Org Divers Evol (2013) 13:605–620
DOI 10.1007/s13127-013-0143-5

ORIGINAL ARTICLE

ORGANISMS DIVERSITY & EVOLUTION

Northern origin and diversification in the central lowlands? – Complex phylogeography and taxonomy of widespread day geckos (*Phelsuma*) from Madagascar

Philip-Sebastian Gehring · Frank Glaw · Marcelo Gehara · Fanomezana Mihaja Ratsoavina · Miguel Vences

Received: 22 January 2013 / Accepted: 2 June 2013 / Published online: 27 June 2013

Abstract Day geckos of the *Phelsuma lineata* group are widespread in Madagascar and have been historically split into numerous species and subspecies based almost exclusively on differences in coloration and body size. We apply phylogenetic and phylogeographic methods to examine the biogeography and taxonomy of these lizards, including explicit tests of various biogeographic predictions and based on a molecular data set covering much of the distribution ranges of all species and subspecies of *P. lineata, P. dorsivittata, P. comorensis, P. hoeschi, P. kely,* and *P. pusilla* in Madagascar (and the Comoros archipelago for *P. comorensis*). Sequences of the mitochondrial *16S rRNA* and the nuclear *RAG-1* gene fragment were determined from 376 samples, and a multigene mtDNA phylogeny of the species group was constructed for the main phylogroups identified in the *16S* haplotype network. We used the *16S* sequences to estimate the geographic location of the ancestor of each major mtDNA clade and to infer their demographic history using a variety of statistical tools. Our phylogeny separates the taxa analyzed into two well-supported major subclades mainly occurring in the north respectively east of the island. Mismatch distribution of samples together with rejection of neutrality, the results of Bayesian Skyline Plots analysis, and a star-like network suggests a recent demographic expansion for the *P. l. lineata* lineage into the eastern lowlands, while the highland (*P. l. elanthana*) and northern clades (*P. dorsivittata* and *P. l. punctulata*) show signatures of rather stable populations. A major genetic discontinuity observed coincided with a northern lowland stretch that separates mid-altitude rainforests in the north from those in the center and south. Our analysis points to numerous unsolved taxonomic problems in this group of geckos, especially in the small-sized taxa (*P. hoeschi, P. kely, P. pusilla*), and provides a basis for a future comprehensive taxonomic revision, which will require integrative analysis of molecular, morphological and chromatic data as well as careful examination of type specimens.

Keywords Squamata, Gekkonidae · *Phelsuma lineata* species group · Biogeography · Multigene mtDNA phylogeny · *16S* rRNA · *RAG-1*

Introduction

Within the last two decades, the tropical island of Madagascar has increasingly moved into the focus of evolutionary studies of the mechanisms that shaped the present patterns of species richness and endemism within this biodiversity hotspot. In this context, various competing hypotheses have been proposed to explain the underlying speciation patterns and the origin of the extraordinary species diversity of Madagascar's unique biota (e.g., Raxworthy and Nussbaum 1995; Wilmé et al. 2006; Boumans et al. 2007; Wollenberg et al. 2008; reviewed by Vences et al. 2009), but whether any of these is prevalent remains uncertain, mainly because of missing taxonomic, phylogenetic, and distributional information (Vences et al. 2009).

Electronic supplementary material The online version of this article (doi:10.1007/s13127-013-0143-5) contains supplementary material, which is available to authorized users.

P.-S. Gehring · M. Gehara · F. M. Ratsoavina · M. Vences (✉)
Division of Evolutionary Biology, Zoological Institute, Technical University of Braunschweig, Mendelssohnstr. 4, 38106 Braunschweig, Germany
e-mail: m.vences@tu-bs.de

F. Glaw
Zoologische Staatssammlung München, Münchhausenstr. 21, 81247 München, Germany

F. M. Ratsoavina
Département de Biologie Animale, Université d'Antananarivo, BP 906, Antananarivo 101, Madagascar

Moreover, several recent studies on Malagasy vertebrates have revealed high levels of genetic differentiation within nominal species (e.g., Boumans et al. 2007; Olivieri et al. 2007; Ratsoavina et al. 2011; Gehring et al. 2011; Lemme et al. 2013) and uncovered numerous new cryptic species (e.g., Vieites et al. 2009; Nagy et al. 2012; Gehring et al. 2012).

Once species inventories are nearing completion and phylogenetic relationships among species have been largely resolved, phylogeographic methods can be employed to test hypotheses on the historical and spatial components of population lineage differentiation. Phylogeography is also important to studies of speciation, because the spatial distribution of alleles might help identify historical barriers to gene flow that could also have influenced earlier splits that led to completion of speciation. Explicit statistical methodology and the incorporation of population genetic analyses have transformed phylogeography into a powerful field of research (e.g., Knowles 2009).

One diversification mechanism discussed but rarely tested for Madagascar is the "northern species-pump" hypothesis, which implies that lineages originate in the northern part of the island and subsequently disperse southwards, causing a primary north–south split between sister species, or between phylogenetic or phylogeographic lineages. Such a north–south split was observed in *Microcebus* mouse lemurs (Yoder and Heckman 2006) and also at various phylogeographic levels in widespread reptile species or species complexes (Boumans et al. 2007).

Along Madagascar's east coast the elevation increases rapidly from the sea westwards toward the central highlands (defined as the zone starting between 800 and 1,000 m a.s.l.), which extends roughly until the northern third of the island but is not continuously connected with the northern mountain massifs. The cause for this separation is a roughly 100-km-long and 50-km-wide stretch of largely low elevations (below 800 m a.s.l.) from Mananara (16°10′12.06″S, 49°46′16.29″E) at the east coast to Mandritsara (15°50′36.10″S, 48°49′35.34″E) in the west (graphical representation is given in Fig. 1). According to the species-pump hypothesis, the northern massifs (Tsaratanana, Manongarivo, Marojejy, and Montagne d'Ambre) could have acted as refugia for humid forests during periods of drier climate (Raxworthy and Nussbaum 1995) when the interspersed lowlands probably became arid grasslands as they partly are today (Bond et al. 2008). Because the steep slopes of these mountains start near sea level, with peaks of 1,400–2,900 m, they appear to be suitable for allopatric speciation of populations that historically remained isolated in such montane refugia (Wollenberg et al. 2008), and they would have also offered opportunities for adaptive speciation along elevational bands. Indications for such adaptive speciation has been observed in other parts of the world (e.g., Endler 1977; Moritz et al. 2000), but explicit studies in Madagascar are still lacking.

Many reptiles are characterized by a rather low individual mobility and/or high substrate specificity, which promotes genetically highly structured populations over short geographical distances, with the retention of signals of historical events that generated current species distributions (reviewed by Camargo et al. 2010). Geographical ranges of many species are small (e.g., Glaw et al. 2012), and this pattern is more distinctly expressed in lizards as compared to snakes (Anderson and Marcus 1992; Böhm et al. 2013).

Most suitable for testing phylogeographic hypotheses on Madagascar are species or species complexes that are widespread along either Madagascar's east or west coast and subdivided into various distinct genealogical lineages. The effect of geographical barriers on their phylogeographic structure can be assessed simultaneously in various such taxa, and cladistic biogeographical inferences can be based on various nodes. One such suitable group of reptile species is formed by the day geckos of the genus *Phelsuma*. With 42 currently recognized species and subspecies, these gekkonids represent the most diverse lizard genus of the western Indian Ocean island Madagascar. The genus most probably originated in Madagascar and subsequently dispersed toward various other archipelagos in the Indian Ocean where they partly underwent small radiations (Austin et al. 2004; Harmon et al. 2008; Rocha et al. 2009, 2010). Most species of *Phelsuma* are colorful diurnal geckos of great morphological homogeneity. Their coloration ranges from bright green, often with red spots and markings, to dull grey or brownish in a few species.

Several molecular studies (Boumans et al. 2007; Raxworthy et al. 2007; Rocha et al. 2009, 2010; Nagy et al. 2012) have indicated the existence of genetically deeply subdivided species complexes within *Phelsuma*. One such example is the *Phelsuma lineata* species group as defined in Rocha et al. (2010). This is a monophyletic group containing nine nominal species (*P. antanosy, P. comorensis, P. dorsivittata, P. kely, P. lineata, P. parva, P. pusilla, P. quadriocellata,* and, as confirmed herein, *P. hoeschi;* information on distribution and phenotypes are given in Fig. 1) of which several include genetically divergent mitochondrial lineages in geographically distant populations (see Boumans et al. 2007; Rocha et al. 2009, 2010; Nagy et al. 2012). In previous studies the taxon sampling for this clade was rather incomplete and missed several important populations, such as the subspecies *P. lineata bombetokensis* (northwest coast), *P. lineata punctulata* (central-northern mountain massifs), *P. pusilla hallmanni* (central-eastern mid-elevation), the nominal species *P. hoeschi* (central-east coast), and a hitherto unstudied species from the island Nosy Be (northwest coast).

According to preliminary molecular data (Rocha et al. 2009, 2010) two clades can be defined within this species group: (1) all *P. quadriocellata* subspecies (sensu Rocha et al. 2010), *P. parva*, and *P. antanosy* and (2) *P. lineata*, *P. dorsivittata, P. pusilla*, *P. kely*, and *P. comorensis*. Although nuclear genes did not support these subgroups, for the faster evolving mtDNA markers they were monophyletic (Rocha et al. 2010). For the

