Received: 27 August 2017 | Revised: 25 February 2018 | Accepted: 12 March 2018

DOI: 10.1111/mec.14573

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Diversification of African tree frogs (genus *Leptopelis*) in the highlands of Ethiopia

Jacobo Reyes-Velasco[1] | Joseph D. Manthey[1] | Xenia Freilich[2] | Stéphane Boissinot[1]

[1]New York University Abu Dhabi, Saadiyat Island, Abu Dhabi, UAE

[2]Department of Biology, Queens College, City University of New York, Flushing, NY, USA

Correspondence
Stéphane Boissinot, New York University Abu Dhabi, Saadiyat Island, Abu Dhabi, UAE.
Email: stephane.boissinot@nyu.edu

Abstract

The frog genus *Leptopelis* is composed of ~50 species that occur across sub-Saharan Africa. The majority of these frogs are typically arboreal; however, a few species have evolved a fossorial lifestyle. Most species inhabit lowland forests, but a few species have adapted to high elevations. Five species of *Leptopelis* occupy the Ethiopian highlands and provide a good opportunity to study the evolutionary transition from an arboreal to a fossorial lifestyle, as well as the diversification in this biodiversity hot spot. We sequenced 14 nuclear and three mitochondrial genes, and generated thousands of SNPs from ddRAD sequencing to study the evolutionary relationships of Ethiopian *Leptopelis*. The five species of highland *Leptopelis* form a monophyletic group, which diversified during the late Miocene and Pliocene. We found strong population structure in the fossorial species *L. gramineus*, with levels of genetic differentiation between populations similar to those found between arboreal species. This could indicate that *L. gramineus* is a complex of cryptic species. We propose that after the original colonization of the Ethiopian highlands by the ancestor of the *L. gramineus* group, episodes of vicariance fragmented the ancestral populations of this group. We also report the re-evolution of arboreality in *L. susanae*, which evolved from a fossorial ancestor, a rare ecological switch in frogs that had previously been reported only once.

KEYWORDS
amphibians, fossoriality, phylogeography, RAD sequencing

1 | INTRODUCTION

The Ethiopian highlands are home to a rich and unique biota. Approximately 10% of all vascular plants in the region are endemic, as well as approximately 120 species of vertebrates (Williams, Vivero Pol, Spawls, Shimelis, & Kelbessa, 2004). However, the region is severely threatened by habitat destruction, which makes it a biodiversity hot spot (Mittermeier, 2004; Myers, Mittermeier, Mittermeier, Da Fonseca, & Kent, 2000). The amphibian fauna of the region is of special interest, as it contains five endemic genera (four anurans and one caecilian), while about half of all amphibian species in the country are endemic (Largen & Spawls, 2010; Mengistu, Nagel, Getahun, Saber, & Loader, 2013). Many of the endemic taxa of these highlands have affinities to lowland species from other regions of Africa (Freilich, Tollis, & Boissinot, 2014; Freilich et al., 2016; Loader et al., 2014; Siu-Ting et al., 2014), while several species that are lowland generalists elsewhere have adapted to high elevations in the Ethiopian highlands (Largen & Spawls, 2010).

The complex geological history of the Ethiopian Highlands might be partially responsible for its high biodiversity and endemism in the region. These highlands do not constitute a single unit, but instead are formed by various massifs of different origins and ages. The region began to rise approximately 75 million years ago, with intense volcanic activity between 45 and 5 million years ago (Williams et al., 2004). Multiple canyons and valleys further shaped the rugged geography of the region. The Great Rift Valley (GRV) crosses the highlands from NE

to SW, dividing them into two main blocks, which are the *Abyssinian massif* to the north, and the *Harar massif* to the south (see Figure 1). The GRV is known to be an important barrier to the dispersal of a multitude of taxa (Evanno, Regnaut, & Goudet, 2005; Freilich et al., 2016; Gottelli, Marino, Sillero-Zubiri, & Funk, 2004; Kebede, Ehrich, Taberlet, Nemomissa, & Brochmann, 2007; Manthey, Reyes-Velasco, Freilich, & Boissinot, 2017; Reyes-Velasco, Manthey, Bourgeois, Freilich, & Boissinot, 2018; Smith, Noonan, & Colston, 2017). Other hydrological features also played an important role in the complex geology of the region and in the diversification of Ethiopian taxa, in particular the Omo River Valley and the Blue Nile Valley (Evans, Bliss, Mendel, & Tinsley, 2011; Freilich et al., 2014; Gottelli et al., 2004; Manthey et al., 2017; Reyes-Velasco et al., 2018).

The frog genus *Leptopelis* is composed of more than 50 species that are distributed across sub-Saharan Africa. Most species are found at low to mid-elevations in moist environments along the African tropical belt, from Senegal and Gambia in the west to Kenya and Tanzania in the east (Schiøtz, 1999). Some species have colonized drier habitats like savannahs and thorn scrub, while a few occur at higher elevations in the East African Highlands and in South Africa (Channing, 2001; Harper et al., 2010; Largen & Spawls, 2010). Most species of *Leptopelis* are arboreal or semi-arboreal (Portik & Blackburn, 2016; Schiøtz, 1999); however, four species show a fossorial lifestyle, which we define as spending most of their adult life underground. These species include *L. bufonides* from the West African savannas; *L. bocagii* from eastern and southern Africa; *L. parbocagii*, which occurs from the DRC to Mozambique; and *L. gramineus*, which is endemic to Ethiopia (Largen, 1977; Largen & Spawls, 2010; Schiøtz, 1999). These species share several morphological similarities that appear to be adaptations to a fossorial lifestyle, including a heavy body, short limbs, reduced webbing between fingers and toes, an enlarged metatarsal tubercle and reduced discs on fingers and toes (Schiøtz, 1999). Because of their morphological similarities, these fossorial *Leptopelis* were considered to be each other's closest relatives (Schiøtz, 1999). *Leptopelis gramineus* is the only fossorial species that occurs at high elevations (>2,500 m) and is the main focus of this study.

Leptopelis gramineus is found on both sides of the Great Rift Valley (GRV), at elevations ranging from ~1,700 to more than 3,500 m (Largen & Spawls, 2010). While *L. gramineus* is restricted to the Ethiopian highlands, it inhabits multiple ecosystems within this region, including mountain grasslands, tropical mountain forests and Afro-alpine moorland above 3,500 m (Largen & Spawls, 2010). With the exception of another Ethiopian endemic, the Ethiopian Mountain toad (*Altiphrynoides malcolmi*), no other African amphibian is able to survive at higher elevations (Largen & Spawls, 2010). A recent study of *L. gramineus* revealed high levels of genetic differentiation in the mitochondrial gene COX1 between populations across its range; however, little to no variation between populations was found when using a small number of nuclear markers (Freilich et al., 2016).

Four other species of *Leptopelis* are endemic to the Ethiopian highlands (Figure 1). The *Leptopelis* from the Ethiopian highlands are a good model to study adaptation into novel environments, as two notable ecological adaptations have occurred in this group: (i) adapting to high elevations (>2,000 m) and all the associated physiological and behavioural challenges that this requires; and (ii) the occurrence

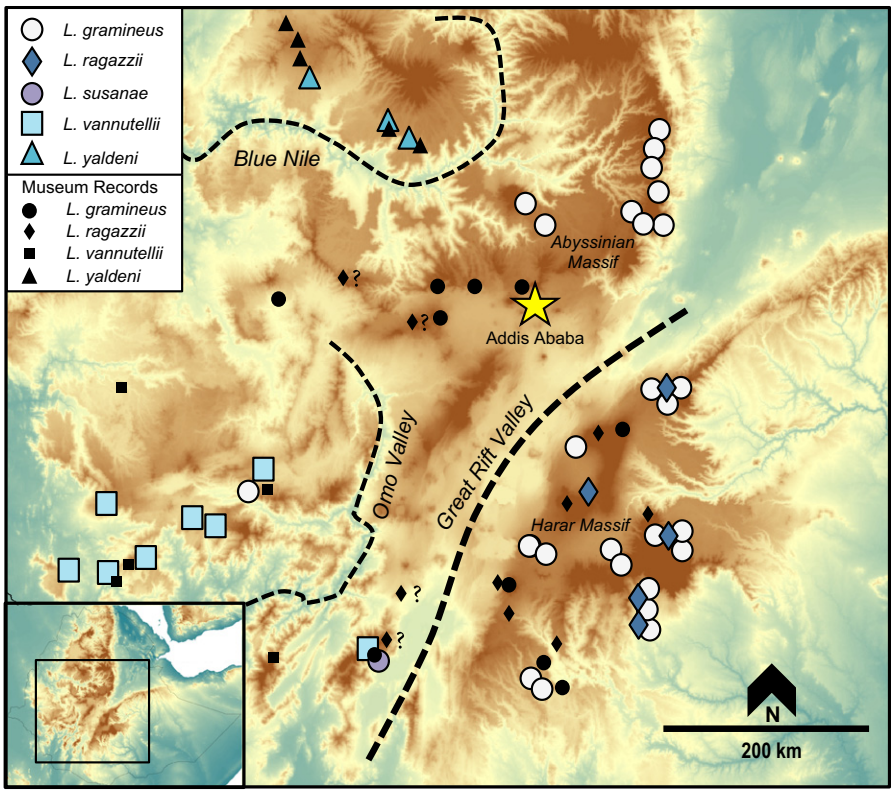


FIGURE 1 Map of Ethiopia showing the localities of *Leptopelis* samples used in this study, as well as previous records in museum collections. Question marks represent questionable records. *Leptopelis gramineus* (grey circles); *L. ragazzii* (blue diamonds); *L. susanae* (purple circle); *L. vannutellii* (blue squares); *L. yaldeni* (blue triangle). Star represents Addis Ababa [Colour figure can be viewed at wileyonlinelibrary.com]

of a terrestrial/fossorial lifestyle at high altitudes in *L. gramineus*. The other four species of *Leptopelis* endemic to Ethiopia are all arboreal and share morphological characters typical of this lifestyle, including a slender body, long legs and expanded disc pads. These endemic species include the following: (i) *L. susanae*, restricted to the Gughe Mountains of southern Ethiopia, at elevations between 2,600 and 3,000 m; (ii) *L. vannutellii*, which is found in the mid-elevation tropical humid forests of southwest Ethiopia; (iii) *L. yaldeni* occurs in the highlands northwest of the Blue Nile River; and (iv) *L. ragazzii*, which is widespread in the eastern highlands between 1,800 and 3,000 m, although it has also been reported in the west (Largen & Spawls, 2010; Perret, 1980).

Previous studies on Ethiopian *Leptopelis* focused on aspects of their natural history, distribution, conservation and taxonomy (Gower et al., 2012, 2013; Largen, 1977; Weinsheimer, Mengistu, & Rödder, 2010). However, the phylogenetic affinities of the species are not well understood, and the only assessment of their phylogenetic relationships is based on a single mitochondrial gene (Portik & Blackburn, 2016). Here we used sequence data of 14 nuclear and 3 mitochondrial loci, as well as double-digest restriction site-associated DNA sequencing (ddRAD-seq), to time the diversification of Ethiopian *Leptopelis*, decipher their evolutionary affinities and assess levels of genetic divergence between species and populations.

