or large (1, 10; mean $\theta = 0.1$) current and ancestral population sizes. Both sets of priors produced the same topology therefore we report the estimates using the large population sizes. The xml file for this analysis is archived in Dryad (http://dx.doi.org/10.5061/dryad.7s7s7). The effective sample size for all parameters was well above 200 and we discarded the first 10% of trees as burn-in prior to summarizing the distribution of topologies in the dataset with TREESETANALYZER. We visualized the distribution of species tree topologies and node heights using DENSITREE (Bouckaert 2010).

To compare relative diversity within and divergence between the island lineages, we used ARLEQUIN version 3.1 (Excoffier et al. 2005) to calculate F_{ST}, the proportion of polymorphic sites (P), theta based on expected homozygosity (θ_H; Zouros 1979; Chakraborty and Weiss 1991), and expected heterozygosity (H_E) versus observed heterozygosity (H_O) on the set of SNPs used in the STRUCTURE analysis ($n = 3644$). We measured mean allelic richness (N_A) with HP-RARE version 1.0 (Kalinowski 2005), which uses rarefaction and hierarchical sampling to adjust for uneven sample sizes across localities.

IDENTIFICATION AND CLASSIFICATION OF HYBRIDS

To quantify the extent of potential hybridization between *H. molleri* and *H. thomensis* on São Tomé, we used NEWHYBRIDS (Anderson and Thompson 2002) to compute the posterior probability that an individual belongs to distinct genotype frequency classes (parental, F_1, F_2, and backcrosses). We used 386 SNPs from our RADseq dataset (we filtered the set of loci from the STRUCTURE analysis to include SNPs with a minor allele frequency >0.2) and performed four replicate runs of 1,000,000 sweeps and a burn-in of 100,000 sweeps with default genotype categories. For individuals with assignment probabilities >0.99 to either the *H. molleri* or *H. thomensis* demes in the STRUCTURE analysis, we specified the corresponding genotype frequency class (parental *H. molleri* or *H. thomensis*) using the z option in the input data file. To account for the potential influence of priors on hybrid classification, we performed two runs with uniform priors and two runs with Jeffrey's priors for the mixing proportions and allele frequencies. We assessed convergence by comparing $P(z)$ values from the replicate runs. This dataset is archived in Dryad (http://dx.doi.org/10.5061/dryad.7s7s7).

Results

MITOCHONDRIAL DIVERSITY AND DIVERGENCE

We recovered three differentiated mitochondrial haplotype groups that largely correspond to *H. thomensis*, the São Tomé lineage of *H. molleri*, and the Príncipe lineage of *H. molleri* (Fig. 1). Although populations of *H. molleri* on both islands are currently considered one species, they do not share any mitochondrial haplotypes. In contrast, *H. thomensis* and *H. molleri* are considered distinct species yet seven *H. molleri* from Lagoa Amélia (LA) and Bom Sucesso (BS) on São Tomé carry *H. thomensis* mitochondrial haplotypes (Fig. 1; Table S1). The three lineages are highly differentiated from one another ($Da = 1.1$–2.7%, $F_{ST} = 0.77$–0.97; Table 1) and genetic diversity is greater within the São Tomé lineage of *H. molleri* than the Príncipe lineage (Table 2).

SNP DATASET

We generated approximately 200 million sequence reads after filtering raw reads for quality, intact restriction sites, and matches to sample barcodes (average of ~2.4 million reads per sample). The STACKS pipeline generated an average of ~28,000 unique loci per sample with an average depth of coverage of 68X per SNP.

Table 1. Estimates of pairwise F_{ST} values between *Hyperolius molleri* (São Tomé and Príncipe Islands), *H. thomensis* (São Tomé Island), and *H. cinnamomeoventris* (Gabon, Central Africa) for mtDNA (*cytochrome-b/16s)* and nuDNA (3644 RADseq SNPs).

	H. cinnamomeoventris	*H. molleri* Príncipe	*H. molleri* São Tomé
H. molleri Príncipe (mtDNA)	0.84/0.85	–	
H. molleri Príncipe (nuDNA)	–	–	
H. molleri São Tomé (mtDNA)	0.87/0.90	0.79/0.86	–
H. molleri São Tomé (nuDNA)	–	0.495	–
H. thomensis São Tomé (mtDNA)	0.77/0.84	0.97/0.96	0.82/0.85
H. thomensis São Tomé (nuDNA)	–	0.705	0.419

Hyperolius cinnamomeoventris are not included for nuDNA comparisons because a small subset of RADseq loci was shared across all four taxa. All values are significant at $P < 0.001$.

Table 2. Summary statistics for mitochondrial loci and nuclear SNPs collected from *Hyperolius molleri* (São Tomé and Príncipe Islands), *H. thomensis* (São Tomé Island), and *H. cinnamomeoventris* (Gabon, Central Africa).

	16s					Cytochrome b					Nuclear SNPs						
Lineage	*N*	bp	N_h	θs	$\theta\pi$	*N*	bp	N_h	θs	$\theta\pi$	*N*	Sites	*P*	N_A	H_O	H_E	θ_H
H. cinnamomeoventris	5	523	3	0.0358	0.0478	6	616	5	0.0647	0.0902	–	–	–	–	–	–	–
H. molleri Príncipe	21	521	1	0.0000	0.0000	22	616	3	0.0009	0.0003	17	3644	0.23	1.15	0.04	0.05	0.050
H. molleri São Tomé	71	521	7	0.0044	0.0026	57	616	8	0.0077	0.0055	54	3644	0.70	1.33	0.08	0.10	0.109
H. thomensis São Tomé	19	521	4	0.0022	0.0011	14	616	5	0.0031	0.0018	7	3644	0.28	1.23	0.07	0.08	0.084

N = number of individuals sampled; bp = sequence length in base pairs; N_h = number of haplotypes; θs = genetic diversity based on the number of segregating sites; $\theta\pi$ = genetic diversity based on pairwise sequence comparisons; P = proportion of polymorphic sites; N_A = allelic richness corrected for uneven sample size; H_O = observed heterozygosity; H_E = expected heterozygosity; θ_H = genetic diversity based on expected homozygosity.