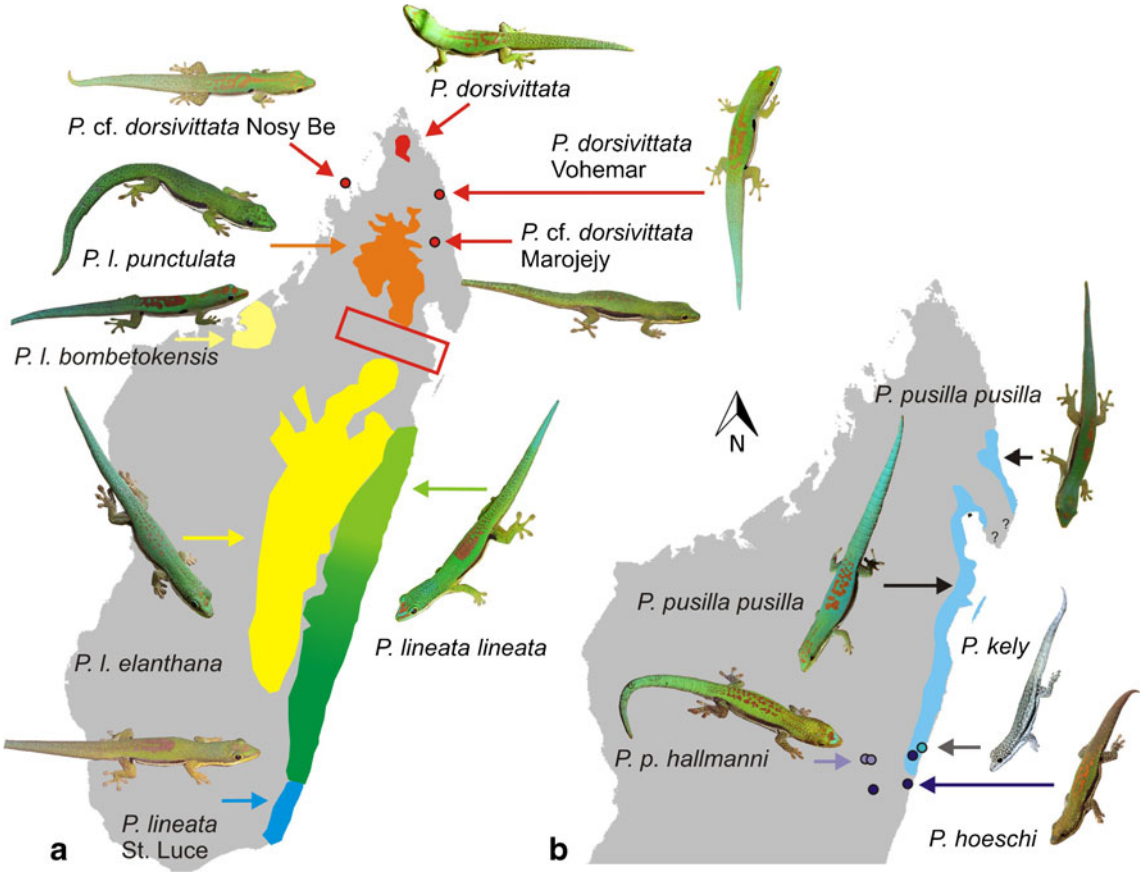

Fig. 1 Distribution map of current nominal taxa and major mtDNA lineages of the *Phelsuma lineata* subgroup. The information on distribution areas is based on Glaw and Vences (2007), Hallmann et al. (2008), Schönecker (2008), Gehring et al. (2010), and our own results. **a** Distribution of *P. dorsivittata* and *P. lineata* ssp. (*sensu lato*), with colors largely in accordance with major mtDNA clades; *red box* roughly highlights the position of the northern lowland stretch separating the central highlands from the northern mountain massifs; **b** The currently known distribution of the small-sized taxa *P. hoeschi*, *P. kely*, *P. pusilla pusilla*, and *P. p. hallmanni*

purpose of the present study, we focus on members of the second of these groups, which we name the *Phelsuma lineata* subgroup. Several of these taxa are rather widespread in both primary and secondary habitats, and a few of these have relatively narrow elevational ranges (Glaw and Vences 2007). This applies to *P. p. pusilla* and to some of the sublineages of *P. lineata*, which however, are rather poorly defined thus far. *Phelsuma kely, P. hoeschi*, and *P. pusilla hallmanni*, as presently defined, are microendemic taxa known only from very small ranges, and *Phelsuma lineata lineata* (sensu lato) presumably is a more generalized taxon and a comparatively good colonizer, suggested by its wide occurrence in both mid- and low-elevation sites along the east coast.

We here reconstruct the phylogeny and phylogeography of these geckos to clarify their evolutionary history and diversification, focusing especially on the large-sized species *P. dorsivittata* and *P. lineata*. We first infer a multigene mtDNA phylogeny for all taxa of the *Phelsuma lineata* subgroup to understand the relationships among the main phylogroups. Based on this tree and a dense population-level sampling for one mitochondrial and one nuclear gene, we reconstruct geographical clade origins and range expansions, and identify taxonomic problems to be clarified in future revisions. Given the geographical setting and the target species of this study, we use the data obtained to test the following hypotheses and predictions:

(1) If diversification of our target clade followed the species-pump mechanism, we expect the origin of the *P. lineata* subgroup to be in northern Madagascar. We also expect a general phylogenetic structure in which clades of successively more southern distribution split off the tree, with the southernmost clades in a nested position.
(2) Because of the apparent restriction of most lineages to limited elevational ranges, we expect that a major genetic discontinuity will be located at the lowland stretch separating northern Madagascar from the central highlands.
(3) If southern Madagascar was colonized recently, then we expect that one or several of the southern lineages will show signs of population expansion.

Material and methods

Terminology and taxonomic scope

To facilitate the presentation of results, we herein use a number of predefined terms to refer to units in our phylogenetic trees and assemblages of taxa. First, we follow previous authors, in particular Rocha et al. (2010), who defined monophyletic species groups among Malagasy day geckos. We refer to genealogical units at or below the species/subspecies level as "lineages" and use the term "clades" or "subclades" preferably for monophyletic groups above the species level according to current taxonomy.

One major problem that became obvious during data analysis was the apparent non-monophyly of the large-sized taxa in the *P. lineata* subgroup (i.e., of a putative group including *P. dorsivittata* plus *P. lineata*), with mtDNA lineages corresponding to several small-sized taxa scattered among the large-sized taxa in the phylogenetic trees. Because our sampling was insufficient for the small-sized taxa (*P. hoeschi, P. kely, P. pusilla*), we deliberately restricted our phylogeographic and demographic analyses to the large-sized taxa. As will be discussed below, we are aware that we thereby miss one important part of the very complex evolutionary history of these geckos, but nevertheless a number of valid inferences can be drawn from the partial data available.

Furthermore, in this work we avoid any taxonomic changes and name all lineages according to current taxonomy and identified using currently used diagnostic characters (Glaw and Vences 2007; Rocha et al. 2010). Given our results as presented below, it is obvious that in some taxa the phylogenetic data are in conflict with current taxonomy and that several nominal species of this group are poorly defined, suggesting the need for a comprehensive taxonomic revision.

Sample collection

Tissue samples were mainly collected by P.-S. Gehring, F. Ratsoavina, E. Rajeriarison, and F. Randrianasolo in April and May 2009 and 2010 along an approximately 1,000-km-long north–south transect along Madagascar's east coast. Additional tissue samples were obtained during fieldwork in Madagascar in the period from 2000–2010 by M. Vences, F. Glaw, S. Hauswaldt (ZCSH), and D. R. Vieites (DRV). FGMV, FGZC, and ZCMV refer to F. Glaw and M. Vences field numbers. PSG refers to field numbers of P.-S. Gehring.

Species identification in the field was mainly based on coloration patterns following Glaw and Vences (2007) and Hallmann et al. (2008). Tissue samples (tail clipping) were taken from geckos and preserved in 95–99 % ethanol. Nearly all animals were immediately released after sampling. Tissue samples were deposited at the Zoological Institute of the Technical University of Braunschweig (Germany). Some selected voucher specimens were euthanized, fixed in 90 % ethanol, and preserved in 70 % ethanol. Specimens are held at the University of Antananarivo, Département de Biologie Animale, Madagascar (UADBA-FGMV, UADBA), and the Zoologische Staatssammlung München, Germany (ZSM). Collecting localities were selected *a priori* with the aims of covering the whole distribution of *P. lineata* and *P. dorsivittata* at regular spatial intervals and of obtaining a representative sample size. Altogether we sampled a total of 376 individuals from ca. 50 localities. Localities were geo-referenced with GPS receivers to document distributions of samples most accurately (see Supplementary Materials Table S1). A georeferenced distribution map of samples including an elevation layer of sample localities in Madagascar was drawn using the program DIVA GIS (version 7.3.0.1.; http://diva-gis.org). Elevation layers were defined as follows: lowland 0–800 m; mid-elevation 800–1,200 highlands 1,200–1,500 m; high-altitude mountains >1,500 m a.s.l.

DNA extraction, amplification, and sequencing

Two molecular data sets were obtained for most of the samples and used for phylogeographic analyses: (1) a small but rather highly variable fragment of the mitochondrial 16S rRNA gene (*16S*) and (2) a fragment of the gene encoding the recombination activation protein (*RAG-1*), a single copy nuclear marker widely used in resolving relationships among vertebrate species (Chiari et al. 2009) and known to be variable among and within *Phelsuma* species (Rocha et al. 2009, 2010).

Total genomic DNA was extracted from the tissue samples using proteinase K digestion (10 mg/ml concentration) followed by a standard salt-extraction protocol (Bruford et al. 1992). A fragment of the mitochondrial 16S rRNA gene (*16S*) was initially amplified using the primers 16SAL and 16SBH (Palumbi et al. 1991). Species-specific primers (16S-Phel-L1 and 16S-Phel-H1) were then designed from the initially obtained sequences and used thereafter (Supplementary Materials Table S2). The aligned sequences had a length of 357 base pairs (bp) (for details, see extended methods in the Supplementary Materials). The *16S* rRNA nucleotide data set contained a hypervariable region between nucleotide position 104 and 148, which we chose to keep in the analysis to improve resolution among closely related lineages.

A fragment of ca. 550 base pairs (bp) of the nuclear *RAG-1* gene was initially amplified using the primers Urop-RAG1-F1 and Urop-RAG1-R1 (Bauer et al. 2011). We used these *RAG-1* sequences and those of other *Phelsuma* species to design new genus-specific primers for this gene (Supplementary Materials Table S2), which amplified a fragment with a final length of 451 bp (for details, see extended methods in the Supplementary Materials).

In addition, to resolve the phylogenetic relationships between major mtDNA clades observed in the *16S* data sets with high statistical support, fragments of three further mitochondrial genes

(*12S* rRNA; cytochrome oxidase subunit 1, *CO1;* and cytochrome b, *COB*) were sequenced for one or two selected representatives of each major *16S* clade (1,731 bp in total). PCR primers are listed in in Supplementary Materials Table S2.

All obtained PCR products were purified by exonuclease 1 and shrimp alkaline phosphatase digestion and largely sequenced by Macrogen Inc. (Seoul, S. Korea, http://www.macrogen.com). Additional samples were sequenced using BigDye v3.1 cycle sequencing chemistry and run on a 3130xl genetic analyzer (Applied Biosystems). Chromatograms were quality checked by eye, and sequences were edited, where necessary, using CodonCode Aligner (v. 3.7.1, Codon Code Corp.). Newly determined sequences were deposited in GenBank (accession nos. KC438417–KC439143; details in Supplementary Materials Table S1).

Haplotype networks and mtDNA phylogeny

The alignment of *16S* sequences was checked manually in MEGA version 5.05 (Tamura et al. 2011); basic analyses of nucleotide variation were performed in DnaSP v5.10 (Librado and Rozas 2009) and converted into the Roehlfile format as needed for the software Network (see below). In some intraspecific analyses, a hierarchical tree format may be inappropriate for representing relationships among haplotypes because the period of time over which the samples have evolved is so short that ancestral and descendant haplotypes exist concurrently (Posada and Crandall 2001; Kratysberg et al. 2004). In such instances, a haplotype network is more appropriate to illustrate relationships among the sampled haplotypes. We assessed the genetic structure through median-joining haplotype networks (Bandelt et al. 1999) using the program Network (version 4.6; http://fluxus-engineering.com), under the MP criterion (ε=10). Networks were manually redrawn and graphically processed in Corel DrawX3.