2 | METHODS

2.1 | Sample collection

We conducted multiple field collecting trips across Ethiopia in the summers of 2011, 2013, 2015 and 2016 to obtain samples of all species of highland *Leptopelis* from across their range (Figure 1; Table S1). Specimens were usually encountered at night, although several individuals were found during the day, usually under logs or rocks. For each specimen, we recorded latitude, longitude and elevation with a hand-held GPS device. Each specimen was photographed and then euthanized with ventral application of benzocaine. We obtained muscle or liver tissue and preserved it in 95% ethanol, cell lysis buffer or RNAlater (Invitrogen). We fixed each specimen in 10% formalin and then transferred it to 70% ethanol for permanent storage. Tadpoles were euthanized by submerging them in 10% ethanol and later preserved in 10% formalin. All specimens are deposited at the Collection of Herpetology, University of Addis Ababa, Ethiopia, and tissue samples are deposited at the Vertebrate Tissue Collection, New York University Abu Dhabi (NYUAD). We included all relevant sequences for *Leptopelis* from GenBank to test phylogenetic hypotheses and additionally sequenced amplicons and ddRAD-seq of the Congolese species *L. christyi* and *L. fiziensis* as outgroups. Samples of these species were kindly provided by Eli Greenbaum at the University of Texas at El Paso.

2.2 | DNA extraction and PCR amplification

We extracted DNA from tissue samples with one of the following techniques: DNeasy blood and tissue kit (Qiagen, Valencia, CA), with the use of Serapure beads (Rohland & Reich, 2012), by standard Phenol-Chloroform extraction, or with the use of standard potassium acetate extraction. For each sample, we measured DNA concentration using a Qubit fluorometer (Life Technologies).

We barcoded all samples of adult *Leptopelis* collected in Ethiopia as well as many juveniles and tadpoles by sequencing a fraction of the 16s rRNA mitochondrial gene ($n = 78$). We used the primers LX12SN1a and LX16S1Ra developed by Zhang et al. (2013), or with the modified primers 16Sar and 16Sbr of Bossuyt and Milinkovitch (2000). We performed polymerase chain reactions (PCRs) in total volumes of 25 μl with the use of regular Taq (Invitrogen), following the manufacturer's suggested protocol. For the first set of primers, we used the following PCR conditions: initial denaturation at 96°C for 2 min, 35 cycles of denaturing at 95°C for 15 s, annealing at 58°C for 1 min, extension at 72°C for 2 min and a final extension at 72°C for 10 min. For the second set of primers, we used the following conditions: initial denaturation at 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, 48°C for 30 s and 72°C for 1 min, and a final extension at 72°C for 10 min. Unpurified PCR products were shipped for sequencing to BGI Tech Solutions (Hong Kong).

For a subset of the samples ($n = 24$), we sequenced 14 single-copy, protein-coding nuclear genes and two mitochondrial genes (Table S1). We chose samples to represent all species of highland *Leptopelis* in Ethiopia and from as many populations as possible. We used parallel-tagged amplicon sequencing of PCR products (Feng, Liu, Chen, Liang, & Zhang, 2016) to sequence multiple long PCR products in the Illumina platform, while reducing the costs associated with traditional Sanger sequencing. We sequenced a subset of the genes proposed by Shen, Liang, Feng, Chen, and Zhang (2013), which are protein-coding, single-copy nuclear genes, and have been shown to be useful in vertebrate systematics. This method makes use of a nested PCR strategy to sequence conserved genes across vertebrates, with the use of a set of two primer pairs per loci (Appendix S1, Table S2). We selected genes with slow, medium, and fast relative evolutionary rates as determined by Shen et al. (2013), using the rate multipliers (m) in MrBayes 3.2 (Ronquist et al., 2012). Additionally, we included the nuclear gene CXCR4 and the mitochondrial genes COX 1 and CytB, as these genes are commonly used in anuran systematics (Zhang et al., 2013).

Library preparation and sequencing for all 24 samples were performed at the Genomics Core Facility at New York University Abu Dhabi, United Arab Emirates. Sequencing was performed on an Illumina MiSeq flow cell.

2.3 | Parallel-tagged Amplicon processing and phylogenetic analysis

We assembled the reads of each gene with the program Geneious v9.1.6 (Biomatters Ltd., Auckland, NZ). Of the 18 genes obtained by parallel-tagged amplicon sequencing, 16 had enough reads to be assembled into complete gene sequences. A single individual of *Leptopelis* was discarded because of low coverage across loci. After assembly of reads for each gene and individual, we created

alignments with the program MUSCLE (Edgar, 2004) in Geneious v9.1.6, using the default settings. We phased each nuclear gene for all individuals using the default settings in the program *Phase* in DnaSP (Rozas, Sánchez-DelBarrio, Messeguer, & Rozas, 2003) and used the resulting haplotypes for all subsequent analyses. For each gene, we randomly choose one of the phased haplotypes before concatenation. We included additional species of *Leptopelis* obtained from GenBank (Table S4). We selected best-fit models of nucleotide substitution for each gene using the Bayesian information criterion (BIC; Table S5), implemented in PartitionFinder v1.1.1 (Lanfear, Calcott, Ho, & Guindon, 2012). We concatenated all genes with the program Sequence Matrix (Vaidya, Lohman, & Meier, 2011). For the mitochondrial gene 16s, which we sequenced separately by standard Sanger sequencing, we manually trimmed the alignment at the 5′ and 3′ ends of all sequences. After alignment in MUSCLE, we used the less stringent selection option in the Gblocks server version 0.91b (Castresana, 2000) to eliminate poorly-aligned positions and divergent regions, resulting in an alignment of 467 bp.

We performed several phylogenetic analyses using Bayesian inference (BI) in MrBayes v3.2.2 (Ronquist et al., 2012) and implemented on the CIPRES Science gateway server (Miller, Pfeiffer, & Schwartz, 2010). Our phylogenetic analyses consisted of four separate data sets: the 16s rRNA only (467 bp), two separate data sets including only nuclear genes (14 genes, 11,082 bp) or mitochondrial genes (three genes, including 16s; 3,279 bp), and all nuclear and mitochondrial genes combined (17 genes, 14,361 bp). Each BI analysis consisted of four runs, each for 10^7 generations with four chains (one cold and three heated), sampled every 1,000 generations. We confirmed that independent runs had converged based on overlap in likelihood and parameter estimates among runs, as well as effective sample size (ESS) and Potential Scale Reduction Factor value estimates (PSRF), which we evaluated in Tracer v1.6 (Drummond & Rambaut, 2007). PSRF values indicated that individual runs had converged by 10^5 generations, and we therefore discarded the first 25% of the runs as burn-in.

2.4 | ddRAD-seq library preparation and sequencing

We performed a double-digest restriction site-associated DNA sequencing (ddRAD-seq) with the enzymes SbfI and MspI (Peterson, Weber, Kay, Fisher, & Hoekstra, 2012) to obtain SNPs and to assess genomewide patterns of diversity and differentiation in Ethiopian *Leptopelis*. A majority of the samples (n = 46) belong to the *L. gramineus* complex, with the remainder (n = 13) representing the three other highland *Leptopelis* species. We digested with both restriction enzymes simultaneously with the recommended buffer (10× CutSmart Buffer, Thermo Scientific) for 7 hr at 37°C. We then purified the fragments with the use of Serapure beads before ligation of barcoded Illumina adaptors (Table S6). We then pooled each column of uniquely barcoded samples and size selected for fragments between 400 and 550 bp on a Pippin Prep (Sage Science, Beverly, MA, USA), followed by amplification of size-selected pools using PCR with Illumina's indexed primers. We used a combination of eight barcodes and 12 Illumina indexes (i.e., dual-indexed) to multiplex samples into a single sequencing lane. Previous to sequencing, we determined fragment size distribution and concentration on a Bioanalyzer 7500 high sensitivity DNA chip (Agilent, Santa Clara, CA, USA), followed by quantitative PCR to check library quantity. All libraries were pooled and sequenced on an Illumina MiSeq or Illumina HiSeq2500 (paired-end reads of different lengths) at the Genome Core Facility of the New York University Abu Dhabi, United Arab Emirates.

2.5 | ddRAD-seq data analyses

We used the FASTX Toolkit (Gordon & Hannon, 2010) to trim restriction sites and to trim all sequences to ensure they were of equivalent length due to sequencing on different Illumina platforms, and to convert the sequences to FASTA format. After trimming, the total length of the combined paired-end reads was 188 bp. We then assembled loci de novo and created SNP data sets from the ddRAD-seq reads using the program ipyrad 0.6.17 (Eaton & Overcast, 2017), an improved version of the program pyRAD (Eaton, 2014). We preferred ipyrad to other RADseq pipelines, because ipyrad (and its predecesor, pyRAD) is specifically designed to be used for phylogenetic studies with highly divergent taxa (Leache et al., 2015), and preliminary analyses of our data showed high molecular divergence between the different species and populations of Ethiopian *Leptopelis*.

For the ipyrad parameters, we required sequences to have an average *phred* score offset of 33, and a maximum of five low-quality bases per read to be included in downstream analyses. We used the default clustering threshold of 85%. We then created two different data sets: The first one required each locus to be present in at least 50% of individuals, as suggested by Streicher, Schulte, and Wiens (2015), while the more stringent data set required loci to be present in at least 70% of individuals. All other parameters were left at default values. Because we were largely interested in diversification within the *L. gramineus* lineage, we created additional SNP data sets specific to *L. gramineus* and another for the arboreal lineage (*L. ragazzii*, *L. vannutelli* and *L. yaldeni*), with the same parameters as above. Preliminary analyses showed *L. susanae* nested within *L. gramineus*, so we included both species in the *L. gramineus* data set.

After quality filtering, we retained a total of ~82 million sequencing reads, with highly variable coverage across individuals (mean = 1.06 million, *SD* = 0.8 million, Table S7). This sequencing produced a mean of ~16,000 RAD-tags (*SD* = 7,000) per individual. Of the six main SNP data sets we created, we obtained between 149 and 5,825 polymorphic loci and between 14,315 and 54,995 SNPs (Table S7).

