

On the diversity, colonization patterns and status of *Hemidactylus* spp. (Reptilia: Gekkonidae) from the Western Indian Ocean islands

Sara Rocha^{1,2,3}, Miguel A. Carretero^{1,2} & D. James Harris^{1,2}

¹CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Portugal

²Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Portugal

³Departamento de Bioquímica, Genética e Inmunología, Facultad de Biología, Universidad de Vigo, Spain

Hemidactylus geckos are probably the most widespread genus of reptiles, with a world-wide distribution and multiple cases of range expansion and transmarine colonization. With an almost cosmopolitan distribution and many species being morphologically similar it has proved difficult to delimit species diversity and distributions. Using a comprehensive analysis of individuals collected across the Western Indian Ocean islands and some locations along the East African coast, we further assess their diversity and the origin of insular populations. Despite four species of *Hemidactylus* being widespread across the Western Indian Ocean islands, most of their range in this area may actually be the result of very recent (possibly human-aided) dispersal events. Most probably, all *Hemidactylus* species occurring in the Comoros and granitic Seychelles archipelagoes are not native. Instances of natural colonization seem to be only the ones of *H. mabouia* to Madagascar and from there to the coralline archipelago of Aldabra. Surprisingly, Aldabra populations reveal a remarkable diversity and structure. Given the degree of divergence observed we propose that insular (Gulf of Guinea, Comoros, Madagascar, Seychelles) populations of *H. mabouia* are recognized as *H. mercatorius*. Cryptic variation is further uncovered in all species in their native range, with *H. platycephalus* and *H. mabouia* harbouring several highly divergent lineages, but further taxonomic assignments should await detailed assessments of distribution and molecular variation.

Key words: 16s rRNA, Comoros, geckos, Seychelles, Madagascar

INTRODUCTION

Hemidactylus geckos occur naturally through much of tropical Asia and Africa, the more arid areas of north-east Africa and southwest Asia, the Mediterranean region and South America, which they reached by natural transmarine colonization (Kluge, 1969). Both naturally and anthropogenically, they also reached various islands of the Americas, West Indies, Australia and the Indian, Pacific and Atlantic oceans, demonstrating more cases of range expansion through transmarine colonization than any other reptile genus (Carranza & Arnold, 2006).

Across the Western Indian Ocean islands at least four species belonging to three deeply differentiated clades (Carranza & Arnold, 2006) are known to occur (*H. mabouia/mercatorius*, *H. platycephalus*, *H. frenatus* and *H. brooki*). The origin and status of Malagasy and Comoran populations has been explored in recent years using molecular data (Vences et al., 2004; Rocha et al., 2005; Boumans et al., 2007). Summarizing, both *H. frenatus* and *H. brooki*, which belong to the tropical Asian clade (Carranza & Arnold, 2006), are present in Madagascar and the Comoros (*H. brooki* only in the Comoros) with patterns that better fit anthropogenic introduction scenarios. *Hemidactylus platycephalus* was

shown to be the most abundant species in the Comoros but the lack of variation observed across these islands (and Madagascar) pointed also to a very recent range expansion, either natural or human-aided. *Hemidactylus mabouia/mercatorius* exhibited the most complex variation pattern and data clearly indicate that the exact species delimitation and phylogeographic structure within this putative “species-complex” is far from being understood. Vences et al. (2004), based on the deep divergence observed between African (*H. mabouia*) and Malagasy individuals, used the *H. mercatorius* (Gray, 1842) designation for the Malagasy populations but results from Rocha et al. (2005) showed that species delimitation is unclear, given that *H. mabouia* haplotypes from the Gulf of Guinea populations cluster with one of the clades of both Malagasy and Comoran populations. Madagascar harbours two distinct clades of this “species” that were further explored by Boumans et al. (2007), in a more comprehensive analysis across its Malagasy distribution. These authors argued in favour of its native status (at least across most of its range) in Madagascar.