The *RAG-1* data set was aligned in MEGA and then separated into haplotypes using the Phase algorithm (Stephens et al. 2001; Stephens and Scheet 2005) implemented in the software DNASp (Librado and Rozas 2009), with three runs of different starting seed numbers with 1,000, 5,000, and 10,000 iterations. We used haplotype reconstructions from the run with the highest average goodness of fit to the underlying coalescent model (no recombination, 10,000 iterations, other parameters set at default values). No indels or stop codons were detected in the nuclear haplotypes. The *RAG-1* gene data sets were checked for recombination in Datamonkey (Delport et al. 2010; Pond and Frost 2005; http://www.datamonkey.org) with GARD (Pond et al. 2006) and in the recombination detection program RDP (Martin et al. 2010) where searches under RDP, Geneconv and Maxchi resulted in no detection of recombination (see Supplementary Materials Table S3). Median-joining haplotype networks were constructed using the previously phased data sets.

Phylogenetic analyses of representatives of the major mtDNA phylogroups were carried out using Bayesian inference (BI), maximum likelihood (ML), and maximum parsimony (MP). The complete mtDNA alignment consisted of 1,731 bp (*16S, 12S* rRNA, *CO1*, and *COB*). Substitution models for ML and BI were determined by AIC{Akaike 1974 #80} in jModelTest 0.1.1 (Posada 2008). The appropriate data partition scheme for BI analysis was determined by a comparison of Bayes factors of various alternative partition schemes according to the procedure outlined in Brandley et al. (2005). Partitioning schemes, selected models, and Bayes factor comparisons are provided in Supplementary Materials Tables S4–S6. Analyses were carried out in MrBayes 3.1.2. (Ronquist and Huelsenbeck 2003) with two runs of 5 million generations (started on random trees) and four chains, sampling the Markov chains at intervals of 1,000 generations.

MP analyses were done in PAUP* 4.0b10 (Swofford 2002) and ML analyses in PhyML 3.0. (Guindon and Gascuel 2003; for settings of the ML and MP analyses, see extended results in Supplementary Materials). In our phylogenetic analyses we used a specimen of *Phelsuma guttata* (Nosy Mangabe, ZCMV 2172) as outgroup and included two specimens of *Phelsuma quadriocellata* from two different localities (Ranomafana, ZCMV 380; Sahatelaka, DRV 5751) as hierarchical outgroups.

Sequence diversity indices and population genetic structure

The haplotype (h_d), nucleotide (π), and average number of nucleotide differences (k) were estimated in DnaSP 5.10 for each taxon and for each of the two major mtDNA lineages found within *P. lineata lineata.* Analyses of molecular variance (AMOVA) were run in order to infer the hierarchical patterns of variation in *16S* and *RAG-1,* utilizing the program ARLEQUIN 3.5 (Excoffier and Lischer 2010). AMOVAs were performed using pairwise differences and 10,000 random permutations; other settings were left as defaults. We performed hierarchical AMOVAs at two different levels to quantify how much variation is partitioned: (1) within the six major phylogroups (*P. dorsivittata* from Montagne d'Ambre*, P. l. elanthana, P. l. punctulata, P. lineata lineata* (north and south), and *P. lineata* St. Luce) and (2) among these major phylogroups.

Demographic history

Evidence for population expansion was examined by three approaches. First, Tajima's D (Tajima 1989) and Fu's F_s (Fu 1997) were calculated for the *16S* data, separately for each major *Phelsuma* mtDNA clade, using ARLEQUIN. These statistics require the that the included samples belong to the same population, and we therefore preferably applied them to mtDNA clades and lineages for which geographic coherence and low molecular variation suggest this assumption might be fulfilled. Tajima's D was originally developed to check for selective

neutrality but is widely used to test for population bottlenecks or rapid range expansions that also cause departures from equilibrium (e.g., Althoff and Pellmyr 2002; Toju and Sota 2006). A population that has experienced expansion may result in a rejection of the null hypothesis of neutrality. Negative F_s values reflect an excess of rare polymorphisms in a population, which is consistent with either positive selection or an increase in population size. Positive F_s values indicate an excess of intermediate-frequency alleles in a population and can result from balancing selection, population bottlenecks, or the presence of population structure. We chose these tests because of their increased statistical power in detecting significant changes in population size when using larger sample sizes (Simonsen et al. 1995; Ramos-Onsins and Rozas 2002).

In addition, mismatch-distribution analyses (Rogers and Harpending 1992; Harpending 1994) were conducted to assess population expansion using ARLEQUIN. This analysis produces an estimate of fit between the observed and expected mismatch (sum of squares deviations, SSD). A population that has experienced a rapid expansion or bottleneck in the recent past shows a smooth wave-like mismatch distribution with a star-like genealogy (Rogers and Harpending 1992; Slatkin and Hudson 1991). Parameters for the demographic and spatial expansion model were estimated from the data assuming theta 1 (population size after expansion) as infinite. Goodness-of-fit of expansion was assessed by calculating the significance of the raggedness index. We assumed no recombination and used 1,000 simulations to estimate the probability of obtaining a raggedness index of less than the observed value. A non-significant raggedness index was interpreted as a relatively good fit of the data to a model of population expansion.

Finally, in order to estimate past population dynamics of several lineages, we performed a Bayesian skyline plot (SLP) analysis in the program BEAST 1.7.2 (Drummond et al. 2005, 2012) using *16S* sequences. Each main mitochondrial lineage was analyzed separately. The best fitting substitution model was estimated using the Akaike information criterion (AIC) (Akaike 1974) as implemented in the software jModelTest 0.1.1 (Posada 2008). The relaxed molecular clock model was rejected for all clades because in exploratory analysis the posterior distribution of standard deviation of the log-normal relaxed clock (ucld.stev) included zero. Hence, in our final analysis we used a strict molecular clock model with a stepwise Bayesian skyline plot model. The number of groups was left as default. We used the program Tracer 1.5 (Rambaut and Drummond 2009) to ascertain chain mixing and a sufficient effective sample size (ESS) for all parameters in the Markov chain Monte Carlo (MCMC) analyses.

Ancestral distribution of haplotypes and estimates of clade origin

To assess the geographic locations of ancestral *Phelsuma* nodes [hypothetically representing the most recent common ancestor (MRCA)], we used the program Phylomapper 1 (Lemmon and Lemmon 2008). We used this method to test *a priori* hypotheses regarding the ancestral distribution in northern Madagascar using the phylogenetic relationships among individuals and their corresponding localities. This program implements a model of migration in a continuous landscape and uses spatial data combined with a rooted genealogy to estimate the coordinates of ancestral nodes under a Likelihood framework. For each internal node, a confidence surface of geographic location is estimated by calculating the likelihood of location in a surface of points representing the area of interest (Lemmon and Lemmon 2008). Using this approach, we estimated confidence surfaces of the nodes specified in Fig. 5. Given that the program estimates surfaces even in the absence of phylogeographic association, we ran the program twice for each of the nodes, with the original data and shuffling the correlation between coordinates and terminal branches of the used genealogy. We then compared the estimated surfaces with shuffled data, which would represent the absence of phylogeographic correlation, in order to check for absence of phylogeographic correlation. We implemented geographic coordinates for 376 specimens from the *16S* network analysis, including all members of the *n* subgroups, plus three individuals of *P. quadriocellata* that were used as outgroup (see Supplementary Materials Table S1). We set a different dispersal class for each main clade based on our observation that some clades appeared more restricted to specific elevations and habitats than others. The number of optimization replicates was set to 1,000. A higher number of replicates did not substantially change the estimated dispersal values, as we ascertained by additional runs of the program. Other parameters were left as default. It is important to note that this method makes a number of assumptions, including a continuous landscape where migration is not restricted, representative sampling of individuals, and constant migration abilities throughout the history of the clade. The results obtained must therefore be cautiously seen as an approximation and interpreted taking other information into account. Furthermore, these and similar ancestral area reconstruction methods are influenced by the accurateness of phylogenetic reconstruction, correspondence of the reconstructed tree with the actual phylogeny, and completeness of geographic coverage of ingroup and outgroup populations. We attempted to avoid these biases by restricting the ancestral area reconstruction to well-supported clades for which a very comprehensive geographic coverage was available.

Results

Mitochondrial DNA phylogeny and sequence diversity

DNA sequences of 357 bp of the mitochondrial *16S* rRNA gene were newly obtained or downloaded from GenBank

from the works of Boumans et al. (2007) and Rocha et al. (2009, 2010) (Supplementary Materials Table S1), from a total of 376 specimens [2 *P. comorensis*, 19 *P. dorsivittata*, 1 *P.* cf. *dorsivittata* from Marojejy, 1 *P.* cf. *dorsivittata* from Nosy Be, 2 *P. dorsivittata* from Vohemar, 5 *P. hoeschi,* 1 *P. kely*, 258 *P. lineata lineata*, 44 *P. l. elanthana*, 18 *P. l. punctulata*, 18 *P. lineata* (St. Luce); 3 *P. pusilla pusilla* (*sensu lato*), 1 *P. p. hallmanni*, and 3 *P. quadriocellata quadriocellata*].

The *16S* data set used for network analyses incorporated 376 specimens with 56 unique haplotypes clustering in seven major haplotype clades and several unique species-specific haplotypes (Fig. 2). High values were found for mean haplotype diversity (h_d=0.8216; 81 segregating sites; 67 parsimony informative sites) and nucleotide diversity (π=0.037; k=10.074). Nucleotide diversity (π) was highest within *P. dorsivittata* (see Table 1).

Based on the variation encountered in *16S*, we selected one or two samples from each major phylogroup for phylogenetic analysis (Fig. 3). This resulted in a concatenated mtDNA data set containing 1,731 total characters, of which 456 were parsimony-informative (for details, see extended results in the Supplementary Materials). Several subclades were identified (Fig. 3), which corresponded to the main clusters in the network. These subclades showed a relatively clear geographic correlation.

All phylogenetic analyses (MP, ML, BI) agreed in recovering a monophyletic group (A) containing the northern *P. dorsivittata* and *P. l. punctulata* with high statistical support. This northern *dorsivittata/punctulata* clade was placed sister to a second (poorly supported) clade containing all other ingroup taxa in our phylogeny. Within this second clade, the following deeply differentiated clades and lineages can be further distinguished: Specimens from the northern coastal lowlands belonging to *P. pusilla pusilla* (B) are placed sister to *P. kely*, *P. hoeschi*, *P. p. hallmanni* and *P. lineata* (sensu lato), and *P. comorensis*; specimens of *P. kely*, *P. hoeschi,* and *P. p. hallmanni* form a monophyletic group (C) with *P. kely* being sister to *P. hoeschi* and *P. p. hallmanni*; (D) *P. lineata* populations from the highlands (Fierenana, Anjozorobe), which correspond to *P. lineata elanthana* according to Krüger (1996), *P. l. bombetokensis,* and the Comoran taxon *P. comorensis*. The northwestern subspecies *P. l. bombetokensis* is recovered sister to *P. l. elanthana* populations from the northwestern highlands (Anjozorobe, Ambohitantely); (E) populations from the southeastern coastal lowlands (St. Luce); (F) populations of *P. l. lineata,* which were further subdivided into two major haplotype lineages (i) from the northern central east (e.g., Mahavelona, Andasibe) and (ii) from the southeast (e.g., Mananjary).