2.6 | Species-tree analysis

We implemented a multispecies coalescent model to estimate the "species-tree" based on our nuclear genes data set. We included all five species of Ethiopian highland *Leptopelis* and subdivided

L. gramineus into seven different populations based on our findings of the concatenated nuclear genes and the ML phylogeny of ddRAD-seq data. We used the species *L. christyi* and *L. fiziensis* as outgroups. We used the program *BEAST (Heled & Drummond, 2010) to estimate a species-tree from the 14 nuclear loci using a HKY + Gamma model of nucleotide substitution for each data partition, and a relaxed lognormal molecular clock. We set the tree prior to a Yule model, and all other values were set to default. We ran the analyses in duplicate, each for 200 million generations, sampling every 20,000 generations, for a total of 10,000 trees. We examined the sampling and convergence results of the *BEAST runs with Tracer v1.6 (Drummond & Rambaut, 2007). We then used TreeAnnotator v1.8.2 (Rambaut, Suchard, Xie, & Drummond, 2014) to discard the first 10% of the sampled trees as burn-in and to map support values of the remaining samples to the species-tree.

2.7 | Divergence time estimates

We estimated divergence times between the different species and populations of Ethiopian highland *Leptopelis* using BEAST2 (Bouckaert et al., 2014). First, we generated a guide tree using a subset of our taxa for which we had complete coverage of the nuclear genes KIAA2003, FICD and RAG-1. We also included 42 additional Afrobatrachian taxa from the Portik and Blackburn (2016) data set (see Table S4) to estimate divergence times. We used two secondary calibration points for our divergence estimates: (i) the most recent common ancestor of the families Arthroleptidae + Hyperoliidae + Hemisotidae + Brevicipitidae to 92.8 Ma ± 5.0 *SD*; and (ii) the most recent common ancestor of the families Hemisotidae + Brevicipitidae to 65.9 Ma ± 6.5 *SD*. These calibration points are based on those provided by Portik and Blackburn (2016), and Roelants et al. (2007), who used fossil and palaeontological evidence as calibration points. We used a normal distribution that overlapped with the confidence intervals reported in previous studies, with no minimum or maximum bound. We set the tree prior to the Yule model of speciation, with an uncorrelated relaxed lognormal clock, and unlinked clock and substitution models. We used the same substitution models as in the *BEAST analysis. Before running our analysis in BEAST2, we manually inserted the guide tree into the xml file generated in BEAUTi. We ran the analysis for 100 million generations, sampling every 10,000 generations, for a total of 10,000 trees. We examined convergence of the runs with Tracer v1.6 (Drummond & Rambaut, 2007) and discarded 25% of the runs as burn-in. Finally, we created a maximum clade credibility tree with the remaining trees in TreeAnnotator v1.8.2 (Drummond & Rambaut, 2007; see Results). As an additional estimate of divergence time, we used the substitution rate of 0.00957/lineage/million years (Crawford, 2003; Freilich et al., 2016) for the mitochondrial gene Cytb to date our phylogeny. We used a strict clock model and a constant population size coalescent tree prior. We modified the XML file to include the resulting topology of our nuclear data set as guide tree for the analysis. We ran the MCMC in BEAST2 for 50 million generations, sampling every 5,000. We again examined convergence of the runs with Tracer v.1.6 and discarded 25% of the runs as burn-in.

2.8 | Phylogenetic analysis of genome-wide SNP data

We used a maximum-likelihood framework (ML) to investigate the evolutionary relationships between all individuals of Ethiopian *Leptopelis*. Initially, we used PAUP* v.4.0.a151 (Swofford, 2003) with the Bayesian information criterion to estimate the best model for characterizing sequence evolution (GTR + I + G) in the ddRAD-seq concatenated data set. We used this model for subsequent phylogenetic analyses. We used RAxML v8 (Stamatakis, 2014) to estimate a phylogeny using ML, with 1,000 quick bootstrap replicates to assess support. We ran RAxML in two separate data sets, one that included only loci recovered in at least 50% of all samples (3,757 loci) and another that included only loci recovered in at least 70% of samples (1,681 loci). The total lengths of the two data sets were ~440,000 bp (50% missing data) and ~63,000 bp (30% missing data). Additionally, to visualize conflicting phylogenetic signal or ambiguities in the genomewide SNP data in populations of *L. gramineus*, we constructed phylogenetic networks, using the NeighborNet algorithm implemented in SplitsTree 4 (Huson & Bryant, 2006).

2.9 | Population structure and nucleotide diversity

We used the STRUCTURE software (Pritchard, Stephens, & Donnelly, 2000) to examine genetic structure among individuals without a priori input. Here, we utilized one SNP from each RAD locus for use in STRUCTURE, using the 50% and 70% coverage SNP matrices of the *L. gramineus* complex, and a separate analysis for the arboreal lineage. To ensure that subsetting the data did not influence results, we created two random samples of SNPs from each data set. We ran STRUCTURE initially to infer lambda with the number of populations (k) limited to one. Subsequent STRUCTURE runs used a constant value of lambda, the admixture model with correlated allele frequencies, a variety of likely k values (k = 1–12, five runs each) and a burn-in period of 50,000 MCMC generations, followed by another 50,000 iterations. To assess the number of populations for which to report results, we used the ΔK method (Evanno et al., 2005) to identify the most likely number of genetic clusters and also used the highest number of genetic clusters that made biological sense (Meirmans, 2015). We used the program VCFTOOLS (Danecek et al., 2011) to calculate Weir and Cockerman mean and weighted F_{ST} values among each population of *L. gramineus* as well as *L. susanae*.

2.10 | Species-tree estimation of SNP data

Using the genetic clusters identified in STRUCTURE analyses (see Results), we used two species-tree methods to identify the relationships between genetic clusters. Each species-tree method was performed on two data sets with special focus on *L. gramineus*. First, we used SVDquartets (Chifman & Kubatko, 2014) implemented in PAUP* v.4.0.a151 (Swofford, 2003). SVDquartets infers unrooted phylogenies for quartets of individuals and then uses sampled

quartets to infer a species-tree. Here, we used all possible quartets for species-tree inference and assessed support with 100 bootstrap replicates. The second species-tree method we used was TreeMix (Pickrell & Pritchard, 2012). TreeMix infers a maximum-likelihood phylogeny and then tries to link genetic clusters that are more closely related to one another than can be explained by the tree topology. These links are candidate migration edges. We ran TreeMix with all SNPs and assessed confidence in the tree topology using 100 bootstrap replicates across bootstrap blocks of 50 SNPs. We added migration edges until they explained >99.8% of the variance in the SNP data (Pickrell & Pritchard, 2012).

3 | RESULTS

3.1 | Estimates of evolutionary relationships

We recovered the Ethiopian Highland *Leptopelis* as a monophyletic group in all phylogenetic analyses, with strong support (Figures 2 and 3). Our data set of 16s rRNA placed all Ethiopian endemics as the sister group to all other members of the genus, with the exception of *L. macrotis* and *L. millsoni* (Figure S1). In the nuclear amplicon data set, the Ethiopian *Leptopelis* formed the sister group to all other members of the genus, with the exclusion of *L. parkeri* (Figure 2a), while in the mitochondrial 3-gene data set we recovered the Ethiopian species as the sister group to all other *Leptopelis*, with the exclusion of *L. barbouri* (Figure 2b). However, the taxon sampling between the two data sets is different, with less taxa sampled for the mitochondrial 3-gene data set. The other fossorial species of *Leptopelis* included in the analysis of the mitochondrial 16s rRNA, *L. bocagii*, was the sister lineage to *L. chrysti* and not closely related to the Ethiopian species (Figure S1). We removed the sequences of 16s and RAG1 of *L. parbocagii* deposited in GenBank from the analysis (JX996026 and KC005991) as they actually represent sequences of a toad of the genus *Sclerophrys*, based on their closest *BlastN* matches in GenBank (*E*-value 0.0; 96% and 99% identity, respectively). The mitochondrial gene CytB did not group *L. parbocagii* with *L. gramineus*. No sequences were available for *L. bufonides*, the only remaining fossorial species in the genus.

In the Ethiopian species, all data sets recovered an "arboreal" lineage composed of *Leptopelis ragazzii*, *L. vannutellii* and *L. yaldeni*. The nuclear loci and ddRAD-seq data sets recovered these species as monophyletic (Figure 2a), while in the mtDNA analysis an individual of *L. ragazzii* (15.144) was the sister lineage to all three species, but with low support (Figure 2b). The relationships within the group differed slightly between data sets, as *L. ragazzii* and *L. vannutellii* were grouped in the concatenated nuclear data set, while *L. yaldeni* and *L. vannutellii* formed a lineage in the mtDNA data set, nuclear species-tree and the ddRAD-seq ML analysis. In all cases, this arboreal lineage was the sister group to *L. gramineus* plus *L. susanae*. We call these later two species the *L. gramineus* complex. In all data sets and analyses, we recovered *L. susanae* nested within *L. gramineus*. We also recovered seven other lineages that correspond to geographically distinct populations of the *L. gramineus* complex (Figures 2, 3, S2 and S3): (i) *Northern lineage*—all samples north and west of the Great Rift Valley and east of the Omo River; (ii) *Jimma lineage*—samples near the town of Jimma in SW Ethiopia, west of the Omo River; (iii) *Bale Mountains lineage*—samples from the Bale Mountains area, east of the Great Rift Valley; (iv) *Assela lineage*—samples from the northeastern parts of the Arsi plateau, near the town of Assela and in the vicinity of Mt. Gugu. Samples from this lineage were not represented in the amplicon data set of nuclear or mitochondrial genes; (v) *Kofele lineage*—samples from the western end of the Arsi plateau, near the town of Kofele; (vi) *Kibre Mengist lineage*—individuals from the southern end of the country, east of the Great Rift Valley. Most individuals were collected near the town of Kibre Mengist; (vii) *Harenna forest lineage*—individuals collected in the Harenna forest, south of the Bale Mountains in southern Ethiopia.

The relationships among the various lineages in the *L. gramineus* complex differed slightly between analyses and data sets. Most analyses and data sets recovered an early split between the *Northern + Jimma* lineages and all populations east of the GRV (plus *L. susanae*; Figures 2a and 3); however, the mtDNA data set recovered the *Bale* and *Kofele* populations as the sister taxa to all other populations (Figures 2b and S1).

The placement of *L. susanae* also differed between analyses, and it was recovered either as the sister group to all other populations east of the GRV (concatenated nuclear, species-tree and SVDquartets; Figures 2a and 3e,f), or with the exclusion of the *Bale Mountains* and *Assela* populations (ML analysis of ddRAD-seq data & TreeMix; Figures 3a,g and S2); however, in all analyses, *L. susanae* was nested within *L. gramineus*. Samples from the *Bale Mountains* and *Assela* lineages were recovered as sister taxa in all analyses, while the major disagreement among analyses and data sets involves the affinities of the *Kofele*, *Harenna* and *Kibre Mengist* lineages, with different sister-taxon relationships recovered by the ML, SVDquartets, species-tree and TreeMix analyses (Figure 3e–g).