Here we use comprehensive sampling across the Western Indian Ocean islands of the Seychelles (both coralline and granitic archipelagos), Zanzibar and Pemba (off Tanzania), some locations across the East African

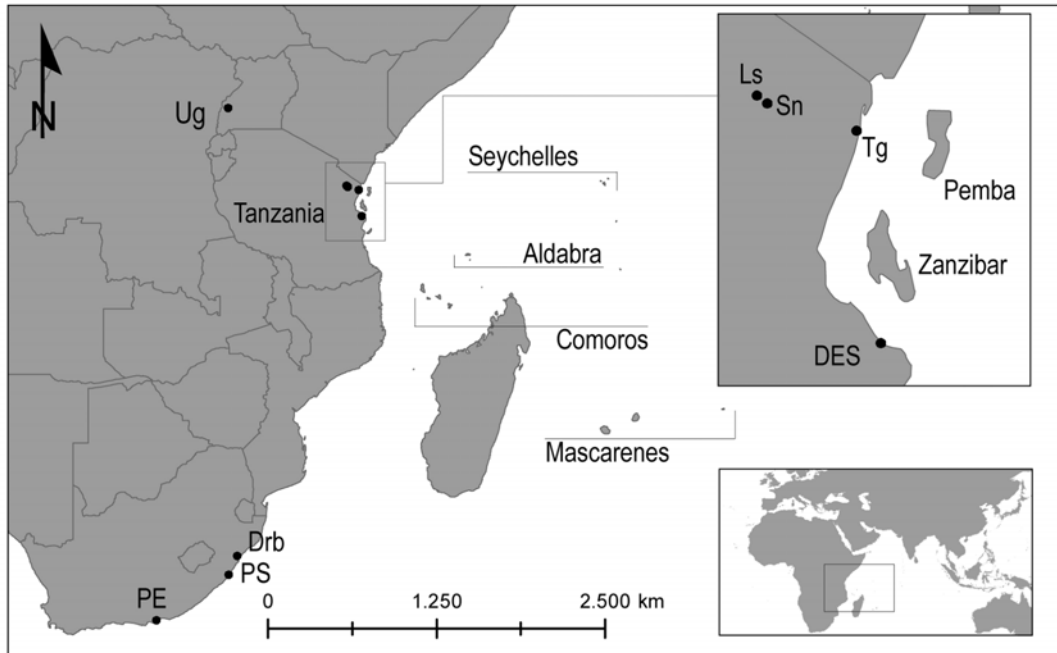


Fig. 1. Map of main sample localities (see Electronic Appendix for detailed information). New localities in continental Africa are highlighted (Ug: Kibale, Uganda; Ls, Sn, Tg, DES: Lushoto, Soni, Tanga and DarEsSalaam, Tanzania; Drb, PS, PE: Durban, Port Shepstone and Port Elizabeth, South Africa). For detailed locations of Comoran and Seychellois samples, see Carretero et al. (2005) and Rocha et al. (2010).

coast (Tanzania, South Africa) and Uganda, together with the already published sequences of these species, to further assess their diversity and the origin of these insular populations and to try to provide some insights into the origins and status of all insular populations of both *H. mabouia* and *H. platycephalus* by the inclusion of continental samples from Africa.

MATERIALS AND METHODS

We collected 130 new samples from the four target species across South Africa, Uganda, Tanzania mainland, Zanzibar and Pemba islands and the Seychelles coralline and granitic archipelagos (Fig. 1 and Electronic Appendix, downloadable at http://www.thebhs.org/pubs_journal_online_appendices.html).

Genetic samples were taken in the field by preserving muscle tissue (autotomized tails) in pure ethanol and total genomic DNA was extracted using standard high-salt (Sambrook et al., 1989) or phenol-chloroform (Taggart et al., 1992) protocols. We used primers 16SA-L and 16SB-H (Palumbi et al., 1991) to amplify a section of approximately 500 bp of the mitochondrial 16S ribosomal RNA gene, using conditions described in Harris et al. (1998). Sequences were produced by a commercial facility (Macrogen, Seoul, Korea).

New sequences were then combined with previously published ones from Vences et al. (2004), Rocha et al. (2005) and Boumans et al. (2007) and aligned in MAFFT v. 6 under the G-INS-I strategy (Kato et al., 2005), an iterative approach that assumes that the total region can be aligned and tries to align them globally. It is appropriate

for gene regions like these, where, despite indel-rich regions, global homology can be observed, and has been shown to perform better than most commonly used pairwise approaches (Wilm et al., 2006). Haplotype networks for all species were constructed using the median-joining method (Bandelt et al., 1999) with MP optimization (Polzin & Daneschmand, 2003) in NETWORK v. 4.510 (www.fluxus-engineering.com). A maximum likelihood (ML) tree was also built for all unique haplotypes of both *H. mabouia-mercatorius* and *H. platycephalus*. The best fitting model was found through jModeltest (Guindon & Gascuel, 2003; Posada, 2008) under the AICc criteria and an ML search with 1000 bootstrap replicates was performed in PhyML (Guindon & Gascuel, 2003).