The strong differentiation among these main groups was confirmed by AMOVA. The percent of variance attributable to differences between major subclades was 88.2 % and significantly greater than zero (P<0.01, 1,000 permutations), but only 11.8 % within subclades (Table 2).

In order to test the biogeographic hypotheses outlined above, we mainly concentrated our analyses on the thoroughly sampled lineages of *P. dorsivittata*, *P. lineata lineata*, *P. l. punctulata*, *P. l. elanthana,* and *P. lineata* (St. Luce). The reconstruction of past demographic events requires a sufficient sample size, which was not given in the remaining taxa of the group. Moreover, the three *P. pusilla pusilla* populations showed deep divergences to each other (4.7–7.5 % uncorrected *p*-distance for *16S*), suggesting that they require additional taxonomic study.

Phylogeography of the northern taxa *P. dorsivittata* and *P. l. punctulata*

Within clade A (Fig. 3), no haplotype sharing was found among populations assigned to either *P. dorsivittata* or *P. l. punctulata*. The uncorrected *p*-distance between both taxa ranges between 5.8 and 8 % in *16S*. We found four different haplotypes within the three sampled *P. dorsivittata* populations, which mostly showed deep divergences among each other (up to 5.4 % uncorrected *p*-distance in *16S;* for a detailed overview on the genetic diversity within taxa, see Table 1). Two lineages distributed in the Montagne d'Ambre mountain massif in northern Madagascar (possibly separated by an elevational gradient), one in the area of Vohemar in northeastern Madagascar (included in the Network analyses only; for a tree-based illustration, see Rocha et al. 2010) and a currently poorly understood form (according to its phenotype initially identified as *P.* cf. *parva*) from the island of Nosy Be in northwestern Madagascar, surprisingly deeply nested within *P. dorsivittata* (localities are given in Fig. 1). Preliminary data on morphological variation of these *P. dorsivittata* lineages indicate the evolutionary discreteness of some of them and suggest they need taxonomic revision. Deep relationships among these lineages, however, were not reliably and congruently resolved by the various analyses.

Two lineages and seven private *16S* haplotypes were recovered by our analyses within *P. l. punctulata*, one lineage represented by the population of the Tsaratanana massif, including two groups of haplotypes separated by elevation, and another separated lineage from the western Makira plateau (Figs. 2 and 3). The uncorrected *p*-distance between the Tsaratanana and Makira lineages reached up to 3.2 % in the *16S* gene fragment (for a detailed overview on the genetic diversity, see Table 1). In terms of geographical distribution, these two lineages are distributed north of the lowland gap at elevations above 800 m.

Phylogeography of *P. l. elanthana*, *P.l. lineata*, and *P. lineata* (Ste Luce)

These taxa correspond roughly to the clades/lineages D, E, and F. Clade D also contains the nominal taxa *P. comorensis*

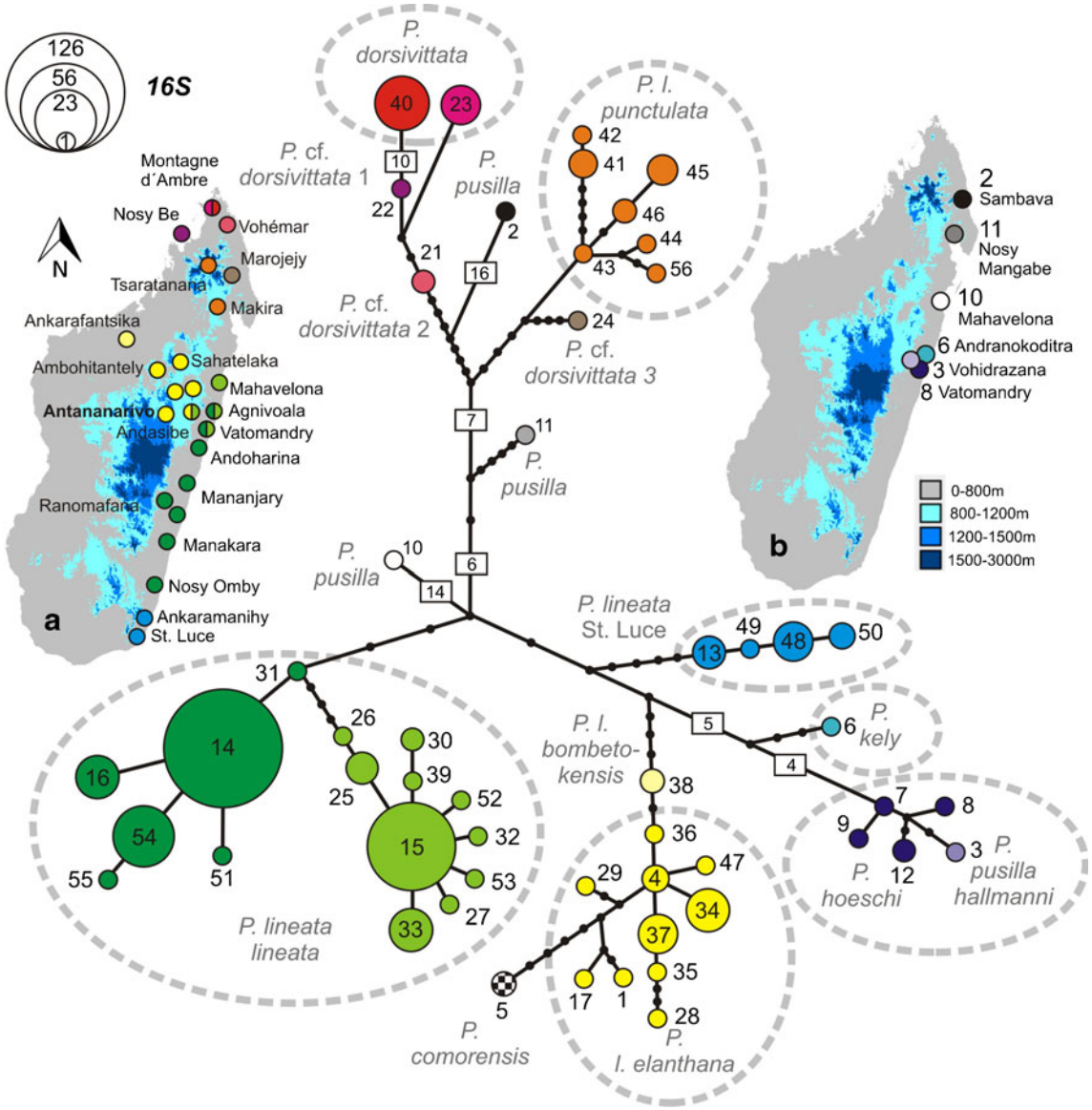

Fig. 2 Median-joining network of *16S* haplotypes in the *P. lineata* subgroup (consensus network of all the shortest trees). *Black dots* and *squares with numbers* represent median vectors and presumed unsampled or missing intermediates. Numbers denote haplotypes (for a detailed list assigning haplotypes to localities, see SI); colors are in accordance with major mtDNA clades as identified by phylogenetic analysis; size is proportional to their frequencies. Geographical location of the sampled populations and their haplotype composition are given in the maps. **a** Distribution of *P. dorsivittata* and *P. lineata* ssp. haplotypes; **b** distribution of *P. p. pusilla* haplotypes (*black, grey*, and *white*); *P. hoeschi* (*purple*), *P. kely* (*turquoise*), and *P. p. hallmanni* (*light purple*) haplotypes. Haplotypes 18–20 (not shown in the network) belong to specimens of *P. quadriocellata*, which were included in the analysis to verify absence of haplotype sharing or close relationships of this outgroup species with any of the ingroup haplotypes

and *P. l. bombetokensis*, which represent two independent colonization events of the Comoros and the dry forest in northwestern Madagascar, respectively. Both taxa possess private haplotypes, although their genetic discreteness is based only on a few mutational steps (Fig. 2).

Within *P. l. elanthana,* two lineages were recovered by our phylogenetic analyses. Compared to each other, they show shallow molecular divergences (1.6–2.0 % uncorrected *p*-distance; for a detailed overview on the genetic diversity, see Table 1). The arrangement of haplotypes of each lineage coincides roughly with the geographical position of the headwaters of the Mangoro River in central eastern Madagascar (Fig. 4). We found discrete haplotypes on either side of the Mangoro river (haplotypes 1, 17, and 34 west of Mangoro River; haplotypes 4, 28, 29, 35, 36, 37, 38, and 47 east of Mangoro River), and no haplotype sharing was found across the river.

Table 1 Molecular diversity in *16S* rRNA and *RAG-1* sequences for each *Phelsuma* (sub-) species and for major mtDNA lineages within *P. lineata lineata*

Species	mtDNA clade	*n*	*H*	*h* (SD)	π	*S*	*k*
16S molecular diversity							
P. dorsivittata (sensu lato)		28	5	0.582 (0.085)	0.022	23	7.003
P. lineata elanthana		44	16	0.902 (0.023)	0.013	27	4.216
P. l. lineata		266	21	0.661 (0.029)	0.014	27	3.772
P. l. lineata	North	77	14	0.719 (0.047)	0.006	24	1.919
P. l. lineata	South	186	8	0.359 (0.042)	0.002	15	0.715
P. lineata St. Luce		16	5	0.717 (0.095)	0.007	6	2.3
P. l. punctulata		16	7	0.717 (0.095)	0.007	6	2.3
RAG-1 molecular diversity							
P. dorsivittata (sensu lato)		14	2	0.143 (0.119)	0.00003	1	0.143
P. lineata elanthana		62	**8**	0.344 (0.075)	0.001	4	0.42
P. l. lineata		488	24	0.404 (0.028)	0.001	22	0.599
P. l. lineata	North	128	4	0.115 (0.04)	0.001	7	0.251
P. l. lineata	South	360	22	0.458 (0.032)	0.001	18	0.667
P. lineata St. Luce		10	4	0.733 (0.12)	0.01	13	4.467
P. l. punctulata		12	5	0.894 (0.054)	0.003	4	1.53

n, Number of studied samples, *H*, number of haplotypes; *h*, haplotype diversity; π, nucleotide diversity; *S*, number of segregating sites; *k*, average number of nucleotide differences, SD, standard deviation

In *P. lineata lineata* (lineage F), a taxon occurring in both mid-elevations and lowlands, we found two distinct sublineages (Fig. 3), a northern and a southern one, separated by a minimum of four and a maximum of ten mutational steps in the network (Fig. 2). For a detailed overview on the genetic diversity of each of these lineages, see Table 1. Both of them showed a star-like pattern in the network, which is characteristic of expanding populations (Clement et al. 2000). The northern lineage comprises ten haplotypes distributed from Mahavelona to Ambatolo makana and Marolambo along the Nosivolo and Mangoro Rivers. The southern lineage comprises six haplotypes distributed from Angivoala (central east coast) to Nosy Omby in the south. The geographical area where these two major lineages get in contact is bisected by the Mangoro River and represents therefore a potential barrier to dispersal and gene flow (Fig. 4). Nevertheless, the river obviously does not constitute an absolute barrier to migration and gene flow in *P. l. lineata* because several haplotypes of both lineages occur on either side of the river, but only in its near proximity.