3.2 | Population structure, migration and nucleotide diversity

STRUCTURE analyses recovered each of the three arboreal species of *Leptopelis* as a single genetic cluster, and we found no population structure within each species (Figure 3b). On the other hand, we found strong population structure among populations in the *L. gramineus* complex, with a value of $K = 8$ as the most supported (Figure 3b). All populations recovered in STRUCTURE correspond to the major lineages recovered in the concatenated ML analysis of ddRAD-seq data (Figures 3a and S2), and there was little evidence for admixture between populations of *L. gramineus*. We found low levels of admixture between populations occurring east of the GRV, in the Arsi plateau and the Bale Mountains. These include admixture between the *Bale Mountains* and the *Assela* lineages, and again between the *Kibre Mengist* and *Kofele* lineages.

We found a large proportion of private alleles in all populations, from ~30% (*Kofele*) to ~50% (*Harenna* and *Jimma*); however,

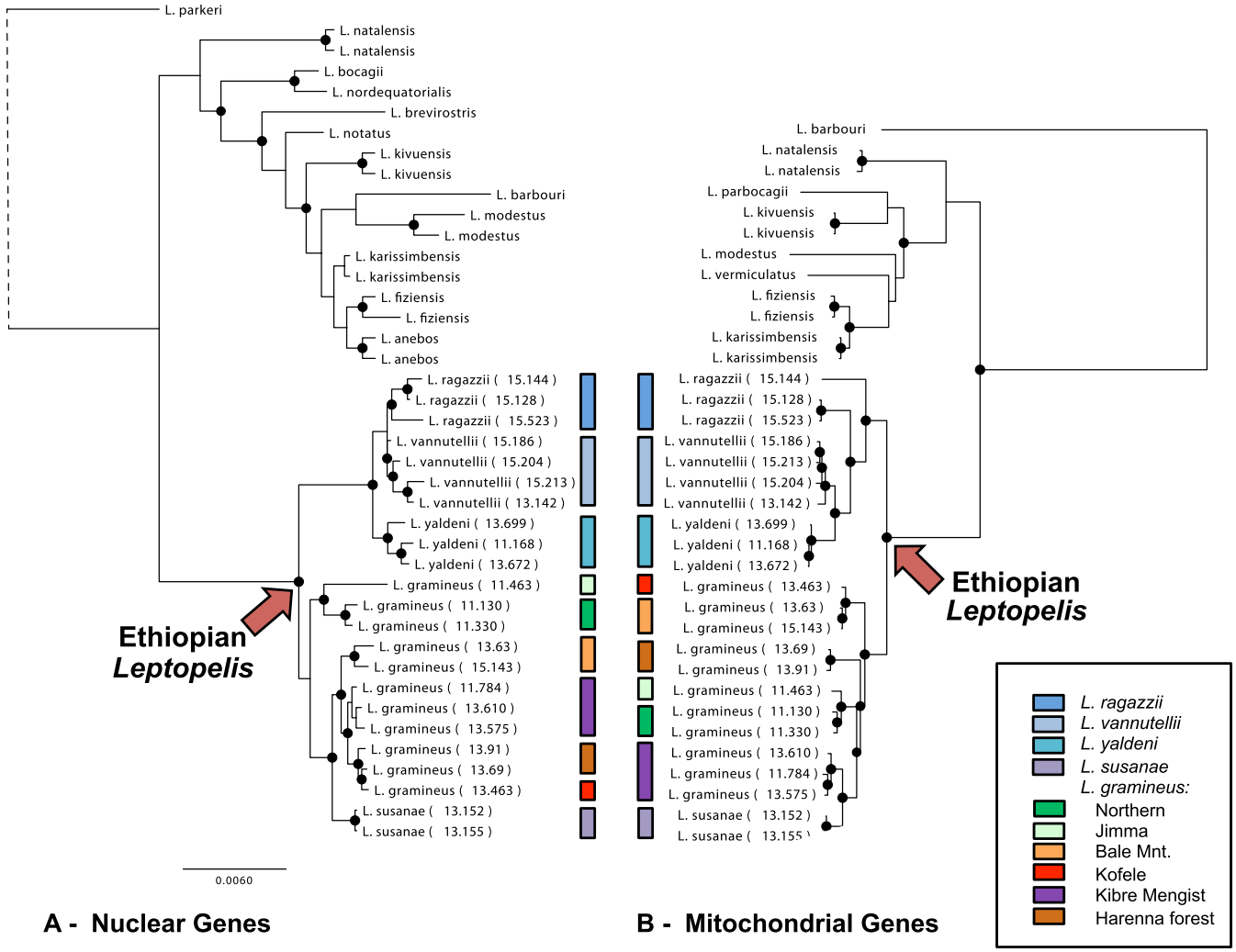
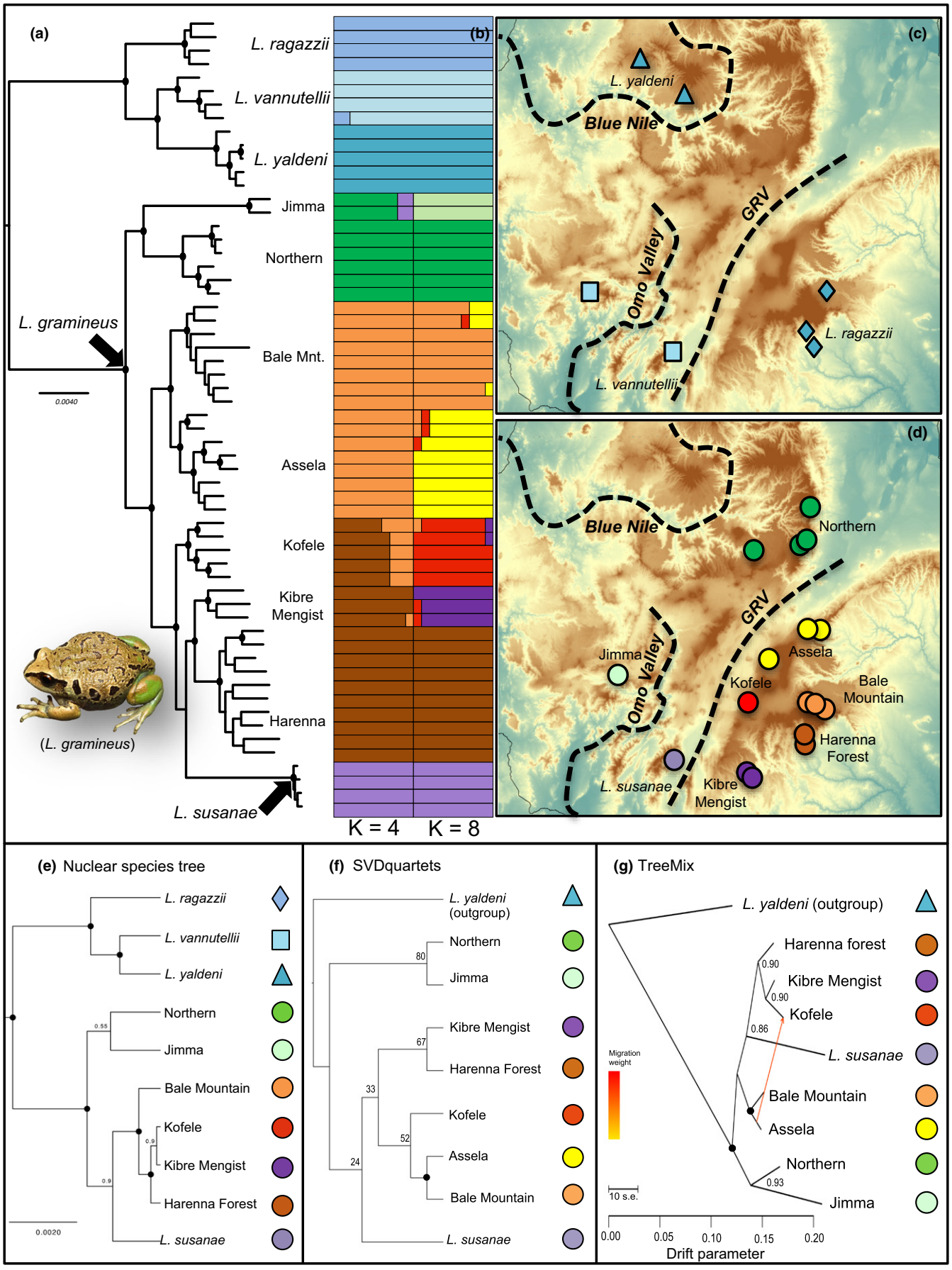


FIGURE 2 Bayesian phylogenetic inference of nuclear (a) and mitochondrial (b) loci used in this study. The nuclear data set consists of 14 protein-coding concatenated nuclear loci and the mitochondrial data set consists of the COX1, Cytb and 16s genes. Nodes with high posterior support (>0.95) are noted with a black circle. Coloured rectangles represent species or populations of particular sample(s) [Colour figure can be viewed at wileyonlinelibrary.com]

fixed differences were less common, and were only abundant in the *Northern* and *Jimma* populations, as well as in *L. susanae* (Figure 4a; Table S10). We detected high levels of population differentiation among *L. gramineus* populations, with high variation in weighted F_{ST} values between populations (Figure 4b; Table S11). We found the highest weighted F_{ST} values between the *Jimma* population and *L. susanae* (0.82), and between *L. susanae* and the *Northern* population. We also found high F_{ST} levels between populations east and west of the GRV (0.50–0.64). East of the GRV, F_{ST} values ranged from 0.15 (between *Kofele* and *Kibre Mengist* populations) to 0.36 (between *Bale* and *Harenna* populations). Despite high levels of shared alleles between *L. gramineus*

FIGURE 3 Phylogeny and genetic structure in Ethiopian Highlands *Leptopelis*. (a) Maximum-likelihood tree of Ethiopian *Leptopelis* based on ~1,600 concatenated loci from the ddRAD-seq data (~35,000 SNPs). Black dots represent nodes with bootstrap support of 100. (b) STRUCTURE plot based on two different analyses. The first included the arboreal *Leptopelis* (*ragazzii*, *vannutellii* and *yaldeni*) and 6,435 unlinked SNPs. The most supported value of *K* was *K* = 3. The second analysis included the members of the *L. gramineus* complex (*gramineus* and *susanae*) and was run with 3,557 independent SNPs. The most supported value for the *L. gramineus* complex was *K* = 8, with minor support for *K* = 4. (c and d) Localities of *Leptopelis* in the arboreal (c) and *L. gramineus* complex (d) lineages. Colours correspond to the lineages in the STRUCTURE analyses. (e) Multilocus coalescence-based tree (species-tree) of Ethiopian *Leptopelis*, inferred from 14 nuclear loci in the program *BEAST. Black circles at nodes represent posterior support >0.95. (f) Species-tree recovered from SVDquartets using ddRAD-seq SNP data, sampling 100,000 quartets. The analysis was run twice, once with all loci present in at least one member of each lineage, and again with loci present it at least 50% of all members of a lineage. The topology did not change between the two analyses. Black circles represent nodes with >95% bootstrap support. (g) TreeMix species-tree inferred using all loci that had at least one individual from each lineage. The phylogeny explains ~99.66% of the variance in SNP data, while the migration edge explains about 0.14%. A migration event from the Assela lineage to the Kofele lineage is depicted by a red arrow. GRV, Great Rift Valley [Colour figure can be viewed at wileyonlinelibrary.com]