The replicated samples (*H. thomensis* CAS251635 and *H. molleri* CAS233703) indicate that the ddRADseq protocol is reasonably repeatable with shared haplotype calls recovered for 91.2% of 4579 and 93.4% of 6050 loci shared across replicate runs, respectively. Discrepancies between replicate runs are mainly attributable to a heterozygous versus a homozygous call for an individual (6.1–7.5% of loci) and the frequency of entirely conflicting calls between replicates was very low (0.5–1.4% of loci). Six loci in the STACKS catalog that matched mitochondrial genes in the BLAST search were excluded from subsequent analyses.

POPULATION STRUCTURE AND PHYLOGENETIC RELATIONSHIPS OF ISLAND ENDEMICS

Our Structure analysis of 3644 SNPs for the island samples recovered three demes corresponding to *H. thomensis*, the São Tomé lineage of *H. molleri*, and the Príncipe lineage of *H. molleri*. Consistent with the mitochondrial haplotype networks, we find no evidence of admixture between the São Tomé and Príncipe lineages of *H. molleri* (Fig. 1). Also consistent with the mtDNA data, several *H. molleri* from sites between 1000 and 1400 m on São Tomé (Lagoa Amélia, Terra Batata, and Bom Sucesso) exhibit substantial admixture with the *H. thomensis* deme (Fig. 1). Despite extensive admixture between *H. molleri* and *H. thomensis* on São Tomé, we recovered considerable genetic differentiation among all three island lineages in our SNP dataset ($F_{ST} = 0.419$–0.705; Table 1). Estimates of heterozygosity, allelic richness, and the proportion of polymorphic sites indicate that genetic diversity is greater within the São Tomé lineage of *H. molleri* than the Príncipe lineage (Table 2).

Divergence at mitochondrial loci between mainland *H. cinnamomeoventris* and the island species ranged from 3.7 to 4.5% for *16s* and 7.7 to 8.5% for *cytochrome b*, consequently we recovered fewer shared RADseq loci across these more divergent lineages that met our criteria for the SNAPP species tree analysis ($n = 467$). Our SNAPP species tree analysis confirms that the island lineages form a monophyletic group that is well differentiated from the mainland sister taxon (Fig. 2). Consistent with the current taxonomy, we recovered a sister relationship between *H. molleri* populations from Príncipe and São Tomé, and monophyly of *H. molleri* relative to *H. thomensis* (Fig. 2). All nodes in the phylogeny are well supported (posterior probability = 0.99).

IDENTIFICATION AND CLASSIFICATION OF HYBRIDS

Using the subset of 386 SNPs with a minor allele frequency >0.2, NEWHYBRIDS identified 17 individuals of *H. molleri* as hybrids with posterior probability >0.99 (six F_2 hybrids, nine *H. molleri* backcross hybrids, and two *H. thomensis* backcross hybrids; Fig. 3). The majority of these hybrids are from three sites

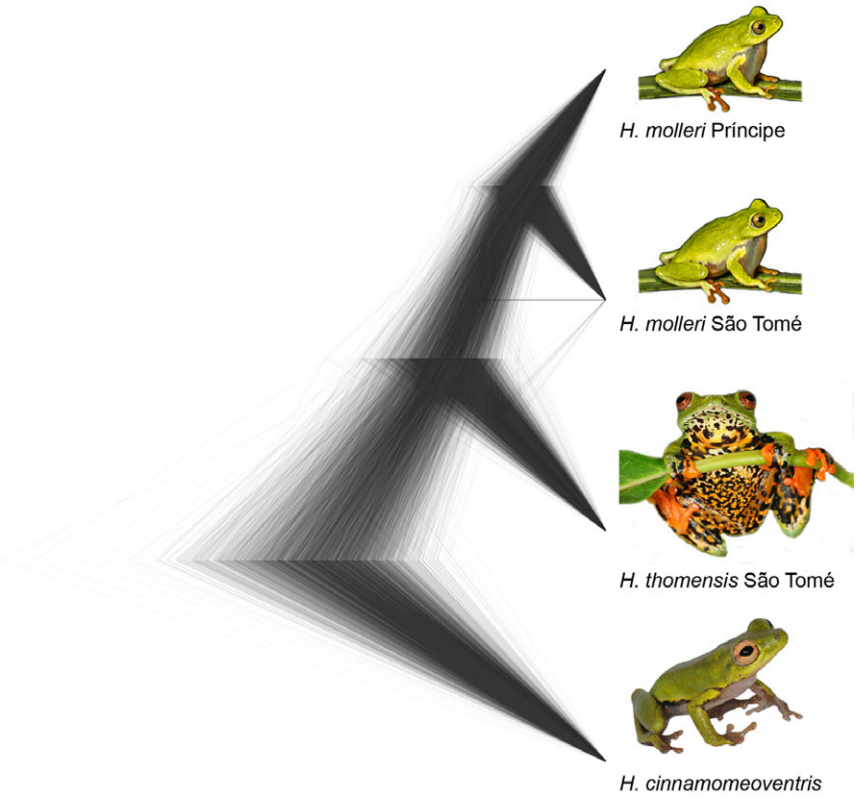

Figure 2. SNAPP **species tree inferred from 467 nuclear bi-allelic SNPs shared among *Hyperolius thomensis* (5), *H. molleri* from São Tomé (5), *H. molleri* from Príncipe (5), and the *H. cinnamomeoventris* (6) complex from Gabon. Branch lengths are a relative measure of substitutions per site. All nodes are supported by posterior probabilities greater than 0.99. Photo credits A. Stanbridge, D. Lin, and B. Stuart.**