Populations from the southeastern coastal lowlands (lineage E from St. Luce and Ankaramanihy) contained four distinct haplotypes and showed no haplotype sharing with any of the other taxa in the *16S* data set (Fig. 2; for a detailed overview on the genetic diversity, see Table 1). These southernmost populations of the *P. lineata* subgroup are separated from *P. l. lineata* and *P. l. elanthana* by a minimum of ten mutational steps each. Their taxonomic status requires additional study.

Phylogeography and differentiation of other taxa

Phelsuma kely, P. pusilla hallmanni, and *P. hoeschi* (clade C) were placed with highest statistical support sister to the clades/lineages D, E, and F. Within clade C, *P. kely* was recovered in a basal position sister to *P. hoeschi* and *P. p. hallmanni*, separated by a minimum of seven mutational steps from *P. hoeschi* (Fig. 2). However, *P. p. hallmanni* did not cluster with any of the other *P. p. pusilla* samples, as *a priori* expected. Instead, it was placed with samples of *P. hoeschi* (Fig. 3). Differentiation between these two taxa was weak (1.1–1.5 % uncorrected *p*-distance in *16S*). *16S* rRNA sequences were assigned to four haplotypes of *P. hoeschi* and one of *P. p. hallmanni*, being distinct by a maximum of three mutational steps, and no haplotype sharing was found between both taxa (Fig. 2). A detailed network of *16S* haplotypes of this clade is given in the Supplementary Materials Fig. S1, with the inclusion of a *P. p. hallmanni* sample from its type locality Andasibe (for a detailed overview on the genetic diversity, see in Supplementary Materials Table S7).

A weakly supported lineage in the phylogenetic analyses comprised the two *P. pusilla pusilla* samples from northeastern Madagascar (lineage B) and placed them in a basal position sister to the clade containing D, E, and F (Fig. 3). The median-joining network confirmed the high divergence within lineage B haplotypes, with a maximum of 51 mutational steps between haplotypes from Mahavelona and Sambava (Fig. 2) and *16S p*-distances of 4.7–7.5 % between Mahavelona, Nosy Mangabe, Antalaha, and Sambava.

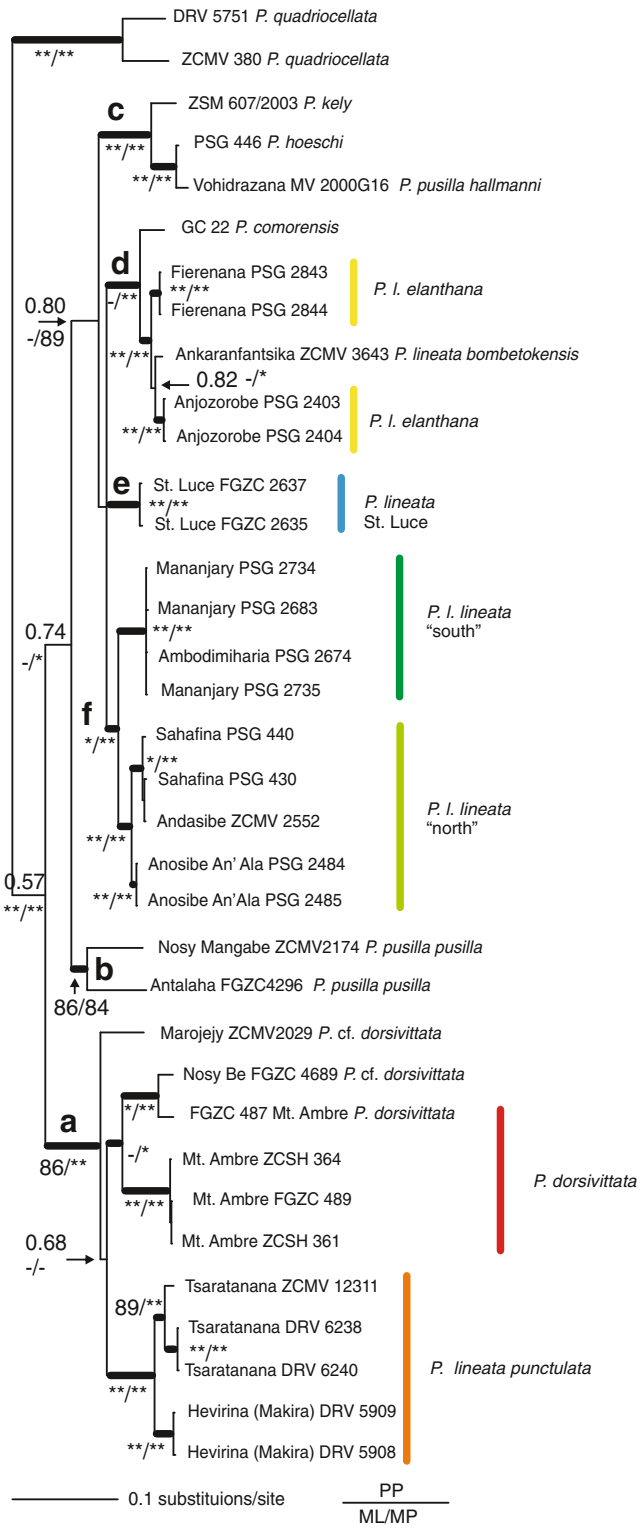

Fig. 3 Results of phylogenetic analyses of major mtDNA lineages within the *Phelsuma lineata* subgroup obtained by Bayesian inference under the best partitioning strategy, based on a total of 1,731 base pairs of four mitochondrial gene fragments. **a**–**e** Clades and lineages discussed in the text. *Bold lines* indicate Bayesian posterior probabilities (≥95), (*below*) bootstrap values of ML (first) and MP (second) analyses (** ≥95; * ≥90 %). Values below a threshold of 0.9 are given in total numbers. A similar topology was recovered in MP and ML analyses. Sequences of *Phelsuma guttata* were used as outgroup (removed for graphical reasons from tree)

Demographics and analyses of mtDNA clade origins

The Bayesian skyline plots (Fig. 5), which summarize estimates of effective population size, show clear signs for a rather recent population expansion of lineage F, while it recovered population stability for lineages D and E, and weak signs of population contraction for *P. dorsivittata* and *P. l. punctulata*. Considering an arbitrary mutation rate of around 1.5 % per site/million years, the population expansion of the *P. l. lineata* lineage would probably have started around 20,000 years ago. While mitochondrial mutation rates for geckos are unknown, this value suggests that this expansion certainly has been a rather young event that might have taken place in the Holocene or Pleistocene.

The statistical methods used to infer population demography were largely in agreement with these results. Fu's F_s and Tajima's D showed in lineage F, and to a lower degree also in *P. l. elanthana*, negative values indicating population expansions, which were significant in lineage F (Table 3). In the other clades, the positive F_s and D values indicated stable populations. The mismatch distribution analyses showed clear signs of population expansion only in lineage F (Table 3 and Supplementary Materials Fig. S2).

In conclusion, different demographic histories are suggested within the *P. lineata* subgroup. Mismatch distribution of samples and rejection of neutrality (Table 2), the results of the Bayesian skyline plots (Fig. 5), and the star-like phylogeny (Fig. 2) suggest a recent demographic expansion for *P. l. lineata* (lineage F) into the eastern lowlands, while the highland (lineage D; *P. l. elanthana*) and southern clades (lineage E; *P. lineata* St. Luce) show signs of rather stable populations. In the northern *P. l. punctulata* and *P. dorsivittata* population, weak signals of population contractions were evident.

Phylomapper (Lemmon and Lemmon 2008) suggested an origin for the basal clade A in northern Madagascar, with its center north of the lowland gap, in the general area of the Tsaratanana mountain massif (Fig. 5). All other areas of origin were plotted south of the lowland gap. Interestingly, for *P. l. elanthana* the results rather indicate an origin in the western highlands.

Nuclear DNA sequence diversity

DNA sequences of a fragment of 451 bp of the nuclear *RAG-1* gene were obtained from 306 specimens of the *P. lineata* subgroup. The *RAG-1* data set lacks samples of the taxa *P. l. bombetokensis*, *P. hoeschi*, *P. p. hallmanni*, and *P. kely,* of which no sequences were obtained because of PCR problems. *RAG-1*-GenBank sequences from Rocha et al. (2009) were not

Table 2 Analysis of molecular variance (AMOVA) examining the partitioning of genetic variation of one mitochondrial (*16S*) and one nuclear marker (*RAG-1*) among six major clades/lineages (as identified in phylogenetic analyses) and within these groups of the *Phelsuma lineata* subgroup

	16S	*RAG-1*
Among clades/lineages	88.24**	21.34**
Within clades/lineages	11.76**	78.66**
Statistical significance * $P<0.05$, ** $P<0.001$		

Covariance components: **highly significant ($p<0.01$); *significant ($p<0.05$)

included in the analysis since the fragment sequenced herein did not fully overlap with the published sequences.