(a)
L. ragazzii
L. vannutellii
L. yaldeni
Jimma
Northern
L. gramineus
Bale Mnt.
0.0040
Assela
Kofele
Kibre Mengist
Harenna
(L. gramineus)
L. susanae
(b)
K = 4
K = 8
(c)
L. yaldeni
Blue Nile
GRV
Omo Valley
L. ragazzii
L. vannutellii
(d)
Blue Nile
Northern
GRV
Jimma
Assela
Omo Valley
Kofele
Bale Mountain
Harenna Forest
L. susanae
Kibre Mengist
(e) Nuclear species tree
L. ragazzii
L. vannutellii
L. yaldeni
Northern
Jimma
Bale Mountain
Kofele
Kibre Mengist
Harenna Forest
L. susanae
0.55
0.9
0.9
0.0020
(f) SVDquartets
L. yaldeni (outgroup)
Northern
Jimma
Kibre Mengist
Harenna Forest
Kofele
Assela
Bale Mountain
L. susanae
80
67
33
52
24
(g) TreeMix
L. yaldeni (outgroup)
Harenna forest
Kibre Mengist
Kofele
L. susanae
Bale Mountain
Assela
Northern
Jimma
0.90
0.90
0.86
0.93
Migration weight
10 s.e.
0.00
0.05
0.10
0.15
0.20
Drift parameter

populations east of the GRV, the TreeMix analyses identified a single migration event between divergent lineages of *L. gramineus* (from Assela to Kofele; Figure 3g), consistent with high levels of shared alleles between *L. gramineus* populations east of the GRV. This migration event explained only ~0.1% of the total variance in the SNP data.

3.3 | Divergence time estimates

Our divergence time estimates suggest that the genus *Leptopelis* diverged from other members of the family Arthroleptidae ~50 million years ago, similar to the ages reported by Portik and Blackburn (2016) and Feng et al. (2017), as would be expected, as many of the samples and calibration points were obtained from those studies. The split between the Ethiopian *Leptopelis* and the rest of the genus (excluding *L. parkeri*) occurred during the Miocene, ~16 mya, followed by the split of the arboreal forms (*L. ragazzii*, *L. vannutellii* and *L. yaldeni*) from *L. gramineus* + *L. susanae*, ~8 mya (Figure 5). The populations of *L. gramineus* from east and west of the Great Rift Valley (GRV) split from each other around the Miocene-Pliocene junction (~6 mya), approximately around the same time that *L. ragazzii* split from the other two arboreal forms, which are separated from this species by the GRV. The different populations of *L. gramineus* east of the GRV (plus *L. susanae*) diverged from one another during the Pliocene (~4 mya), which is a similar time frame as the splitting between the arboreal species *L. vannutellii* and *L. yaldeni*. Our divergence time estimates obtained using the mtDNA substitution rate are consistent with those using the nuclear data set with calibration points (Figure S5) and the 95% highest posterior densities (HPD) of the estimated dates overlapped between both dating strategies. The split between the Ethiopian *Leptopelis* and other members of the genus was estimated to be ~15.6 mya (vs. 16 mya), while the split between the arboreal clade and the *gramineus* group was estimated at ~9.6 mya with the mitochondrial estimate (vs. 8 mya). In *L. gramineus*, the divergence estimate between populations from east vs. west was 5 mya (vs. 6 mya).

4 | DISCUSSION

4.1 | Origin and diversification of the Ethiopian *Leptopelis*

Our analyses of mitochondrial and nuclear loci suggest that the endemic highland *Leptopelis* of Ethiopia form a monophyletic group that split from the rest of *Leptopelis* during the Miocene (Figures 5 and S5). However, not all species in the genus were included in our analyses, and it is possible that an unsampled species is closer to the Ethiopian *Leptopelis* than are the other *Leptopelis* species included in this study. Our analyses, however, showed that *L. gramineus* is not closely related to the fossorial species *L. bocagii* and *L. parbocagii*; therefore, its fossorial lifestyle likely evolved in situ and independently from other species in the genus. The split between the Ethiopian *Leptopelis* and their sister taxon from other regions occurred about 16 million years ago, as the Ethiopian species appear to be the sister group to the majority of other *Leptopelis* (Figures 2 and 5). These result emphasizes the unique nature of the Ethiopian Highlands as an ancient centre of diversification for amphibians, as was previously shown in several genera of anurans, including *Balebreviceps* (Loader et al., 2014), *Ericabatrachus* (Siu-Ting et al., 2014) and *Ptychadena* (Freilich et al., 2014; Reyes-Velasco et al., 2018; Smith et al., 2017). This is consistent with the large fraction of endemic frog taxa in the region, including a number of endemic genera such as *Altiphrynoides*, *Balebreviceps*, *Ericabatrachus* and *Paracassina* (Largen & Spawls, 2010).

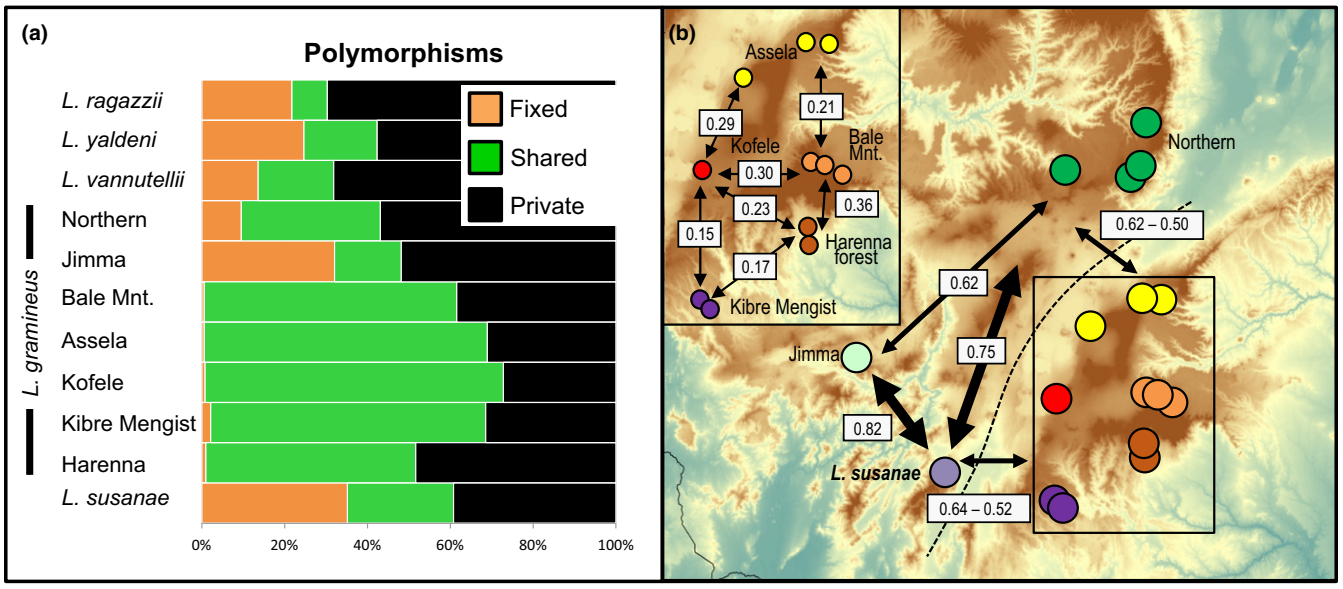


FIGURE 4 (a) Proportion of fixed, shared and private alleles in species and populations of *Leptopelis* based on the ddRAD-seq loci data set. The total number of polymorphisms for each species and population is found in Table S10. (b) F_{ST} values for different populations of the *L. gramineus* complex. Numbers in boxes represent weighted F_{ST} values between each pair of populations [Colour figure can be viewed at wileyonlinelibrary.com]

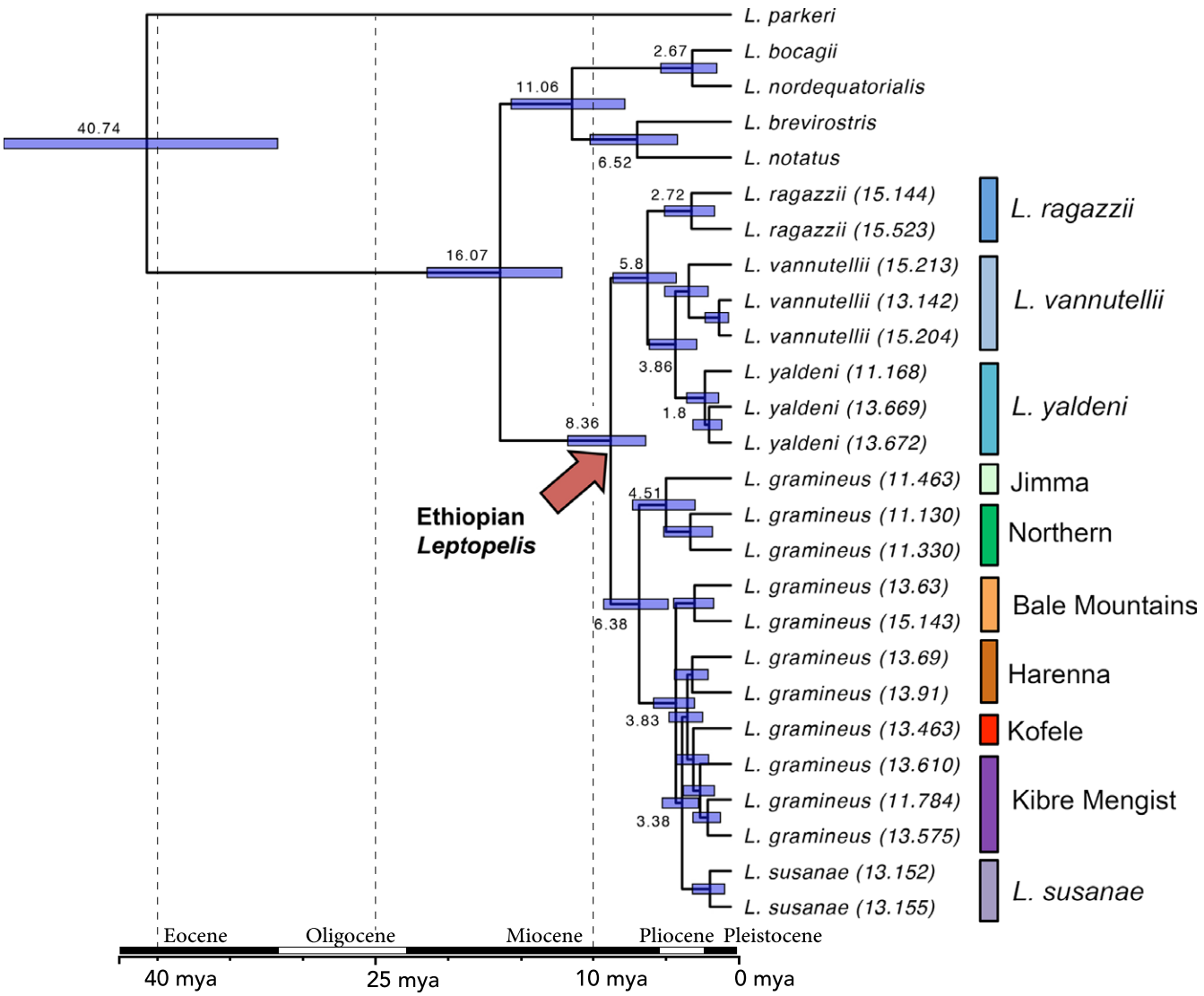


FIGURE 5 Divergence time estimates of members of the genus *Leptopelis* in the Ethiopian highlands. The time calibration was inferred using Bayesian inference in BEAST v2.4.3, with a relaxed lognormal clock model and two calibration points (not shown). Dates at nodes represent divergence date estimates, while purple bars indicate 95% highest posterior density region of dates. We constrained the topologies previous to the divergence estimate, using the resulting topology from the species-tree estimate of nuclear loci [Colour figure can be viewed at wileyonlinelibrary.com]