where the ranges of *H. thomensis* and *H. molleri* are sympatric on São Tomé (Lagoa Amélia, Bom Sucesso, and Terra Batata; Fig. 3) and a few hybrids are from allopatric sites (Bem Posta, Monte Café, and Quisinda; Fig. 3). Of the seven *H. molleri* that carry *H. thomensis* mitochondrial haplotypes, one was classified by NEWHYBRIDS as *H. thomensis* backcross ($F_1 \times$ *H. thomensis*), five as *H. molleri* backcross ($F_1 \times$ *H. molleri*), and one as *H. molleri* parental (Fig. 3). The individual classified as *H. molleri* parental is a male we collected in an agricultural field (Bom Sucesso) that was assigned to the *H. molleri* deme in the STRUCTURE analysis with $Q = 0.92$. Given the mixed mitochondrial and nuclear ancestry of this male, we consider that mixed ancestry in this individual likely results from multiple generations of backcrossing with *H. molleri*.

Discussion

REED FROG DISPERSAL AND COLONIZATION IN THE GULF OF GUINEA DOES NOT FOLLOW THE PROGRESSION RULE

Although the island of São Tomé is much younger than Príncipe (13 Myr vs. 30 Myr), our species tree indicates that *Hyperolius* initially colonized São Tomé and subsequently dispersed to Príncipe. Estimates of genetic diversity (e.g., number of polymorphic sites, θ_H, and allelic richness; Table 2) for *H. molleri* on Príncipe are much lower than for São Tomé populations, which is consistent with this colonization order. Most instances of interisland dispersal in well-studied island systems such as the Canary, Hawaiian, and Galapagos archipelagos follow the progression rule (Wagner and Funk 1995), proceeding from older to younger islands (Juan et al. 2000; Cowie and Holland 2008; Parent et al. 2008). This biased direction of colonization is often attributed to the greater availability of ecological niche space on younger islands (Gillespie and Roderick 2002). Exceptions to this pattern indicate that other physical attributes of islands, including wind patterns, ocean currents, and migration routes, must also shape overall patterns of interisland dispersal (Cowie and Holland 2006). For *Hyperolius* and other organisms that rely on rafting to disperse overseas, we expect that dispersal between islands is largely determined by ocean currents, which flow from south to north (Annobón to São Tomé to Príncipe) in the Gulf of Guinea. Few phylogenetic studies are available for such taxa in the Gulf of Guinea, but mitochondrial studies of island *Afroablepharus* skinks and *Lygodactylus* geckos are consistent with a south to north dispersal pattern (Jesus et al. 2006, 2007).