The software package BEAST v. 1.4.8 (Drummond & Rambaut, 2007) was used to obtain Bayesian estimates (posterior distributions) of the substitution rates underlying the data, assuming the Aldabra group clade of *H. mabouia-mercatorius* to be approximately 125,000 years old (Thomson & Walton, 1972) (see Results and Discussion). For this we again found the best fitting model for the respective dataset (all Aldabra group individuals) using jModeltest and chose the one with the best AICc score (GTR+I) among those available in BEAST. Because we were dealing with a short intraspecific dataset (low information content and little reason to expect rate heterogeneity), we used a strict clock model with a coalescent tree prior and an exponential growth population model. We used a uniform prior on the tree root of 118,000 yr to 140,000 yr (according to Thomson & Walton, 1972) as well as a uniform prior on the mutation rate, which we allowed to vary from 0.005 substitutions per site per mil-

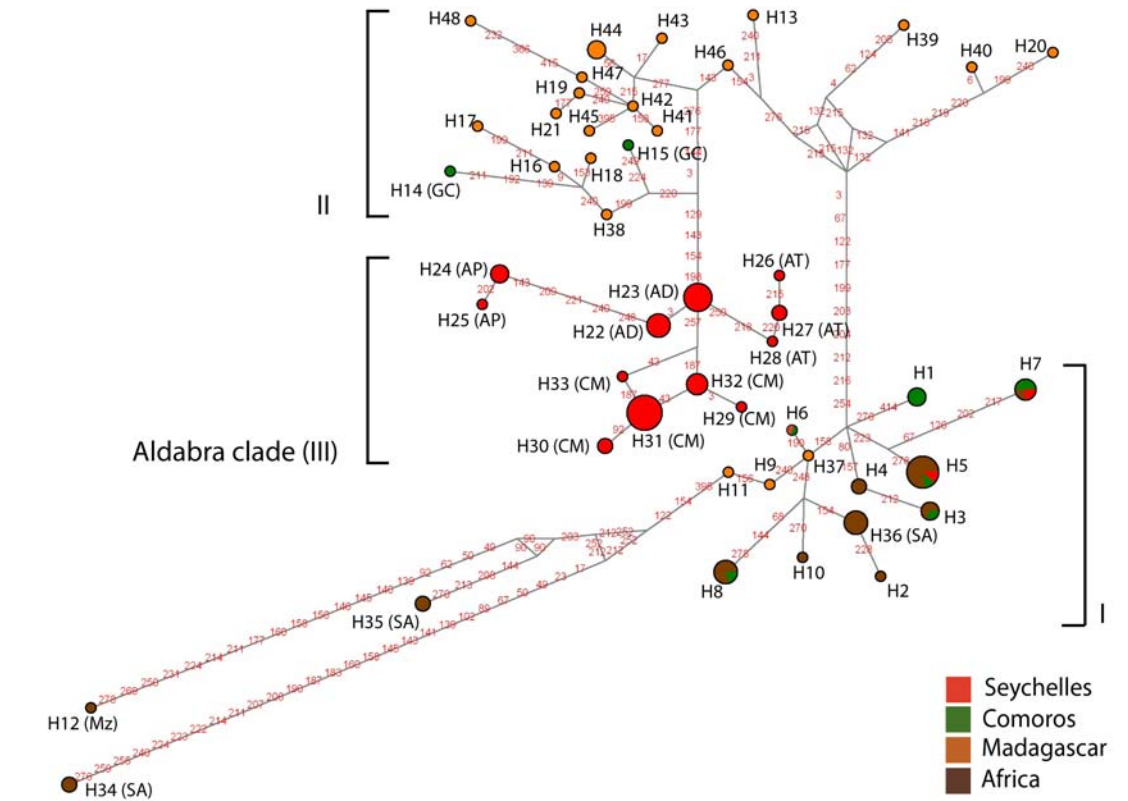


Fig. 2. MJ network for *H. mabouia-mercatorius* individuals. Haplotypes are colour-coded geographically and positions presenting substitutions are represented along the branches. The correspondence between haplotypes and individuals can be found in the Electronic Appendix.