The split between members of the *L. gramineus* complex and the arboreal species (*L. ragazzii*, *L. vannutellii* and *L. yaldeni*) occurred during the late Miocene. This period coincides with a time of intense tectonic activity and subsequent periods of climatic shifts in Ethiopia (Sepulchre et al., 2006). After the split between the arboreal *Leptopelis* and the *L. gramineus* complex, the Great Rift Valley (GRV) separated both species groups into eastern and western lineages during the early Pliocene, a period that corresponds to the major development of the Rift system (Davidson & Rex, 1980). Populations of *L. gramineus* were split into eastern and western lineages, while in the arboreal species, *L. ragazzii* was isolated to the east of the GRV from the ancestor of *L. vannutellii* and *L. yaldeni*. The GRV has shaped the genetic structure of a multitude of plants and animals. These include the frog genera *Amietia*, *Ptychadena* and *Xenopus* (Evans et al., 2011; Freilich et al., 2016; Manthey et al., 2017), mammals, for example, the Gelada baboon and the Ethiopian wolf (Belay & Mori, 2006; Gottelli et al., 2004), as well as plants (Kebede et al., 2007; Silvestrini et al., 2007). The split between populations on either side of the GRV appears to be older in *Leptopelis* (~4–6 mya) than in other genera of frogs such as *Amietia* (0.4 mya), *Ptychadena* (0.4–0.5 mya) and *Xenopus* (~1–3.5 mya).

4.2 | Diversification and cryptic lineages in *Leptopelis gramineus*

Our results show that the different populations of *L. gramineus* separated approximately 3–6 million years ago (Figures 5 and S5). The multiple lineages appear to be allopatric and show high levels of

genetic differentiation, even though some of those populations are geographically in close proximity, with no obvious biogeographic barrier separating the different lineages. This is particularly true for the five populations east of the GRV that were identified by our phylogenetic and STRUCTURE analyses (Figures 2 and 3). Comparisons with codistributed anurans suggest that this pattern is unique to *L. gramineus*. The arboreal species *L. ragazzii* occurs in both the Bale Mountains and the Harenna forest, and we found little genetic structure between these populations (Figure 3b), while there appears to be no gene flow between populations from these two areas in *L. gramineus*. Studies of other frog species that inhabit the Arsi plateau revealed no evidence of population structure within the plateau, for example, within *Amietia nutti* (Freilich et al., 2014; Manthey et al., 2017), *Xenopus clivii* (Evans et al., 2011), *Ptychadena cooperi* and *P.* cf. *neumanni* 2 (Freilich et al., 2014, 2016; Reyes-Velasco et al., 2018). *Ptychadena* cf. *neumanni* 4 occurs in the same forests as the *Kibre Mengist* and *Harenna* populations of *L. gramineus*, yet exhibits no population structure (Freilich et al., 2014; Reyes-Velasco et al., 2018). Similarly, west of the GRV, *P. cf. neumanni* 1 occurs in the central plateau as well as in the area around Jimma, but again shows no population structure (Freilich et al., 2014), which is in contrast to what we found in *L. gramineus*. The studies cited above used similar data sets as the present study (ddRAD-seq as well as nuclear and mitochondrial loci), so it is possible to make direct comparisons between studies.

The divergence times among many of the *L. gramineus* populations predate the Last Glacial Maximum (Figures 5 and S5) and correspond to a time of cooler and dryer conditions than those present today. Based on our results, we propose the following biogeographic scenario for the diversification of *L. gramineus* in the Ethiopian highlands. Following the split from the ancestor of the arboreal species during the late Miocene and the colonization of the Ethiopian highlands, several vicariant events fragmented the ancestral population of *L. gramineus*. First, environmental changes in the GRV (Trauth, Larrasoana, & Mudelsee, 2009) made it inhospitable for the species and thus created a barrier to gene flow across the Rift, isolating *L. gramineus* into northern and southern lineages. In the populations north of the GRV, the Omo River valley, with was already present ~4.5 mya (Brown & Heinzelin, 1983), probably created a barrier between populations east and west of it (*Northern* and *Jimma* lineages, respectively). The populations of *L. gramineus* east of the GRV probably differentiated from each other around the late Pliocene and early Pleistocene (Figures 5 and S5). These estimates are older than those reported by Freilich et al. (2016), who used an approximate nuclear mutation rate and a small number of slow-evolving nuclear genes to infer divergence time between populations. In the Harar Massif of southern Ethiopia, multiple glaciation events occurred during the Pleistocene (Mark & Osmaston, 2008; Mohammed & Bonnefille, 1998). These events probably isolated the multiple lineages of *L. gramineus* (Figures 2 and 3), as the glaciers and permafrost pushed down their habitats and created barriers between populations. Glaciers in the Bale Mountains in particular could have isolated populations east and west of it (*Bale* and *Assela + Kofele* lineages, respectively), while these cooler conditions forced some lineages to move to lower elevations (e.g., *Kibre Mengist* and *Harenna*). These areas probably served as refugia during these periods. Glaciation events in the Bale Mountains were also suggested to be involved in the diversification of frogs of the genus *Ptychadena* (Reyes-Velasco et al., 2018).

The levels of divergence between the lineages of *L. gramineus* suggest that these lineages in fact represent different species. This hypothesis is supported by the fact that the divergences between *L. gramineus* lineages are similar to those found between species of arboreal *Leptopelis* (Figure 5). Multiple studies have shown that many species pairs of African frogs diversified around the Miocene/Pliocene junction. For example, Freilich et al. (2014) report divergence times between species in the *Ptychadena neumanni* species complex to be between 2 and 6 million years, while Portik and Blackburn (2016) report similar ages for many species pairs of Afrobatrachians.

Following the vicariant events that separated the multiple populations of *L. gramineus*, the current distribution of the multiple lineages could have been restricted for several reasons. First, it is possible that after climatic changes restored more favourable conditions, upon secondary contact, the vicariant populations exclude each other by incumbency (Schenk, Rowe, & Steppan, 2013), even in the absence of physical barriers to dispersal. Second, it remains possible that several geological or topographical features of the Ethiopian Highlands play a role in limiting dispersal in *L. gramineus* but not in other taxa. For instance, the samples from *Bale Mountains* are isolated from other populations by the headwaters of the Shebelle River to the west and the Sanetti plateau to the south. Similarly, the *Kibre Mengist* and *Harenna* lineages are separated by the low and dry reaches of the Ganale Doria River valley. Yet, no clear geographic barriers separate the populations from Assela, Kofele and Kibre Mengist. Finally, a role of ecology cannot be excluded. Recent studies have shown that fossorial species tend to have low dispersal abilities and smaller geographic distributions than nonfossorial frogs (Eggert, 2002; Penner & Rödel, 2017; Sánchez-Montes, Wang, Ariño, & Martínez-Solano, 2017; Székely, Cogălniceanu, Székely, & Denoël, 2017) and it is thus possible that fossoriality played a role in the limited distribution of the different linages in *L. gramineus*.

4.3 | Re-evolution of an arboreal lifestyle

Our analyses show that *L. susanae* (an arboreal species) is nested within populations of the fossorial species *L. gramineus*. A shift from a fossorial lifestyle to an arboreal one in *L. susanae* is a more parsimonious explanation than repeated instances of a switch from an arboreal ecology to fossoriality in multiple populations of *L. gramineus*. *Leptopelis susanae* possesses larger digital discs than *L. gramineus*, which might aid it for climbing in vegetation (Largen, 1977). The metatarsal tubercle (which is used for digging in multiple species of frogs) is reduced in *L. susanae* (Largen, 1977), while it remains enlarged in *L. gramineus*. While conducting fieldwork, we found immature individuals of *L. gramineus* climbing on vegetation on several occasions, which is rarely observed in adults. It is possible

that the retention of some neotenic and plesiomorphic characters could have aided *L. susanae* in re-adapting to an arboreal lifestyle. Fossoriality is common in frogs and is present in at least 19 frog families (Nomura, Rossa-Feres, & Langeani, 2009). In some families like the African rain frogs (Brevicipitidae), a fossorial ecology is the norm, and only a few species exhibit a different lifestyle (Portik & Blackburn, 2016). Shifts in the ecological preferences of frogs are common, for example, switching between terrestrial, arboreal or aquatic lifestyles. In contrast, a shift from a fossorial lifestyle to an arboreal lifestyle is very rare (Moen, Morlon, & Wiens, 2016; Portik & Blackburn, 2016) and to our knowledge has been reported only for the family Microhylidae (Blackburn et al., 2013). Portik and Blackburn (2016) performed a thorough analysis on the evolution of ecological characters in the family Arthroleptidae, which includes the genus *Leptopelis*. They found that an arboreal ecology is the ancestral state in the genus *Leptopelis*, with only a few of species adapted to a fossorial ecology, including *L. gramineus*. Little is known about the ecology of *L. susanae*, as few specimens have ever been collected. However, adults are known to call from high up in the vegetation, just as other arboreal species of the genus.

4.4 | Taxonomic implications, endemism and conservation

The amphibian fauna of Africa has received much less attention than other areas in the world, and this is particularly true when it comes to studies using genetic data (Miraldo et al., 2016). We identified several cryptic lineages in *L. gramineus* with the use of multiple genetic markers. Some of the splits in *L. gramineus* are very old, and the genetic differences between them so large that they likely deserve species status. While we do not recommend splitting the different populations of *L. gramineus* into separate species at this point, it is clear that a detailed study on the morphological and mate-call variation across the range of the species is warranted, and it is currently under way.