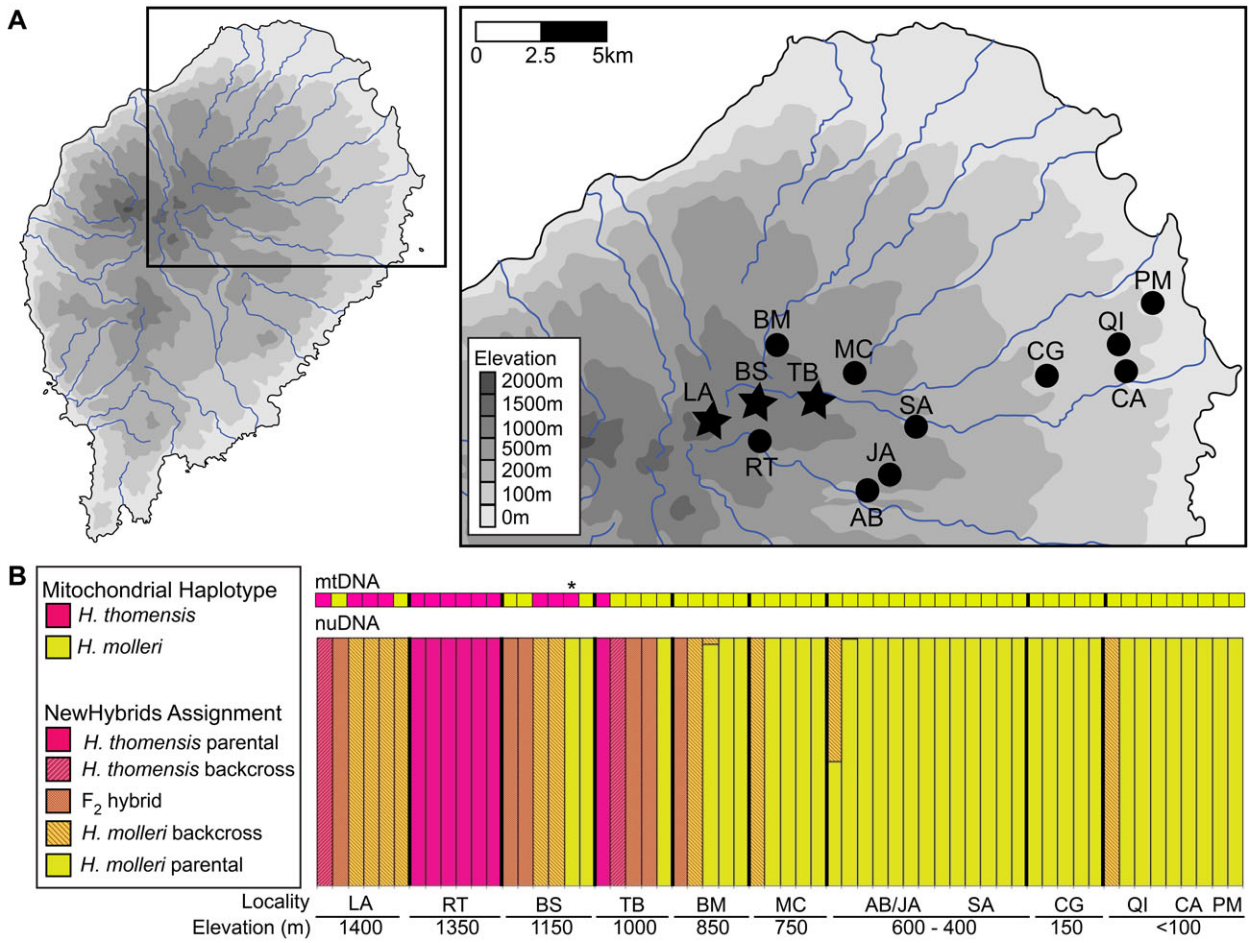

Figure 3. **(A) Sampling localities on the island of São Tomé; stars denote localities with high proportions of F_2 and backcross hybrid individuals. (B) *Hyperolius thomensis* and *H. molleri* mitochondrial haplotype group and hybrid classification from the NewHybrids analysis of 386 SNP genotypes. The asterisk denotes an individual classified by NewHybrids as *H. molleri* parental that carries an *H. thomensis* mitochondrial haplotype. AB, Abade; BM, Bem Posta; BS, Bom Sucesso; CA, Caxueira; CG, Caxão Grande; JA, Java; LA, Lagoa Amélia; MC, Monte Café; PM, Praia Melão; QI, Quisinda; RT, Radio Tower; SA, Santy; TB, Terra Batata.**

INTERISLAND DISPERSAL AND ALLOPATRIC DIVERGENCE IN *H. MOLLERI*

Hyperolius molleri populations on São Tomé and Príncipe share a most recent common ancestor and are strongly differentiated at mtDNA and nuDNA, which confirms that dispersal between the islands occurred in the past but is not currently ongoing. Although the islands are only separated by approximately 150 km, none of the six other endemic amphibians that occur on Príncipe or São Tomé have successfully dispersed between the islands, further indicating that such dispersal events are uncommon for amphibians. We previously estimated divergence between populations of *H. molleri* on the two islands at approximately 1.1 Myr to 270 kyr (Bell et al. 2015), indicating that *H. molleri* colonized Príncipe very recently in the island's 30 Myr evolutionary history. Successful dispersal and recruitment on older islands is typically limited by the availability of ecological niches (Gillespie and Roderick 2002), but Príncipe only hosts two other amphibian species (a large-bodied treefrog, *Leptopelis palmatus* and a leaf litter species, *Phrynobatrachus dispar*) that are unlikely to compete with *H. molleri*. Therefore, though in situ diversification eventually eclipses dispersal in the accumulation of biodiversity on older islands (Emerson and Oromí 2005), dispersal may continue to play an important role for groups that rarely disperse overseas and remain relatively depauperate on oceanic islands.

Populations of *H. molleri* on the two islands are currently considered a single species because they are phenotypically similar and occupy similar habitats (Drewes and Stoelting 2004). Our study clearly indicates that they represent evolutionarily distinct lineages, however, as they do not share mtDNA haplotypes and form entirely distinct genetic demes in our Structure analysis of genome-wide SNPs (Fig. 1). The consistency of these results despite fairly recent population divergence, as well as lower genetic diversity in Príncipe *H. molleri* compared to São Tomé populations (Table 2), indicates that founder effects and genetic drift

have likely augmented genetic differentiation between the two lineages. These microevolutionary processes result in large shifts in allele frequencies (Nei et al. 1975; Dlugosch and Parker 2008) and accelerate rates of lineage sorting (Kimura and Ohta 1969), which can generate phenotypic divergence over short evolutionary timescales when coupled with divergent ecological selection (Velo-Antón et al. 2011). The absence of phenotypic differentiation between the island populations of *H. molleri* may therefore indicate that the selective environments on São Tomé and Príncipe are similar. Alternatively, closer examination of the morphology and ecology of *H. molleri* on the two islands may reveal previously unrecognized phenotypic differentiation between these genetically diverged lineages.

BREEDING SITE AVAILABILITY AND IN SITU DIVERSIFICATION ON SÃO TOMÉ

Our species tree analysis confirms that divergence between *H. molleri* and *H. thomensis* occurred in situ on the island of São Tomé. Where the two species are sympatric, the STRUCTURE analysis recovers individuals with intermediate assignments to the *H. molleri* and *H. thomensis* genetic demes, which is consistent with our observations of individuals with intermediate phenotypes at these sites (Bell and Drewes, unpubl. data). Both incomplete lineage sorting and secondary introgression can produce a pattern of shared ancestry in descendent species (Maddison 1997; Hudson and Coyne 2002). When multiple ancestral alleles are retained in descendent species due to incomplete lineage sorting, this variation should be randomly distributed among descendant populations. In contrast, when multiple ancestral alleles are present in descendant species due to introgression, this variation is concentrated in populations near geographic points of contact between previously isolated species. Therefore, the geographic pattern of divergence we recover in the STRUCTURE analysis (intermediate assignment probabilities decrease with increasing distance from the zone of sympatry) is more consistent with allopatric speciation and secondary contact than with incomplete lineage sorting (McGuire et al. 2007). Despite relatively recent divergence between the two species (1.7–0.5 Myr; Bell et al. 2015), *H. thomensis* is 50% larger than *H. molleri* and breeds exclusively in water-filled tree cavities (Drewes and Stoelting 2004), implicating a role for divergent ecological selection in driving divergence between the species.