lion years (half of the assumed mitochondrial vertebrate long-term mutation rate of 2%; Wilson, 1985) and 0.2 (a twenty-fold increase on the latter rate). Default priors were used for population size, modern day population size and exponential growth rate parameters. The analysis was run at least two times for 50×10^6 steps, sampled each 1000 steps, and 10% were discarded as burn-in. The software TRACER v. 1.4.1 (Rambaut & Drummond, 2007) was used to visualize the runs, checking for stationarity, convergence and adequacy of effective sample sizes (ESS) of parameters and to obtain posterior estimates and 95% credibility intervals for estimates.

RESULTS AND DISCUSSION

Hemidactylus mabouia “complex”

Examination of the haplotype networks concerning the *H. mercatorius-mabouia* “complex” reveal even more complicated patterns of variation than those previously described (Fig. 2). Clearly multiple cryptic lineages exist within this “species”, and the current taxonomic arrangement is completely inadequate. Haplotypes from Mozambique and South Africa define three deep, highly divergent lineages from the remaining group (clade A in Figure 4). Both PCR amplification and sequences were clean, without evidence of multiple bands or peaks, so we have no reason to be suspicious about the possibility of analysing paralogous fragments in these individuals (nu-

clear copies). Furthermore, base composition of these sequences was similar to the remaining ones in the dataset (average 18% T, 26% C, 31% A, 22% G), with values that fit well with those expected for vertebrate mitochondrial genomes (Asakawa et al., 1991). Differentiation between these main lineages (4) is high: all distance comparisons involving the South African lineages (H35 and H34) and the individual from Mozambique (H12) reveal distances of 2.7–8.3%, 7.3–8.5% and 5.8–7.7%, respectively (uncorrected average distances between groups) towards remaining individuals. If we invoke the 3% cut-off value for this same fragment used by Vieites et al. (2009) we may be in the presence of (unconfirmed) candidate species (UCS, *sensu* Vieites et al., 2009).

Within clade A, at least three further groups can tentatively be defined, each with considerable levels of diversity. One of the sub-clades (I in Figures 2 and 4) harbours haplotypes from the north of Madagascar, the Comoros (Mayotte), Tanzania including the islands of Zanzibar and Pemba, South Africa and the Seychelles (Mahé), as well as Uganda and the Gulf of Guinea islands, without any geographic structure. This species is not abundant either in the Comoros, or in the granitic Seychelles – just a few individuals were found during several weeks of fieldwork, on the island of Mahé, and they all belong to this clade, sharing haplotypes with African and Comoran individuals. From the distribution of the African haplotypes, it may be that both Mayotte and Mahé (and