Overall variation in *RAG-1* was low, and sequences contained no heterozygote sites. Only 32 different haplotypes out of the complete set of 296 samples of the *P. lineata* subgroup were identified (h_d=0.431; 34 segregating sites; 20 parsimony informative sites; Fig. 6). The data set for *P. dorsivittata* contained only the Montagne d'Ambre population, which turned out to be genetically rather homogeneous (two haplotypes; h_d=0.143) with only very shallow molecular diversity in the studied nuclear gene fragment and only one private haplotype (Fig. 6; Table 1). Most of the *P. dorsivittata* samples shared the same haplotype with *P. l. elanthana* and *P. l. lineata* (haplotype 3; Fig. 5). In our analyses, we found no haplotype sharing between the northern *P. dorsivittata* and *P. l. punctulata*. In the other northern lineage of day geckos (*P. l. punctulata*), we found five different haplotypes (containing one private haplotype), which were separated by a maximum of three mutational steps (for details on molecular diversity, see Table 1). In the central highland taxon *P. l. elanthana,* we found eight haplotypes (one private haplotype), which were shared with all other taxa, especially with *P. l. punctulata* and *P. l. lineata* (four shared haplotypes; for details on molecular diversity, see Table 1). The highest total number of haplotypes was found in *P. l. lineata* (lineage F; 24 haplotypes) with 19 private haplotypes separated by a maximum of eight mutational steps. Haplotype diversity (h_d) was second highest within this clade (for details on molecular diversity, see Table 1). Haplotype 3 was most common and distributed in almost all *P. l. lineata* populations. Some evidence for geographic structure was found, with 15 exclusive haplotypes restricted to populations

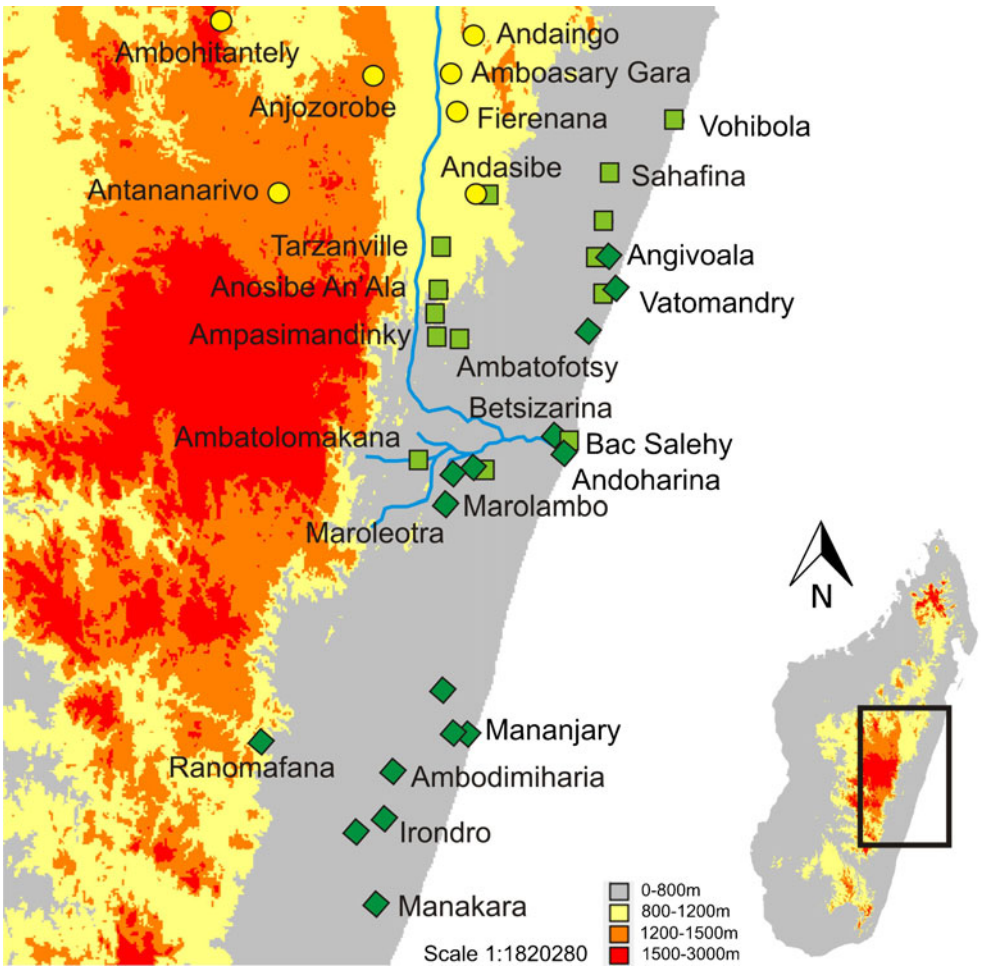

Fig. 4 Map of central eastern Madagascar showing the detailed distribution of haplotypes of *P. l. lineata* (*squares and diamonds*) and *P. l. elanthana* (*circles*) along the Mangoro and Nosivolo Rivers (*blue lines*). Colors are in accordance with major mtDNA clades as identified by phylogenetic analyses (*light green squares*: northern *P. l. lineata*; *dark green diamonds*: southern *P. l. lineata*)

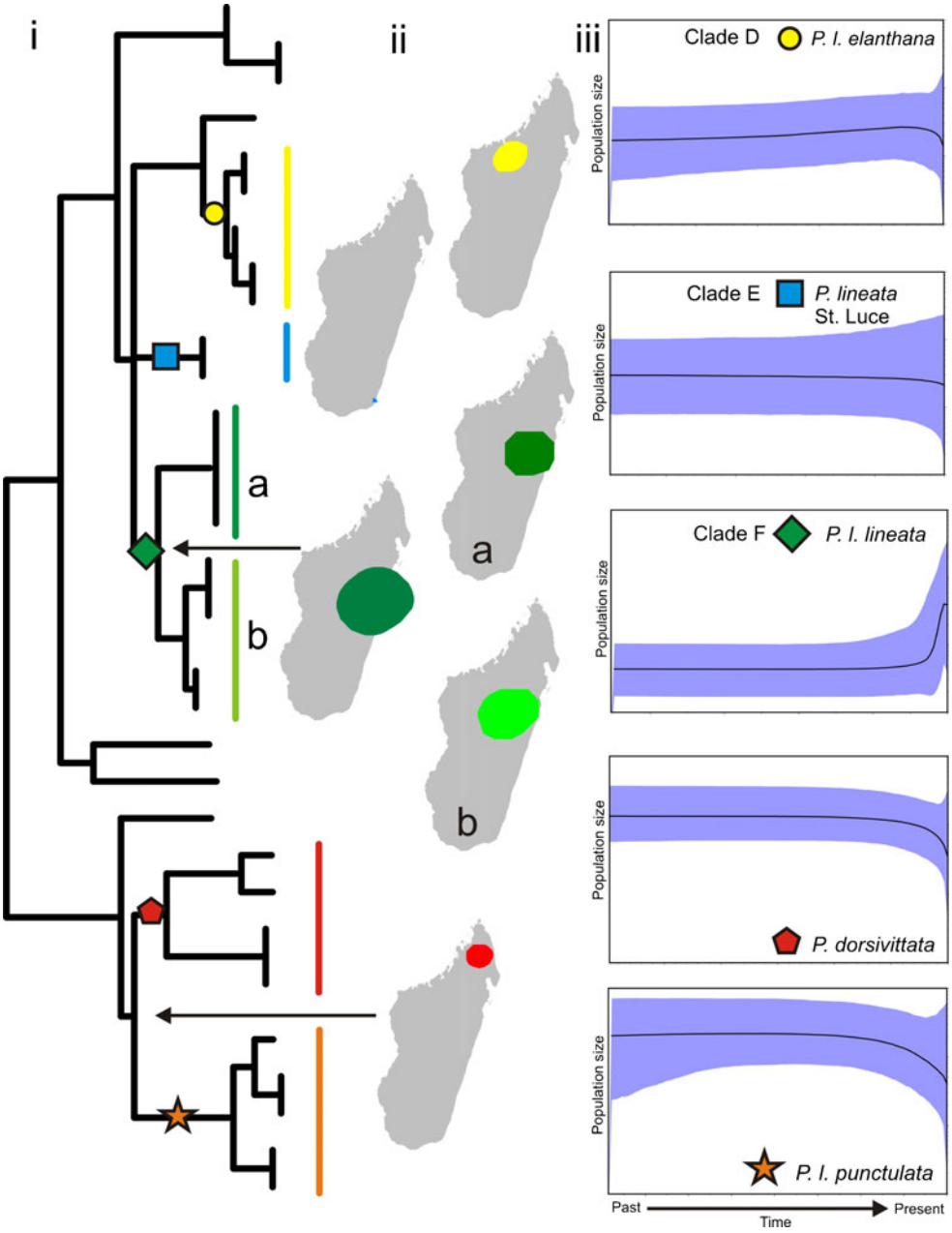

Fig. 5 Results of Phylomapper and Bayesian skyline plot (BSLP) analysis. (*i*) A simplified phylogram of the *Phelsuma lineata* subgroup; *symbols* indicate the nodes that were used for the BSLP analyses. (*ii*) Results of the Phylomapper analysis showing the 95 % confidence surface of the center of origin of different *P. lineata* subgroup clades calculated with one dispersal class and three dispersal classes (for details see text). (*iii*) Bayesian skyline plots for selected major mtDNA clades of the *Phelsuma lineata* subgroup. The *x*-axis is in units of years from the past to the present, and the *y*-axis is the estimated effective population size. The *thick solid line* is the median estimate of the demographic history; the *blue underlay* is the estimated median and 95 % HPD limits of the posterior distribution of the demographic history (see text for details)

Table 3 Results of demographic analyses of *Tajima's D* (Tajima 1989) and *Fu's* F_s (Fu 1997)

mtDNA lineage	F_s	*D*	*SSD*	r_{Harp}
P. dorsivittata	11.216	2.29	0.392**	0.457
P. lineata lineata (clade F)	0.233	0.633		
P. lineata lineata north	−2.717	−1.099	0.298**	0.40
P. lineata lineata south	−2.331	−1.621*	0.077	0.337
P. l. elanthana (clade D)	0.391	−0.328	0.053	0.142*
P. l. punctulata	2.105	0.514	0.044	0.056
P. lineata St. Luce (clade E)	2.643	1.93	0.167	0.521**

SSD, Sum of squared deviations between the observed mismatch distribution and the distribution expected under a sudden demographic expansion model; r_{Harp}, raggedness index of the mismatch distribution of major mtDNA lineages in the *Phelsuma lineata* subgroup as defined by Harpending (1994); statistical significance: *P<0.05; **P<0.01

along the estuary of the Mangoro River, but almost all of them separated by just one or two substitutions from the more common haplotypes (Fig. 6). In the southern populations of *P. l. lineata*, the analyses revealed only three exclusive haplotypes (haplotypes 32, 31, and 13 from Ranomafana and Mananjary). The highest nucleotide diversity was found in the southernmost lineage E comprising four haplotypes, of which three were restricted to this lineage (for details on molecular diversity, see Table 1).

Discussion

Biogeography of the *Phelsuma lineata* subgroup

Our phylogenetic reconstructions for the *P. lineata* subgroup are largely in agreement with previous studies (e.g., Boumans et al. 2007; Rocha et al. 2009, 2010) and complement some major taxon-sampling gaps within this species-rich clade of day geckos. Although a number of biogeographic inferences are possible from our results, not all of our main predictions could be satisfyingly tested.

In agreement with our first biogeographic prediction of a basal north–south split, the phylogeny recovered a basalmost position in the *P. lineata* subgroup of a northern clade (clade A) with two taxa largely restricted to northern mountain massifs (*P. dorsivittata* and *P. l. punctulata*) and another branch containing northeastern haplotypes here assigned to *P. pusilla* splitting off next (clade B). However, the respective nodes received only weak statistical support and therefore do not yet provide conclusive evidence for the north-south split. In general, it seems that deep lineages in the northernmost portion of Madagascar occupy smaller ranges than in the center and south.