Our results show that *Leptopelis susanae* is nested within *L. gramineus*, and not the sister taxon to it as previously reported (Portik & Blackburn, 2016). Yet, we maintain this taxon as valid, as it represents a unique lineage on its own evolutionary trajectory, presents unique and distinguishable characters that separate it from all populations of *L. gramineus* and has re-evolved an arboreal lifestyle.

The geographic areas where we identified the distinct lineages of *L. gramineus* are known to harbour several endemic species of amphibians and other vertebrates. The caecilian *Sylvacaecilia grandisonae* is endemic to the forest near the town of Jimma, while four undescribed species of frogs of the genus *Ptychadena* are endemic to this region (Reyes-Velasco et al., 2018). This area also harbours some of the last remaining large patches of forest in Ethiopia (J. Reyes-Velasco, X. Freilich & S. Boissinot, personal observation, August 2015). The Harenna forest and the Bale Mountains in southern Ethiopia are important biodiversity centres for amphibians and other vertebrates, including two genera and three species of frogs that are strictly endemic to the Harenna forest. One genus and two species of toads are endemic to the Bale Mountains National Park (Largen & Spawls, 2010), while three endemic species of grass frogs in genus *Ptychadena* are also found there (Freilich et al., 2014; Reyes-Velasco et al., 2018). Three species of reptiles (Gower et al., 2016; Largen, 1995; Tilbury, 1998) and a species of vervet monkey (Neumann, 1902) are also endemic to the Bale Mountains National Park. Our discovery of novel lineages of *L. gramineus* in the forests of southern Ethiopia provides further evidence that these areas constitute important biodiversity hot spots that deserve urgent measures for their conservation.

The population of Ethiopia has increased more than ten times in the last 60 years, which has influenced the need for arable land to feed the still growing population (Williams et al., 2004). The increased need for rangeland is also a major concern, especially in areas like the Bale Mountains or the Harenna forest, as the density of livestock is extremely high and the boundaries of the national park are not respected (Gower et al., 2013). It is unknown how the species and populations of Ethiopian *Leptopelis* will cope with climate change (Gower et al., 2013), but we hypothesize that many of the populations of *Leptopelis* that are adapted to high elevations will be in jeopardy, especially those at the Bale Mountains National Park. The continuing encroachment on the few remaining natural areas and current national parks by populations centres and agriculture could jeopardize the survival of some of the *Leptopelis* in Ethiopia, particularly those with small and isolated ranges like *L. susanae*.

5 | CONCLUSIONS AND FUTURE STUDIES

Our study is among a few that have used high-throughput sequencing to study genetic diversity of east African amphibians, particularly in the Ethiopian highlands. Our results indicate that the *Leptopelis* of the Ethiopian highlands are older and more diverse than previously estimated, with the possible existence of several cryptic species in *L. gramineus*. However, a thorough revision of morphological characters is necessary before applying species names to the different lineages that we identified. Studies on the advertisement calls between populations are also warranted, as mating calls may serve as a premating isolation mechanism. We also show that an arboreal lifestyle re-evolved in *L. susanae*, which might have involved the retention of neotenic and plesiomorphic characters in this species. Our study highlights areas of future study. For example, it will be necessary to ask whether the genetic differences between populations of *L. gramineus* are spread evenly across the genome or whether variation is restricted to particular genomic regions under selection due to local conditions.

ACKNOWLEDGEMENTS

We thank the Ethiopian Wildlife Conservation Authority for providing us with permits to work in Ethiopia. Collection of the samples would not have been possible without the help of many field hands over the years, including Robert Ruggiero, Yann Bourgeois, Sudikchya Shrestha,

Jigar Govind, Mona Nehme, Victor Okoth and Jayson Slovak. Yvonne and Guy Levene provided lodging and friendship while at Bale Mountains National Park, and Sami, Megersa and Solomon of Rock Hewn Tours were indispensable for the planning, transportation, translation and fieldwork during many years. Isaac Overcast provided valuable assistance with the use of ipyrad. Eli Greenbaum kindly provided tissue samples. We also thank Marc Arnoux and Nizar Drou, from the Genome Core Facility and the Bioinformatics group at NYUAD. We thank four anonymous reviewers for their comments, which greatly improved the quality of this manuscript.

AUTHOR CONTRIBUTIONS

J.R.-V., J.D.M. and S.B. designed the research. J.R.-V. and J.D.M. performed the analyses and analysed the data. J.R.-V., J.D.M., X.F. and S.B. participated in fieldwork and wrote the study.

DATA ACCESSIBILITY

All mitochondrial and nuclear DNA Sanger sequencing data were uploaded to NCBI's GenBank. GenBank Accession nos are provided in Table S1. All Illumina reads were deposited at NCBI's Sequence Read Archive with Accession no. SRP129301.

REFERENCES

Belay, G., & Mori, A. (2006). Intraspecific phylogeographic mitochondrial DNA (D-loop) variation of Gelada baboon, *Theropithecus gelada*, in Ethiopia. *Biochemical Systematics and Ecology*, *34*, 554–561. https://doi.org/10.1016/j.bse.2006.01.004

Blackburn, D. C., Siler, C. D., Diesmos, A. C., McGuire, J. A., Cannatella, D. C., & Brown, R. M. (2013). An adaptive radiation of frogs in a Southeast Asian island archipelago. *Evolution*, *67*, 2631–2646. https://doi.org/10.1111/evo.12145

Bossuyt, F., & Milinkovitch, M. C. (2000). Convergent adaptive radiations in Madagascan and Asian ranid frogs reveal covariation between larval and adult traits. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 6585–6590. https://doi.org/10.1073/pnas.97.12.6585

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., … Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Brown, F. H., Heinzelin, J. D. (1983). The lower Omo basin. In J. D. Heinzelin (Ed.), *The Omo Group*. Archives of the International Omo Research Expedition, Annales des Sciences Geologiques (pp. 7–24). Tervuren, Belgium: Musee Royal de l'Afrique Centrale.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, *17*, 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334

Channing, A. (2001). *Amphibians of central and southern Africa*. Ithaca, NY: Comstock Pub. Associates.

Chifman, J., & Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics*, *30*, 3317–3324. https://doi.org/10.1093/bioinformatics/btu530

Crawford, A. J. (2003). Relative rates of nucleotide substitution in frogs. *Journal of Molecular Evolution*, *57*, 636–641. https://doi.org/10.1007/s00239-003-2513-7

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … McVean, G. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Davidson, A., & Rex, D. (1980). Age of volcanism and rifting in southwestern Ethiopia. *Nature*, *283*, 657–658. https://doi.org/10.1038/283657a0

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*, 1.

Eaton, D. A. (2014). PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, *30*, 1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Eaton, D. A., & Overcast, I. (2017). *ipyrad: interactive assembly and analysis of RADseq*. Retrieved from http://ipyrad.readthedocs.io/

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*, 1792–1797. https://doi.org/10.1093/nar/gkh340

Eggert, C. (2002). Use of fluorescent pigments and implantable transmitters to track a fossorial toad (*Pelobates fuscus*). *Herpetological Journal*, *12*, 69–74.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Evans, B. J., Bliss, S. M., Mendel, S. A., & Tinsley, R. C. (2011). The Rift Valley is a major barrier to dispersal of African clawed frogs (*Xenopus*) in Ethiopia. *Molecular Ecology*, *20*, 4216–4230. https://doi.org/10.1111/j.1365-294X.2011.05262.x

Feng, Y. J., Blackburn, D. C., Liang, D., Hillis, D. M., Wake, D. B., Cannatella, D. C., & Zhang, P. (2017). Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous–Paleogene boundary. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, E5864–E5870. https://doi.org/10.1073/pnas.1704632114

Feng, Y. J., Liu, Q. F., Chen, M. Y., Liang, D., & Zhang, P. (2016). Parallel tagged amplicon sequencing of relatively long PCR products using the Illumina HiSeq platform and transcriptome assembly. *Molecular Ecology Resources*, *16*, 91–102. https://doi.org/10.1111/1755-0998.12429

Freilich, X., Anadón, J. D., Bukala, J., Calderon, O., Chakraborty, R., & Boissinot, S. (2016). Comparative Phylogeography of Ethiopian anurans: Impact of the Great Rift Valley and Pleistocene climate change. *BMC Evolutionary Biology*, *16*, 206. https://doi.org/10.1186/s12862-016-0774-1

Freilich, X., Tollis, M., & Boissinot, S. (2014). Hiding in the highlands: Evolution of a frog species complex of the genus *Ptychadena* in the Ethiopian highlands. *Molecular Phylogenetics and Evolution*, *71*, 157–169. https://doi.org/10.1016/j.ympev.2013.11.015

Gordon, A., & Hannon, G. (2010). *Fastx-toolkit*. FASTQ/A short-reads preprocessing tools (unpublished). Retrieved from http://hannonlab.cshl.edu/fastx_toolkit

Gottelli, D., Marino, J., Sillero-Zubiri, C., & Funk, S. M. (2004). The effect of the last glacial age on speciation and population genetic structure of the endangered Ethiopian wolf (*Canis simensis*). *Molecular Ecology*, *13*, 2275–2286. https://doi.org/10.1111/j.1365-294X.2004.02226.x

Gower, D. J., Aberra, R. K., Schwaller, S., Largen, M. J., Collen, B., Spawls, S., … Gebresenbet, F. (2013). Long-term data for endemic frog genera reveal potential conservation crisis in the Bale Mountains, Ethiopia. *Oryx*, *47*, 59–69. https://doi.org/10.1017/S0030605311001426

Gower, D. J., Doherty-Bone, T. M., Aberra, R. K., Mengistu, A., Schwaller, S., Menegon, M., … Loader, S. P. (2012). High prevalence of the amphibian chytrid fungus (*Batrachochytrium dendrobatidis*) across multiple taxa and localities in the highlands of Ethiopia. *The Herpetological Journal*, *22*, 225–233.

Gower, D. J., Wade, E. O., Spawls, S., Böhme, W., Buechley, E. R., Sykes, D., & Colston, T. J. (2016). A new large species of *Bitis* Gray, 1842 (Serpentes: Viperidae) from the Bale Mountains of Ethiopia. *Zootaxa*, *4093*, 41–63. https://doi.org/10.11646/zootaxa.4093.1.3

Harper, E. B., Measey, G. J., Patrick, D. A., Menegon, M., Vonesh, J. R., & Swilla, I. (2010) *Field guide to the amphibians of the Eastern Arc Mountains and Coastal Forests of Tanzania and Kenya* (114 pp.). Nairobi, Kenya: Camerapix Publishers International.