We propose that geographic segregation in the availability and type of breeding habitats on São Tomé may have driven initial allopatric divergence between *H. molleri* and *H. thomensis.* *Hyperolius molleri* breed along slow moving streams and water-filled ditches, which are typical breeding sites for *Hyperolius* species, including the mainland sister taxon *H. cinnamomeoventris* (Schiøtz 1999). The absence of slow-moving streams at higher elevations on São Tomé may underlie the evolution of tree cavity (phytotelm) breeding in *H. thomensis* although this specialized reproductive mode is typically associated with avoiding predation and competition encountered in stream or pond habitats (Lehtinen et al. 2004). Breeding site availability may also explain what has brought these previously allopatric lineages into secondary contact. The hybrid zone, which extends from approximately 1000 to 1400 m elevation on Pico de São Tomé, coincides with a transition from agricultural land to primary forest. Most of the *H. molleri* breeding sites at these elevations are artificial and associated with agriculture (e.g., cisterns); thus, the expansion of agriculture may have increased the availability of *H. molleri* breeding sites at higher elevations, enhancing the chance of secondary contact among previously spatially segregated populations. This region is also coincident with the well-studied *Drosophila santomea/D. yakuba* hybrid zone (Lachaise et al. 2000; Llopart et al. 2005; Matute et al. 2009). These two species of *Drosophila* are ecologically isolated and differences in habitat and temperature preference contribute to both premating and postmating reproductive barriers (Matute et al. 2009; Matute and Coyne 2010); therefore, the expansion of agriculture at mid to high elevations on São Tomé may have promoted secondary contact and hybridization in both *Drosophila* and *Hyperolius*. Sympatry and the potential for hybridization between *H. molleri* and *H. thomensis* may predate agricultural development on São Tomé, however, because one of our sample sites is a natural crater lake (Lagoa Amélia) at approximately 1400 m elevation on the Pico de São Tomé that hosts a large breeding population of *H. molleri*.

Hybridization between *H. molleri* and *H. thomensis* is common at the sympatric sites we sampled but the extent of introgression is geographically constrained; we do not find *H. thomensis* mitochondrial haplotypes beyond Bom Sucesso and the proportion of individuals classified as hybrids (F_2 or backcross) in the NEWHYBRIDS analysis drops precipitously where the species are allopatric. These patterns may reflect selection against hybridization (Servedio and Noor 2003). With the exception of one individual, the 17 hybrids we sampled were breeding adults that were classified as F_2 and backcross hybrids, indicating that hybrid progeny are likely viable and fertile (Blair 1964; Sasa et al. 1998; Malone and Fontenot 2008); however, these hybrids may exhibit lower fitness relative to parental phenotypes due to ecological or sexual selection against intermediate phenotypes (extrinsic postzygotic isolation; Hatfield and Schluter 1999; Lemmon and Lemmon 2010). Therefore, although premating isolation (differences in preferred breeding sites) may be the primary reproductive barrier for these two species, environment-dependent selection against intermediate phenotypes may contribute to species isolation and may ultimately accelerate the evolution of intrinsic reproductive barriers (Bolnick et al. 2006).

Although our sampling of *H. thomensis* is limited (20 individuals from two sites), we did not find any *H. thomensis* carrying

H. molleri mitochondrial haplotypes indicating that hybridization is likely asymmetrical. This apparent asymmetry in hybridization may result from sexual differences in dispersal and mating behavior (Lamb and Avise 1986; Cahill et al. 2013) such that male *H. thomensis* breed exclusively in tree cavities whereas female *H. thomensis* visit both *H. molleri* and *H. thomensis* breeding sites. The male advertisement calls of the two species are not dramatically different and we collected a female *H. thomensis* in amplexus with a male *H. molleri* in a cistern at Terra Batata (an agricultural field at 1000 m; Fig. 1); therefore, it is feasible that artificial breeding sites between 1000 and 1400 m and sexual differences in mating behavior have facilitated asymmetrical hybridization between these species. Alternatively, the absence of *H. thomensis* carrying *H. molleri* mitochondrial haplotypes and higher prevalence of *H. molleri* backcross hybrids relative to *H. thomensis* backcross hybrids may indicate strong selection against progeny from *H. thomensis* male and *H. molleri* female matings (Coyne and Orr 1998). Such asymmetric postmating isolation is relatively common in plants and animals and may result from asymmetric incompatibilities in nuclear-cytoplasmic, maternal-zygotic, or sex-chromosome/autosome interactions (Turelli and Moyle 2006).

In summary, our results indicate that initial population divergence between *H. molleri* and *H. thomensis* on São Tomé was likely allopatric due to the geographic separation of available breeding sites and that secondary contact has resulted in hybridization and extensive introgression between the species where their ranges currently overlap. The evolution of gigantism and a specialized reproductive mode in *H. thomensis* despite fairly recent divergence between *H. molleri* and *H. thomensis* (1.7–0.5 M.Y.B.P.; Bell et al. 2015) highlight a role for divergent ecological or sexual selection in driving rapid phenotypic differentiation between the species. Future studies quantifying selection on these phenotypes across the hybrid zone (e.g., Hopkins et al. 2014) may identify the selective pressures that initially drove divergence between *H. molleri* and *H. thomensis* and highlight mechanisms that underlie the evolution of gigantism on islands (Lomolino 1985) and the evolution of phytotelm breeding in frogs (Lehtinen et al. 2004).