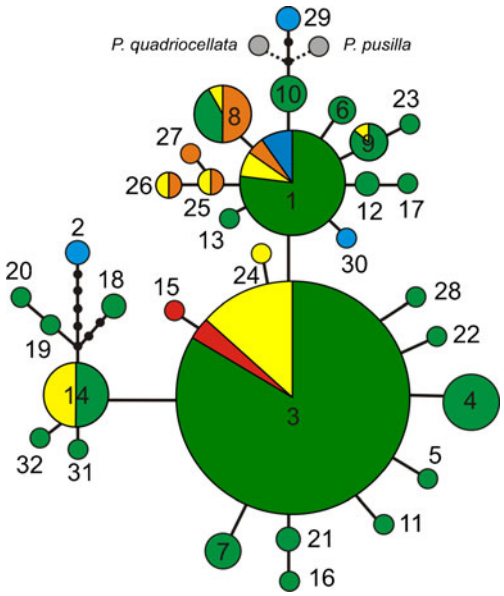

Fig. 6 Median-joining network of 32 *RAG-1* haplotypes of the *Phelsuma lineata* subgroup (*P. dorsivittata*, *red*; *P. l. lineata, green*; *P. l. elanthana, yellow*; *P. l. punctulata, orange*; *P. lineata* (St. Luce), *blue*). The consensus network of all the shortest trees is shown. *Numbers* denote unique haplotypes; *black dots* are median vectors or presumed unsampled or missing intermediates; size is proportional to their frequencies

Furthermore, our data agreed with our second prediction in supporting the existence of a major genetic discontinuity located at the northern lowland stretch, separating northern Madagascar from the central highlands. The rapid population expansion of *P. lineata lineata* into the central and southern lowland areas fulfills our third prediction. This event possibly coincided with the last glacial maximum and the beginning of the deglaciation of the Northern Hemisphere (Clark et al. 2009). No signature of range expansion was detected for the other lineages, but for several of these our sampling was too limited for meaningful demographic analyses.

Various studies have emphasized that northern Madagascar is characterized by a rich species diversity of reptiles and other vertebrates, including many regional endemics (e.g., Glaw et al. 2012; Vences et al. 2009; Wollenberg et al. 2008). The region is characterized by a high elevational heterogeneity (Wollenberg et al. 2008), with lowlands directly adjacent to several high mountain massifs in this area (e.g., Montagne d'Ambre, Anjanaharibe, and Marojejy). These may act as biogeographic refugia and centers of endemism (Vences et al. 2009). In the day geckos studied herein, we found at least four distinct northern lineages of *P. dorsivittata* (sensu lato) and two lineages of *P. l. punctulata*, most of them restricted to only one of the mountain massifs in the area. The *P.* cf. *dorsivittata* population from Marojejy represents the most basal lineage of clade A, and could also be considered as *P.* cf. *l. punctulata*. The molecular differentiation of this population is congruent with the assumed high level of endemism of Marojejy, which might have acted as a mountain refuge (Raxworthy and Nussbaum 1995; Wollenberg et al. 2008). For example, the existence of a flower (*Takhtajania perrieri)* in the adjacent Anjanaharibe mountain massif that otherwise is only known as a Cretaceous fossil supports the role of a possible stable rainforest refugium in this mountain area, since this only extant Africa/Madagascar representative of the family Winteraceae lacks water-conducting cells, which allow plants to withstand drought. Therefore, this mountain massif must have remained in moist conditions for millions of years (Schatz et al. 1998).

The microendemism of mtDNA haplotypes found in the northern mountains contrasts with the pattern in *P. l. lineata* (F), especially in the southern lineage of this taxon. Here we found a close correspondence between observed and predicted mismatch distributions, significantly negative values of Tajima's D test (Table 2), a star-like phylogeny (Fig. 2), and a clade origin in the northern portion of its distribution. This implies, in the recent past, a fast population expansion into a rather large area of southeastern Madagascar. Alternatively, such star-like phylogenies can also result from the selective sweep of an advantageous allele followed by accumulation of neutral variants (Maruyama

and Birky 1991). The origin of this lineage is reconstructed around the Mangoro River on the central east coast. The highest genetic diversity (in both mtDNA and nucDNA) was found in close proximity to the Mangoro River, which may have acted as a dispersal corridor from the central highlands into the eastern lowland regions (Wilmé et al. 2006). Localities east of the Mangoro (e.g., Andaingo, Amboasary Gara, Fierenana, Andasibe) are located at mid-elevations above 800 m a.s.l. and below 1,200 m a.s.l., while the western localities (Anjozorobe, Ambohitantely and Antananarivo) are located in the central highland at elevations above 1,200 m a.s.l. Therefore, these localities also differ considerably in their elevational level. Whether it is primarily the river barrier that is limiting gene flow or population differentiation might be influenced by elevation remains unsolved without fine-scale sampling closer to the river where, however, natural vegetation has been almost completely destroyed.

Higher endemism was found in the small-sized lowland taxa (e.g., *P. hoeschi*, *P. kely*, probably to some degree also in the different *P. pusilla* lineages) restricted to small geographic areas along the east coast. In *P. pusilla*, we found at least three independent colonization events into Madagascar's northeastern coastal lowlands.

One further biogeographic pattern worth mentioning is the occurrence of *P. l. bombetokensis* in northwestern Madagascar and its very close genetic relationships with *P. l. elanthana* from the central highlands. The distribution of the *elanthana-bombetokensis* lineage almost perfectly coincides with the river systems of the eastern Mangoro and the western Betsiboka, which together form a retreat-dispersion watershed after Wilmé et al. (2006).

The placement of *P. comorensis* in the clade occurring in the northwest of Madagascar (clade D in Fig. 3) agrees with the pattern observed in the iguanid lizard *Oplurus cuvieri comorensis,* which, like *P. comorensis*, is endemic to a small area in the north of the island Grand Comoro (Hawlitschek et al. 2011). In both cases, the sister lineage of the Comoroan taxon occurs in northwestern Madagascar (Münchenberg et al. 2008).

Prospects for future taxonomical research

Clearly, our results indicate that a systematic revision of the *P. lineata* group is necessary. Such a revision will require intensive work because various important questions remain unresolved: (1) *Phelsuma lineata*, being paraphyletic relative to *P. kely*, *P. comorensis,* and *P. pusilla*, exhibits consistent deep molecular differentiation (Boumans et al. 2007; Rocha et al. 2009, 2010; Nagy et al. 2012; this study). A taxonomic consequence could be to accept a paraphyletic species *P. lineata* or, given the known color differences between populations (Krüger 1996), its partitioning into several species. In this study we refrain from any taxonomic conclusions, nevertheless our results suggest that the taxonomy of these day geckos is highly complex and requires careful revisionary work.

Besides the unclarified status of the large-sized taxa *lineata, elanthana*, and *punctulata*, an even more complex situation is observed in the small-bodied species *P. pusilla, P. hoeschi,* and *P. kely*. Especially in *P. pusilla* concordance between morphological variation (Berghof and Trautmann 2009; Hallmann et al. 2008; Meier 1989) and molecular diversity has not yet been thoroughly assessed. The mitochondrial phylogeny suggests at least three independent small-bodied clades specialized to Madagascar's northeastern coastal lowlands. In its current taxonomic position, *P. pusilla* is paraphyletic given that *P. p. hallmanni* is deeply nested within the *P. kely/hoeschi* clade. Further studies should incorporate a large geographic sampling of various *P. pusilla* populations from the entire distribution range. Furthermore, the relationships between the nominal taxa *P. hoeschi* and *P. p. hallmanni* need to be studied in detail. Considering the weak genetic differentiation between both taxa, we cannot exclude that these two taxa might be conspecific. A resolution of this taxonomic conundrum will require inclusion of nuclear gene sequences, preferably of faster evolving genes, and a thorough comparison of type specimens of all small-sized taxa of the *P. lineata* subgroup.

Acknowledgements We are grateful to a large number of friends, colleagues, and students who assisted during field work, in particular to Jason L. Brown, Jörn Köhler, James and Carol Patton, Emile Rajeriarison, Theo Rajoafiarison, Roger-Daniel Randrianiaina, and David R. Vieites. Hans-Peter Berghof, Angelica Crottini, Arne Hartig, Susanne Hauswaldt, Thomas Hofmann, Alexandra Lima, Maciej Pabijan, and Patrick Schönecker provided pictures and additional samples. Gabi Keunecke, Meike Kondermann, and Eva Saxinger provided crucial help in the laboratory. This work was carried under a collaboration accord between TU Braunschweig and the Université d'Antananarivo, Département de Biologie Animale. We are grateful to the Malagasy authorities for research and export permits. Financial support was provided by the Volkswagen Foundation and by the Deutsche Gesellschaft für Herpetologie und Terrarienkunde (DGHT) to PSG, FG, and MV.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716–723.

Althoff, D. M., & Pellmyr, O. (2002). Examining genetic structure in a bogus yucca moth: a sequential approach to phylogeography. *Evolution, 56*, 1632–1643.

Anderson, S., & Marcus, L. F. (1992). Aerography of Australian tetrapods. *Australian Journal of Zoology, 40*, 627–651.

Austin, J. J., Arnold, E. N., & Jones, C. G. (2004). Reconstructing an island radiation using ancient and recent DNA: the extinct and living day geckos (*Phelsuma*) of the Mascarene islands. *Molecular Phylogenetics and Evolution, 31*, 102–109.

Bandelt, H.-J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution, 16*, 37–48.

Bauer, A., Glaw, F., Gehring, P.-S., & Vences, M. (2011). New species of *Blaesodactylus* (Squamata: Gekkonidae) from Ankarafantsika National Park in north-western Madagascar. *Zootaxa, 2942*, 57–68.

Berghof, H. P., & Trautmann, G. (2009). Eine neue Art der Gattung *Phelsuma* Gray, 1825 (Sauria: Gekkonidae) von der Ostküste Madagaskars. *Sauria, 31*, 5–14.

Böhm, M., et al. (2013). The conservation status of the world's reptiles. *Biological Conservation, 157*, 372–385 (261 authors).

Bond, W., Silander, J. A., Jr., Ranaivonasy, J., & Ratsirarson, J. (2008). The antiquity of Madagascar's grasslands and the rise of C4 grassy biomes. *Journal of Biogeography, 35*, 1743–1758.

Boumans, L., Vieites, D. R., Glaw, F., & Vences, M. (2007). Geographical patterns of deep mitochondrial differentiation in widespread Malagasy reptiles. *Molecular Phylogenetics and Evolution, 45*, 822–839.

Brandley, M. C., Schmitz, A., & Reeder, T. W. (2005). Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. *Systematic Biology, 54*, 373–390.