Heled, J., & Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, *27*, 570–580. https://doi.org/10.1093/molbev/msp274

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*, 254–267. https://doi.org/10.1093/molbev/msj030

Kebede, M., Ehrich, D., Taberlet, P., Nemomissa, S., & Brochmann, C. (2007). Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopian and Tropical East African mountains. *Molecular Ecology*, *16*, 1233–1243. https://doi.org/10.1111/j.1365-294X.2007.03232.x

Lanfear, R., Calcott, B., Ho, S. Y., & Guindon, S. (2012). PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, *29*, 1695–1701. https://doi.org/10.1093/molbev/mss020

Largen, M. (1977). The status of the genus *Leptopelis* (Amphibia Anura Hyperoliidae) in Ethiopia, including descriptions of two new species. *Monitore Zoologico Italiano. Supplemento*, *9*, 85–136.

Largen, M. J. (1995). A new species of chameleon (Reptilia Sauria Chamaelecnidae) from montane forest in Ethiopia. *Tropical Zoology*, *8*, 333–339. https://doi.org/10.1080/03946975.1995.10539290

Largen, M., & Spawls, S. (2010). *The amphibians and reptiles of Ethiopia and Eritrea*. Frankfurt am Main: Edition Chimaira.

Leache, A. D., Chavez, A. S., Jones, L. N., Grummer, J. A., Gottscho, A. D., & Linkem, C. W. (2015). Phylogenomics of phrynosomatid lizards: Conflicting signals from sequence capture versus restriction site associated DNA sequencing. *Genome Biology and Evolution*, *7*, 706–719. https://doi.org/10.1093/gbe/evv026

Loader, S. P., Sara Ceccarelli, F., Menegon, M., Howell, K. M., Kassahun, R., Mengistu, A. A., … Larson, J. G. (2014). Persistence and stability of Eastern Afromontane forests: Evidence from brevicipitid frogs. *Journal of Biogeography*, *41*, 1781–1792. https://doi.org/10.1111/jbi.12331

Manthey, J., Reyes-Velasco, J., Freilich, X., & Boissinot, S. (2017). Diversification in a biodiversity hotspot: Genomic variation in the river frog *Amietia nutti* across the Ethiopian Highlands. *Biological Journal of the Linnean Society*, *122*, 801–813. https://doi.org/10.1093/biolinnean/blx106

Mark, B. G., & Osmaston, H. A. (2008). Quaternary glaciation in Africa: Key chronologies and climatic implications. *Journal of Quaternary Science*, *23*, 589–608. https://doi.org/10.1002/jqs.1222

Meirmans, P. G. (2015). Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, *24*, 3223–3231. https://doi.org/10.1111/mec.13243

Mengistu, A. A., Nagel, P., Getahun, A., Saber, S. A., & Loader, S. P. (2013). Updated review of amphibian diversity, distribution and conservation in Ethiopia. *Ethiopian Journal of Biological Sciences*, *12*, 81–116.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE): 14 Nov 2010*, New Orleans, (pp. 1–8).

Miraldo, A., Li, S., Borregaard, M. K., Flórez-Rodríguez, A., Gopalakrishnan, S., Rizvanovic, M., … Nogués-Bravo, D. (2016). An Anthropocene map of genetic diversity. *Science*, *353*, 1532–1535. https://doi.org/10.1126/science.aaf4381

Mittermeier, R. A. (2004). *Hotspots revisited*. Cemex.

Moen, D. S., Morlon, H., & Wiens, J. J. (2016). Testing convergence versus history: Convergence dominates phenotypic evolution for over 150 million years in frogs. *Systematic Biology*, *65*, 146–160. https://doi.org/10.1093/sysbio/syv073

Mohammed, M., & Bonnefille, R. (1998). A late Glacial/late Holocene pollen record from a highland peat at Tamsaa, Bale Mountains, south Ethiopia. *Global and Planetary Change*, *16*, 121–129. https://doi.org/10.1016/S0921-8181(98)00025-3

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*, 853. https://doi.org/10.1038/35002501

Neumann, O. (1902). Ueber neue nordost-und ostafrikanische Saugethiere. *Sitzungsberichte der Gesellschaft Naturforschender Freunde zu Berlin*, *1902*, 49–59.

Nomura, F., Rossa-Feres, D. C., & Langeani, F. (2009). Burrowing behavior of *Dermatonotus muelleri* (Anura, Microhylidae) with reference to the origin of the burrowing behavior of Anura. *Journal of Ethology*, *27*, 195. https://doi.org/10.1007/s10164-008-0112-1

Penner, J., & Rödel, M.-O. (2017). Keep it simple? Dispersal abilities can explain why species range sizes differ, the case study of West African amphibians. *Acta Oecologica*.

Perret, J.-L. (1980). Sur Quelques *Ptychadena* (Amphibia Ranidae): D'Ethiopie. *Monitore Zoologico Italiano. Supplemento*, *13*, 151–168. https://doi.org/10.1080/00269786.1980.11758552

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, *7*, e37135. https://doi.org/10.1371/journal.pone.0037135

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, *8*, e1002967. https://doi.org/10.1371/journal.pgen.1002967

Portik, D., & Blackburn, D. (2016). The evolution of reproductive diversity in Afrobatrachia: A phylogenetic comparative analysis of an extensive radiation of African frogs. *Evolution*, *70*, 2017–2032. https://doi.org/10.1111/evo.12997

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Rambaut, A., Suchard, M. A., Xie, D., & Drummond, A. J. (2014). *TreeAnnotator v1.8.2*. Retrieved from http://beast.bio.ed.ac.uk/TreeAnnotator

Reyes-Velasco, J., Manthey, J. D., Bourgeois, Y., Freilich, X., & Boissinot, S. (2018). Revisiting the phylogeography, demography and taxonomy of the frog genus *Ptychadena* in the Ethiopian highlands with the use of genome-wide SNP data. *PLoS One*, *13*, e0190440. https://doi.org/10.1371/journal.pone.0190440

Roelants, K., Gower, D. J., Wilkinson, M., Loader, S. P., Biju, S. D., Guillaume, K., … Bossuyt, F. (2007). Global patterns of diversification in the history of modern amphibians. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 887–892. https://doi.org/10.1073/pnas.0608378104

Rohland, N., & Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research*, *22*, 939–946. https://doi.org/10.1101/gr.128124.111

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., … Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, *61*, 539–542. https://doi.org/10.1093/sysbio/sys029

Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X., & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, *19*, 2496–2497. https://doi.org/10.1093/bioinformatics/btg359

Sánchez-Montes, G., Wang, J., Ariño, A. H., & Martínez-Solano, Í. (2017). Mountains as barriers to gene flow in amphibians: Quantifying the differential effect of a major mountain ridge on the genetic structure of four sympatric species with different life history traits. *Journal of Biogeography*, *45*, 318–331.

Schenk, J. J., Rowe, K. C., & Steppan, S. J. (2013). Ecological opportunity and incumbency in the diversification of repeated continental colonizations by muroid rodents. *Systematic Biology*, *62*, 837–864. https://doi.org/10.1093/sysbio/syt050

Schiøtz, A. (1999). *Treefrogs of Africa*. Frankfurt am Main: Ed. Chimaira.

Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tiercelin, J. J., & Brunet, M. (2006). Tectonic uplift and Eastern Africa aridification. *Science*, *313*, 1419–1423. https://doi.org/10.1126/science.1129158

Shen, X. X., Liang, D., Feng, Y. J., Chen, M. Y., & Zhang, P. (2013). A versatile and highly efficient toolkit including 102 nuclear markers for vertebrate phylogenomics, tested by resolving the higher level relationships of the Caudata. *Molecular Biology and Evolution*, *30*, 2235–2248. https://doi.org/10.1093/molbev/mst122

Silvestrini, M., Junqueira, M. G., Favarin, A. C., Guerreiro-Filho, O., Maluf, M. P., Silvarolla, M. B., & Colombo, C. A. (2007). Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica* L. accessions using microsatellites markers. *Genetic Resources and Crop Evolution*, *54*, 1367–1379. https://doi.org/10.1007/s10722-006-9122-4

Siu-Ting, K., Gower, D. J., Pisani, D., Kassahun, R., Gebresenbet, F., Menegon, M., … Loader, S. P. (2014). Evolutionary relationships of the Critically Endangered frog *Ericabatrachus baleensis* Largen, 1991 with notes on incorporating previously unsampled taxa into large-scale phylogenetic analyses. *BMC Evolutionary Biology*, *14*, 44. https://doi.org/10.1186/1471-2148-14-44

Smith, M. L., Noonan, B. P., & Colston, T. J. (2017). The role of climatic and geological events in generating diversity in Ethiopian grass frogs (genus *Ptychadena*). *Royal Society Open Science*, *4*, 170021. https://doi.org/10.1098/rsos.170021

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Streicher, J. W., Schulte, J. A., & Wiens, J. J. (2015). How should genes and taxa be sampled for phylogenomic analyses with missing data? An empirical study in iguanian lizards. *Systematic Biology*, *65*, 128–145.

Swofford, D. L. (2003). {PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4.}.

Székely, D., Cogălniceanu, D., Székely, P., & Denoël, M. (2017). Out of the ground: Two coexisting fossorial toad species differ in their emergence and movement patterns. *Zoology*, *121*, 49–55. https://doi.org/10.1016/j.zool.2016.12.003

Tilbury, C. (1998). Two new chameleons (Sauria: Chamaeleonidae) from isolated Afromontane forests in Sudan and Ethiopia. *Bonner Zoologische Beitrage*, *47*, 293–300.

Trauth, M. H., Larrasoana, J. C., & Mudelsee, M. (2009). Trends, rhythms and events in Plio-Pleistocene African climate. *Quaternary Science Reviews*, *28*, 399–411. https://doi.org/10.1016/j.quascirev.2008.11.003

Vaidya, G., Lohman, D. J., & Meier, R. (2011). SequenceMatrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, *27*, 171–180. https://doi.org/10.1111/j.1096-0031.2010.00329.x

Weinsheimer, F., Mengistu, A. A., & Rödder, D. (2010). Potential distribution of threatened *Leptopelis* spp. (Anura, Arthroleptidae) in Ethiopia derived from climate and land-cover data. *Endangered Species Research*, *9*, 117–124. https://doi.org/10.3354/esr00231

Williams, S., Vivero Pol, J. L., Spawls, S., Shimelis, A., & Kelbessa, E. (2004). Ethiopian highlands. *Hotspots Revisited: Earths Biologically Richest and Most Endangered Ecoregions.–Mexico City (CEMEX)*.

Zhang, P., Liang, D., Mao, R. L., Hillis, D. M., Wake, D. B., & Cannatella, D. C. (2013). Efficient sequencing of anuran mtDNAs and a mitogenomic exploration of the phylogeny and evolution of frogs. *Molecular Biology and Evolution*, *30*, 1899–1915. https://doi.org/10.1093/molbev/mst091

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information section at the end of the article.

How to cite this article: Reyes-Velasco J, Manthey JD, Freilich X, Boissinot S. Diversification of African tree frogs (genus *Leptopelis*) in the highlands of Ethiopia. *Mol Ecol*. 2018;27:2256–2270. https://doi.org/10.1111/mec.14573