ACKNOWLEDGMENTS

For fieldwork in Gabon, we thank the Centre National de la Recherche Scientifique et Technologique and Agence Nationale des Parcs Nationaux for research permits, the Direction de la Faune et des Aères Protégées for export permits, the Wildlife Conservation Society Gabon Program for logistical support, and B. Stuart, N. Emba-Yao, F. Moiniyoko, B. Hylayre, E. Ekomy, A. Dibata, T. Ogombet, U. Eyagui, P. Endazokou, for assistance in the field. For fieldwork on São Tomé and Príncipe, we thank the Ministry of Environment (Director General A. de Ceita Carvalho, V. Bonfim, and S. Sousa Pontes) for permission to collect and export specimens for study, STeP Up São Tomé (E. N. Seligman, R. dos Santos, and Q. Quade Cabral) and HBD of Bom Bom Island Resort and the Omali Lodge for logistical support, and J. Vindum, A. Stanbridge, J. P. Pio, B. Simison, and V. Schnoll for assistance in the field. We thank K. Marshall and M. Hydeman for assistance in the molecular laboratory, S. Bogdanowicz and M. K. Fujita for advice on laboratory protocols and analyses, H. W. Greene, R. G. Harrison, B. B. Johnson, J. I. Lovette, D. Rodriguez, N. Polato, D. Matute, K. Bi, and S. A. Taylor for thoughtful discussion, and two anonymous reviewers whose comments improved the manuscript. The STRUCTURE analyses were performed at Cornell Computational Biology Service Unit, a facility partially funded by the Microsoft Corporation. Funding for this study was provided by grants from the Explorer's Club, American Philosophical Society, Sigma Xi, Cornell chapter of Sigma Xi, Society of Systematic Biologists, Mario Einaudi Center for International Studies at Cornell University, Cornell University Graduate School, Andrew W. Mellon Student Research Grant, Paul P. Feeny Graduate Research fund in Ecology and Evolutionary Biology, Andrew and Margaret Paul Graduate Fellowship, and a National Geographical Young Explorer's Grant (to R.C.B.), and the California Academy of Sciences Gulf of Guinea Fund. The authors have no conflict of interest to declare.

DATA ARCHIVING

The doi for our data is 10.5061/dryad.7s7s7. Sequences and datasets are deposited in GenBank (*16s*: KP137113-KP137228; *cytochrome-b*: KJ865997-KJ865998, KJ866004-KJ866011, KP137229-KP137316), NCBI Sequence Read Archive (BioProject ID PRJNA268025), and Dryad (http://dx.doi.org/10.5061/dryad.7s7s7).

LITERATURE CITED

Ackerman, J. D., J. C. Trejo-Torres, and Y. Crespo-Chuy. 2007. Orchids of the West Indies: predictability of diversity and endemism. J. Biogeogr. 34:779–786.

Anderson, E. C., and E. A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217–1229.

Ballard, J. W. O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13:729–744.

Bell, R. C., R. C. Drewes, A. Channing, V. Gvozdik, J. Kielgast, S. Lötters, B. L. Stuart, and K. R. Zamudio. 2015. Overseas dispersal of Hyperolius reed frogs from Central Africa to the oceanic islands of São Tomé and Príncipe. J. Biogeogr. 42:65–75. doi:10.1111/jbi.12412.

Birky, C. W., P. Fuerst, and T. Maruyama. 1989. Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics 121:613–627.

Blair, W. F. 1964. Isolating mechanisms and interspecies interactions in anuran amphibians. Q. Rev. Biol. 39:334–344.

Bolnick, D. I., T. J. Near, and P. C. Wainwright. 2006. Body size divergence promotes post-zygotic reproductive isolation in centrarchids. Evol. Ecol. Res. 8:903–913.

Bouckaert, R. R. 2010. DensiTree: making sense of sets of phylogenetic trees. Bioinformatics 26:1372–1373.

Bryant, D., R. Bouckaert, J. Felsenstein, N. A. Rosenberg, and A. RoyChoudhury. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Mol. Biol. Evol. 29:1917–1932.

Cahill, J. A., R. E. Green, T. L. Fulton, M. Stiller, F. Jay, N. Ovsyanikov, R. Salamzade, J. St John, I. Stirling, M. Slatkin, et al. 2013. Genomic evidence for island population conversion resolves conflicting theories of polar bear evolution. PLoS Genet. 9:e1003345.

Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait. 2011. Stacks: building and genotyping Loci de novo from short-read sequences. Genes Genom. Genet. 1:171–182.

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. 22:3124–3140.

Chakraborty, R., and K. M. Weiss. 1991. Genetic variation of the mitochondrial DNA genome in American Indians is at mutation-drift equilibrium. Am. J. Phys. Anthropol. 86:497–506.

Chiba, S. 1999. Accelerated evolution of land snails Mandarina in the oceanic Bonin Islands: evidence from mitochondrial DNA sequences. Evolution 53:460–471.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9:1657–1659.

Cowie, R. H., and B. S. Holland. 2006. Dispersal is fundamental to biogeography and the evolution of biodiversity on oceanic islands. J. Biogeogr. 33:193–198.

———. 2008. Molecular biogeography and diversification of the endemic terrestrial fauna of the Hawaiian Islands. Phil. Trans R. Soc. B 363:3363–3376.

Coyne, J. A., and H. A. Orr. 1998. The evolutionary genetics of speciation. Phil. Trans R. Soc. B 353:287–305.

Coyne, J. A., and T. D. Price. 2000. Little evidence for sympatric speciation in island birds. Evolution 54:2166–2171.

Dlugosch, K. M., and I. M. Parker. 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Mol. Ecol. 17:431–449.