Bruford, M. W., Hanotte, O., Brookfield, J. F. Y., & Burke, T. (1992). Singlelocus and multilocus DNA fingerprint. In A. R. Hoelzel (Ed.), *Molecular genetic analysis of populations: A practical approach* (pp. 225–270). Oxford: IRL Press.

Camargo, A., Sinervo, B., & Sites, J. W., Jr. (2010). Lizards as model organisms for linking phylogeographic and speciation studies. *Molecular Ecology, 19*, 3250–3270.

Chiari, Y., van der Meijden, A., Madsen, O., Vences, M., & Meyer, A. (2009). Base composition, selection, and phylogenetic significance of indels in the recombination activating gene-1 in vertebrates. *Frontiers in Zoology*, *6*, article 32.

Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., et al. (2009). The last glacial maximum. *Science, 325*, 710–714.

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology, 9*, 1657–1659.

Delport, W., Poon, A. F., Frost, S. D. W., & Pond, S. L. K. (2010). Pond Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics, 26*, 2455–2457.

Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution, 22*, 1185–1192.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution, 29*, 1969–1973.

Endler, J. A. (1977). *Geographic variation, speciation, and clines*. Princeton University Press.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources, 10*, 564–567.

Fu, Y. X. (1997). Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics, 147*, 915–925.

Gehring, P.-S., Crottini, A., Glaw, F., Hauswaldt, S., & Ratsoavina, F. M. (2010). Notes on the natural history, distribution and malformations of day geckos (*Phelsuma*) from Madagascar. *Herpetology Notes, 3*, 321–327.

Gehring, P.-S., Ratsoavina, F. M., Vences, M., & Glaw, F. (2011). *Calumma vohibola*, a new chameleon species (Squamata: Chamaeleonidae) from the littoral forests of eastern Madagascar. *African Journal of Herpetology, 60*, 130–154.

Gehring, P.-S., Tolley, K., Eckhardt, F. S., Townsend, T. M., Ziegler, T., Ratsoavina, F. M., et al. (2012). Hiding deep in the trees: discovery of divergent mitochondrial lineages in Malagasy chameleons of the *Calumma nasutum* group. *Ecology and Evolution, 2*, 1468–1479.

Glaw, F., & Vences, M. (2007). *A field guide to the amphibians and reptiles of Madagascar* (3rd ed.). Köln (Cologne): Vences & Glaw Verlag. 496 pp.

Glaw, F., Köhler, J., Townsend, T. M., & Vences, M. (2012). Rivaling the world's smallest reptiles: Discovery of miniaturized and microendemic new species of leaf chameleons (*Brookesia*) from northern Madagascar. *PLoS ONE, 7*, doi:10.1371/journal.pone.0031314.

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology, 52*, 696–704.

Hallmann, G., Krüger, J., & Trautmann, G. (Eds.). (2008). *Faszinierende Taggeckos. Die Gattung Phelsuma (2. Auflage)*. Münster: Natur und Tier-Verlag. 253 pp.

Harmon, L. J., Melville, J., Larson, A., & Losos, J. B. (2008). The role of geography and ecological opportunity in the diversification of Day Geckos (*Phelsuma*). *Systematic Biology, 57*, 562–573.

Harpending, H. C. (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology, 66*, 591–600.

Hawlitschek, O., Brückmann, B., Berger, J., Green, K., & Glaw, F. (2011). Integrating field surveys and remote sensing data to study distribution, habitat use, and conservation status of the herpetofauna of the Comoro Islands. *Zookeys, 144*, 21–79.

Knowles, L. L. (2009). Statistical phylogeography. *Annual Review of Ecology, Evolution, and Systematics, 40*, 593–612.

Kratysberg, Y., Schwartz, M., & Brown, T. A. (2004). Recombination of human mitochondrial DNA. *Science, 304*, 981.

Krüger, J. (1996). Zur Nomenklatur der *lineata*-Gruppe in der Gattung *Phelsuma* (Reptilia: Sauria: Gekkonidae). *Sauria, 18*, 37–42.

Lemme, I., Erbacher, M., Kaffenberger, N., Vences, M., & Köhler, J. (2013). Molecules and morphology suggest cryptic species diversity and an overall complex taxonomy of fish scale geckos, genus *Geckolepis*. *Organisms, Diversity and Evolution, 13*, 87–95.

Lemmon, A., & Lemmon, E. (2008). A likelihood framework for estimating phylogeographic history on a continuous landscape. *Systematic Biology, 57*, 544–561.

Librado, P., & Rozas, J. (2009). DNASP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics, 25*, 1451–1452.

Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D., & Lefeuvre, P. (2010). RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics, 26*, 2462–2463.

Maruyama, T., & Birky, C. W., Jr. (1991). Effects of periodic selection on gene diversity in organelle genomes and other systems without recombination. *Genetics, 127*, 449–451.

Meier, H. (1989). Eine neue Form aus der *lineata*-Gruppe der Gattung *Phelsuma* auf Madagaskar. *Salamandra, 25*, 230–236.

Moritz, C., Patton, J. L., Schneider, C. J., & Smith, T. B. (2000). Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics, 31*, 533–563.

Münchenberg, T., Wollenberg, K. C., Glaw, F., & Vences, M. (2008). Molecular phylogeny and geographic variation of Malagasy iguanas (*Oplurus* and *Chalarodon*). *Amphibia-Reptilia, 29*, 319–327.

Nagy, Z. T., Sonet, G., Glaw, F., & Vences, M. (2012). First large-scale DNA barcoding assessment of reptiles in the biodiversity hotspot of Madagascar, based on newly designed COI primers. *PLoS ONE, 7*, doi:10.1371/journal.pone.0034506

Olivieri, G., Zimmermann, E., Randrianambinina, B., Rasoloharijaona, S., Rakotondavony, D., Guschanski, K., et al. (2007). The ever-increasing diversity in mouse lemurs: three new species in north and northwestern Madagascar. *Molecular Phylogenetics and Evolution, 43*, 309–327.

Palumbi, S., Martin, A., Ramano, S., McMillan, W. O., Stice, L., & Grabowski, G. (1991). *The simple fool's guide to PCR, version 2*. Honolulu, Hawaii: University of Hawaii Zoology Department. 46 pp.

Pond, S. L. K., & Frost, S. D. W. (2005). Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics, 21*, 2531–2533.

Pond, S. L. K., Posada, D., Gravenor, M. B., Woelk, C. H., & Frost, S. D. W. (2006). Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology and Evolution, 23*, 1891–1901.

Posada, D. (2008). JModelTest: phylogenetic model averaging. *Molecular Biology and Evolution, 25*, 1253–1256.

Posada, D., & Crandall, K. A. (2001). Selecting the best-fit model of nucleotide substitution. *Systematic Biology, 50*, 580–601.

Rambaut, A., & Drummond, A. J. (2009). Tracer v1.5 Available at http://beast.bio.ed.ac.uk/Tracer

Ramos-Onsins, R., & Rozas, R. (2002). Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution, 19*, 2092–2100.

Ratsoavina, F. M., Louis, E. E., Jr., Crottini, A., Randrianiaina, R. D., Glaw, F., & Vences, M. (2011). A new leaf tailed gecko species from northern Madagascar with a preliminary assessment of molecular and morphological variability in the *Uroplatus ebenaui* group. *Zootaxa, 3022*, 39–57.

Raxworthy, C. J., & Nussbaum, R. A. (1995). Systematics, speciation and biogeography of the dwarf chameleons (*Brookesia*, Reptilia, Squamata, Chamaeleontidae) of northern Madagascar. *Journal of Zoology, 235*, 525–558.

Raxworthy, C. J., Ingram, C. M., Rabibisoa, N., & Pearson, R. G. (2007). Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using Day Geckos (*Phelsuma*) from Madagascar. *Systematic Biology, 56*, 907–923.

Rocha, S., Vences, M., Glaw, F., Posada, D., & Harris, D. J. (2009). Multigene phylogeny of Malagasy day geckos of the genus *Phelsuma*. *Molecular Phylogenetics and Evolution, 52*, 530–537.

Rocha, S., Rösler, H., Gehring, P.-S., Glaw, F., Posada, D., Harris, D. J., et al. (2010). Phylogenetic systematics of day geckos, genus *Phelsuma*, based on molecular and morphological data (Squamata: Gekkonidae). *Zootaxa, 2429*, 1–28.

Rogers, A. R., & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution, 9*, 552–569.

Ronquist, F., & Huelsenbeck, J. P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics, 19*, 1572–1574.

Schatz, G. E., Lowry, P. P., II, & Ramisamihantanirina, A. (1998). *Takhtajania perrieri* rediscovered. *Nature, 391*, 133–134.

Schönecker, P. (2008). Terralog Vol. 12: geckos of Madagascar, the Seychelles, Comoros and Mascarene Islands – Geckos Madagaskars, der Seychellen, Komoren und Maskarenen. Edition Chimaira, Frankfurt am Main, 144 p.

Simonsen, K., Churchill, G., & Aquadro, C. (1995). Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics, 141*, 413–429.

Slatkin, M., & Hudson, R. R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics, 129*, 555–562.

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

Stephens, M., Smith, N. J., & Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics, 68*, 978–989.

Swofford, D. L. (2002). *PAUP*. Phylogenetic analysis using parsimony (*and Other Methods). Version 4.0b10*. Sunderland, Massachusetts: Sinauer Associates.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics, 123*, 585–595.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution, 28*, 2731–2739.

Toju, H., & Sota, T. (2006). Phylogeography and the geographic cline in the armament of a seed-predatory weevil: effects of historical events vs. natural selection from the host plant. *Molecular Ecology, 15*, 4161–4173.

Vences, M., Wollenberg, K. C., Vieites, D. R., & Lees, D. C. (2009). Madagascar as a model region of species diversification. *Trends in Ecology & Evolution, 24*, 456–465.

Vieites, D. R., Wollenberg, K. C., Andreone, F., Köhler, J., Glaw, F., & Vences, M. (2009). Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. *Proceedings of the National Academy of Sciences of the U.S.A., 106*, 8267–8272.

Wilmé, L., Goodman, S. M., & Ganzhorn, J. U. (2006). Biogeographic evolution of Madagascar's microendemic biota. *Science, 312*, 1063–1065.

Wollenberg, K. C., Vieites, D. R., van der Meijden, A., Glaw, F., Cannatella, D. C., & Vences, M. (2008). Patterns of endemism and species richness in Malagasy cophyline frogs support a key role of mountainous areas for speciation. *Evolution, 62*, 1890–1907.

Yoder, A. D., & Heckman, K. (2006). Mouse lemur phylogeography revises a model of ecogeographic constraint in Madagascar. In *Primate biogeography: Progress and prospects.* Fleagle, J., Lehman, S. M., (eds), pp. 255–268, Kluwer.