Drewes, R. C., and J. A. Wilkinson. 2004. The California Academy of Sciences Gulf of Guinea expedition (2001) I. The taxonomic status of the genus *Nesionixalus* Perret, 1976 (Anura: Hyperoliidae), treefrogs of São Tomé and Príncipe, with comments on the genus *Hyperolius*. Proc. Calif. Acad. Sci. 55:395–407.

Drewes, R. C., and R. Stoelting. 2004. The California Academy of Sciences Gulf of Guinea Expedition (2001) II. Additions and corrections to our knowledge of the endemic amphibians of São Tomé and Príncipe. Proc. Calif. Acad. Sci. 55:573–587.

Earl, D. A., and B. M. vonHoldt. 2011. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4:359–361.

Emerson, B. C., and P. Oromí. 2005. Diversification of the forest beetle genus Tarphius on the Canary Islands, and the evolutionary origins of island endemics. Evolution 59:586–598.

Endler, J. A. 1977. Geographic variation, speciation, and clines. Princeton Univ. Press, Princeton.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Mol. Ecol. 14:2611–2620.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. Online 1–47.

Falush, D., M. Stephens, and J. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587.

Gillespie, R. G., and G. K. Roderick. 2002. Arthropods on islands: colonization, speciation, and conservation. Annu. Rev. Entomol. 47:595–632.

Hatfield, T., and D. Schluter. 1999. Ecological speciation in sticklebacks: environment-dependent hybrid fitness. Evolution 53:866–873.

Heaney, L. R. 2000. Dynamic disequilibrium: a long-term, large-scale perspective on the equilibrium model of island biogeography. Glob. Ecol. Biogeogr. 9:59–74.

Hohenlohe, P. A., S. Bassham, P. D. Etter, N. Stiffler, E. A. Johnson, and W. A. Cresko. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet. 6:e1000862.

Hopkins, R., R. F. Guerrero, M. D. Rausher, and M. Kirkpatrick. 2014. Strong reinforcing selection in a Texas wildflower. Curr. Biol. 24:1995–1999.

Hudson, R. R., and J. A. Coyne. 2002. Mathematical consequences of the genealogical species concept. Evolution 56:1557–1565.

Jesus, J., A. Brehm, and D. J. Harris. 2006. Phylogenetic relationships of Lygodactylus geckos from the Gulf of Guinea islands: rapid rates of mitochondrial DNA sequence evolution? Herpetol. J. 16:291–295.

Jesus, J., D. J. Harris, and A. Brehm. 2007. Relationships of *Afroablepharus* Greer, 1974 skinks from the Gulf of Guinea islands based on mitochondrial and nuclear DNA: patterns of colonization and comments on taxonomy. Mol. Phylogenet. Evol. 45:904–914.

Jesus, J., Z. T. Nagy, W. R. Branch, M. Wink, A. Brehm, and D. J. Harris. 2009. Phylogenetic relationships of African green snakes (genera *Philothamnus* and *Hapsidophrys*) from São Tomé, Príncipe and Annobón islands based on mtDNA sequences, and comments on their colonization and taxonomy. Herpetol. J. 19:41–48.

Jones, P. J. 1994. Biodiversity in the Gulf of Guinea: an overview. Biodivers. Conserv. 3:772–784.

Juan, C., B. C. Emerson, P. Oromí, and G. M. Hewitt. 2000. Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. Trends Ecol. Evol. 15:104–109.

Kalinowski, S. T. 2005. hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. Mol. Ecol. Notes 5:187–189.

Kimura, M., and T. Ohta. 1969. The average number of generations until fixation of a mutant gene in a finite population. Genetics 61:763–771.

Kisel, Y., and T. G. Barraclough. 2010. Speciation has a spatial scale that depends on levels of gene flow. Am. Nat. 175:316–334.

Lachaise, D., M. Harry, M. Solignac, F. Lemeunier, V. Benassi, and M. L. Cariou. 2000. Evolutionary novelties in islands: *Drosophila santomea*, a new melanogaster sister species from Sao Tome. Proc. R. Soc. Lond. B 267:1487–1495.

Lamb, T., and J. C. Avise. 1986. Directional introgression of mitochondrial DNA in a hybrid population of tree frogs: the influence of mating behavior. Proc. Natl. Acad. Sci. USA 83:2526–2530.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, et al. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948.

Lehtinen, R. M., M. J. Lannoo, and R. J. Wassersug. 2004. Phytotelm-breeding anurans: past, present and future research. Ecology and evolution of phytotelm-breeding anurans. Misc. Publ. Mus. Zool. Univ. Mich. 193:1–9.

Lemmon, E. M., and A. R. Lemmon. 2010. Reinforcement in chorus frogs: lifetime fitness estimates including intrinsic natural selection and sexual selection against hybrids. Evolution 64:1748–1761.

Llopart, A., D. Lachaise, and J. A. Coyne. 2005. An anomalous hybrid zone in drosophila. Evolution 59:2602–2607.

Lomolino, M. V. 1985. Body size of mammals on islands: the island rule reexamined. Am. Nat. 125:310–316.

Losos, J. B., and C. E. Parent. 2009. The speciation–area relationship. Pp. 415–438 *in* J. B. Losos and R. E. Ricklefs (eds.) The theory of island biogeography revisited. Princeton Univ. Press, Princeton.

Losos, J. B., and D. Schluter. 2000. Analysis of an evolutionary species-area relationship. Nature 408:847–850.

MacArthur, R. H., and E. O. Wilson. 1963. An equilibrium theory of insular zoogeography. Evolution 17:373–387.

Maddison, W. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.

Malone, J. H., and B. E. Fontenot. 2008. Patterns of reproductive isolation in toads. PLoS One 3:e3900.

Matute, D. R., and J. A. Coyne. 2010. Intrinsic reproductive isolation between two sister species of Drosophila. Evolution 64:903–920.

Matute, D. R., C. J. Novak, and J. A. Coyne. 2009. Temperature-based extrinsic reproductive isolation in two species of Drosophila. Evolution 63:595–612.

McGuire, J. A., C. W. Linkem, M. S. Koo, D. W. Hutchison, A. K. Lappin, D. I. Orange, J. Lemos-Espinal, B. R. Riddle, and J. R. Jaeger. 2007. Mitochondrial introgression and incomplete lineage sorting through space and time: phylogenetics of crotaphytid lizards. Evolution 61:2879–2897.

Measey, G. J., M. Vences, R. C. Drewes, Y. Chiari, M. Melo, and B. Bourles. 2007. Freshwater paths across the ocean: molecular phylogeny of the frog *Ptychadena newtoni* gives insights into amphibian colonization of oceanic islands. J. Biogeogr. 34:7–20.

Melo, M., B. H. Warren, and P. J. Jones. 2011. Rapid parallel evolution of aberrant traits in the diversification of the Gulf of Guinea white-eyes (Aves, Zosteropidae). Mol. Ecol. 20:4953–4967.

Mendelson, T. C., and K. L. Shaw. 2005. Sexual behaviour: rapid speciation in an arthropod. Nature 433:375–376.

Miller, E. C., A. B. Sellas, and R. C. Drewes. 2012. A new species of *Hemidactylus* (Squamata: Gekkonidae) from Príncipe Island, Gulf of Guinea, West Africa with comments on the African-Atlantic clade of *Hemidactylus* geckos. Afr. J. Herpetol. 61:40–57.

Nei, M., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. Evolution 29:1–10.

Papadopulos, A. S., W. J. Baker, D. Crayn, R. K. Butlin, R. G. Kynast, I. Hutton, and V. Savolainen. 2011. Speciation with gene flow on Lord Howe Island. Proc. Natl. Acad. Sci. USA 108:13188–13193.

Parent, C. E., A. Caccone, and K. Petren. 2008. Colonization and diversification of Galapagos terrestrial fauna: a phylogenetic and biogeographical synthesis. Phil. Trans R. Soc. B 363:3347–3361.

Parent, C. E., and B. J. Crespi. 2006. Sequential colonization and diversification of Galapagos endemic land snail genus Bulimulus (Gastropoda, Stylommatophora). Evolution 60:2311–2328.

Paulay, G. 1985. Adaptive radiation on an isolated oceanic island: the Cryptorhynchinae (Curculionidae) of Rapa revisited. Biol. J. Linn. Soc. 26:95–187.

———. 1994. Biodiversity on oceanic islands: its origin and extinction1. Integr. Comp. Biol. 34:134–144.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7:e37135.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Rabosky, D. L., and R. E. Glor. 2010. Equilibrium speciation dynamics in a model adaptive radiation of island lizards. Proc. Natl. Acad. Sci. USA 107:22178–22183.

Rambaut, A., M. A. Suchard, D. Xie, and A. J. Drummond. 2013. Tracer v. 1.5. Available at http://tree.bio.ed.ac.uk/software/tracer/

Ricklefs, R. E., and I. J. Lovette. 1999. The roles of island area per se and habitat diversity in the species–area relationships of four Lesser Antillean faunal groups. J. Anim. Ecol. 68:1142–1160.

Roderick, G. K., and R. G. Gillespie. 1998. Speciation and phylogeography of Hawaiian terrestrial arthropods. Mol. Ecol. 7:519–531.

Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge Univ. Press, New York, NY.

Sasa, M. M., P. T. Chippindale, and N. A. Johnson. 1998. Patterns of postzygotic isolation in frogs. Evolution 1811–1820.

Savolainen, V., M.-C. Anstett, C. Lexer, I. Hutton, J. J. Clarkson, M. V. Norup, M. P. Powell, D. Springate, N. Salamin, and W. J. Baker. 2006. Sympatric speciation in palms on an oceanic island. Nature 441:210–213.

Schiøtz, A. 1999. Treefrogs of Africa. Edition Chimaira, Frankfurt am Main.

Servedio, M. R., and M. A. Noor. 2003. The role of reinforcement in speciation: theory and data. Annu. Rev. Ecol. Evol. Syst. 34:339–364.

Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10:512–526.

Turelli, M., and L. C. Moyle. 2006. Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. Genetics 176:1059–1088.

Velo-Antón, G., K. R. Zamudio, and A. Cordero-Rivera. 2011. Genetic drift and rapid evolution of viviparity in insular fire salamanders (*Salamandra salamandra*). Heredity 108:410–418.

Wagner, W. L., and V. A. Funk (eds). 1995. Hawaiian biogeography. Evolution on a hot spot archipelago. Smithsonian Institution, Washington, DC.

Whittaker, R. J., K. A. Triantis, and R. J. Ladle. 2008. A general dynamic theory of oceanic island biogeography. J. Biogeogr. 35:977–994.

Zouros, E. 1979. Mutation rates, population sizes and amounts of electrophoretic variation of enzyme loci in natural populations. Genetics 92:623–646.

Associate Editor: J. Good
Handling Editor: T. Lenormand

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. Sampling localities and voucher information.
Table S2. Primer sequences and amplification conditions for mitochondrial sequences collected from *Hyperolius molleri* (São Tomé and Príncipe Islands)
Figure S1. Frequency of SNPs in the catalog of loci called by the Stacks pipeline with respect to the position of the SNP within the Stack.
Figure S2. Distribution of missing data in the Structure analysis shown as the proportion of individuals in each percentage of missing data category.