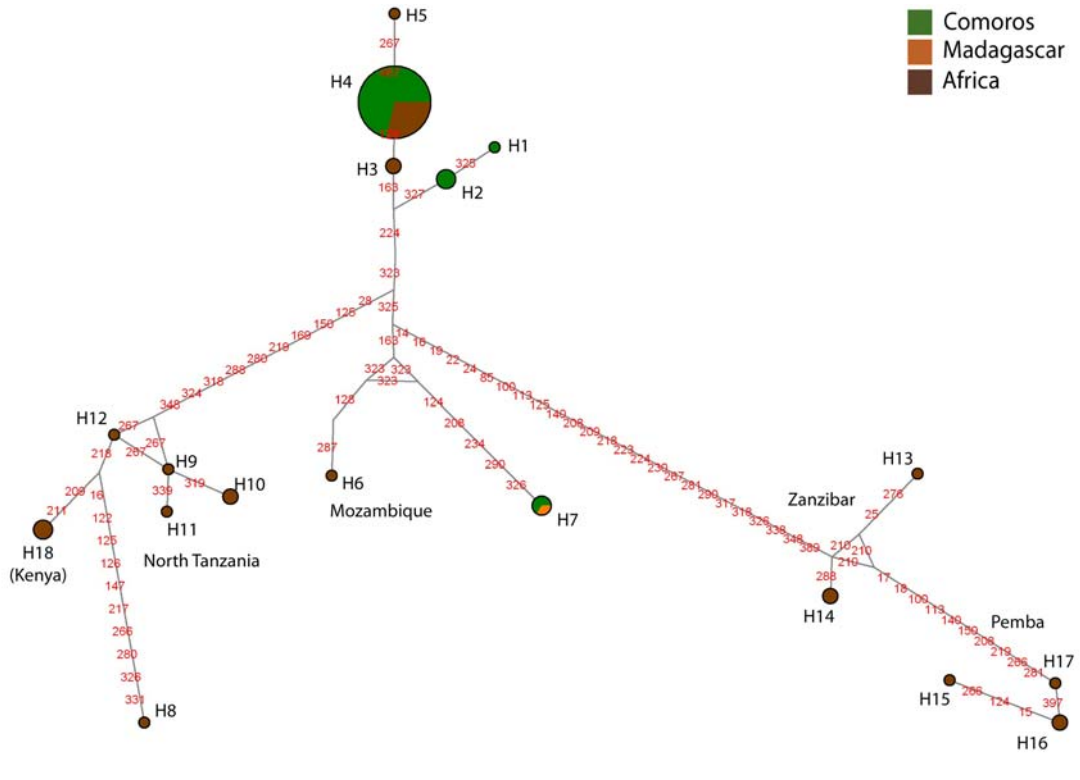


Fig. 3. MJ network for *H. platycephalus* individuals. Haplotypes are colour coded geographically and positions presenting substitutions represented along the branches. The correspondence between haplotypes and individuals can be found in the Electronic Appendix.

north Madagascar) were colonized recently multiple times from the East African coast or/and that dispersals from Mayotte, Mahé or North Madagascar back to Africa also do happen (possibly all corresponding to anthropogenic movements). It is also clear that cryptic diversity is present along the East African coast (haplotypes from South Africa belong to three different and highly divergent lineages), where variation within this species and its geographic distribution must be further investigated. The second group (II) of this Western Indian Ocean clade is defined mostly by Malagasy samples (and two individuals from Grand Comoro, Comoros), harbouring considerable haplotype diversity and structure, in what could be considered an endemic Malagasy clade (Boumans et al., 2007). Comoran haplotypes from this clade (two, from Grand Comoro) may again result from occasional introductions.

Closely related to this, haplotypes from the coralline archipelago of Aldabra (Aldabra, Cosmoledo, Astove, Assumption) define a well-supported monophyletic group (III) with considerable diversity and a remarkable phylogeographic structure, without any of these islands sharing any haplotype. This is a remarkable and surprising result given their geographical proximity and that Aldabra was supposedly completely submerged only about 125,000 years ago (Thomson & Walton, 1972) and the remaining islands in this group, lower than Aldabra itself, were also most probably submerged during this last interglacial period. The geographic structure of the distri-

bution of the haplotypes nevertheless clearly supports the native status of this species in these islands. As a test of the likelihood of this clade being native to the Aldabra archipelago we estimated the mutation rate needed to invoke if we assume a colonization of the island right after its emergence, implementing the 118,000–140,000-year interval from Thomson & Walton (1972) as a uniform prior on the root height and a uniform prior on the mutation rate from 0.005 to 0.2 substitutions/site/my. We recovered a roughly “normal-shaped” posterior distribution on the substitution rate, differing considerably from the prior used, with a mean of 0.07 substitutions/site/my but with a large 95% credibility interval (0.019–0.1457), which although higher than the traditional phylogenetic estimate of the rate of evolution, could correspond to the higher reported values of the mutation rate observed in “recent” (= low divergence) intraspecific datasets (Howell et al., 2003; Burridge et al., 2008), regardless of whether they are real or dataset or analysis artefacts (for discussion of this topic, see Ho et al., 2005; Emerson, 2007; Debruyne & Poinar, 2009). Explorative estimates using a relaxed clock model led to similar results (though with an even wider CI).

We thus have no reason to reject the hypothesis that this is a real endemic clade of the Aldabra archipelago, which is notable, given the recent age of these islands and the overall evidence for an otherwise anthropogenic-driven distribution of this species across the Western Indian Ocean islands.

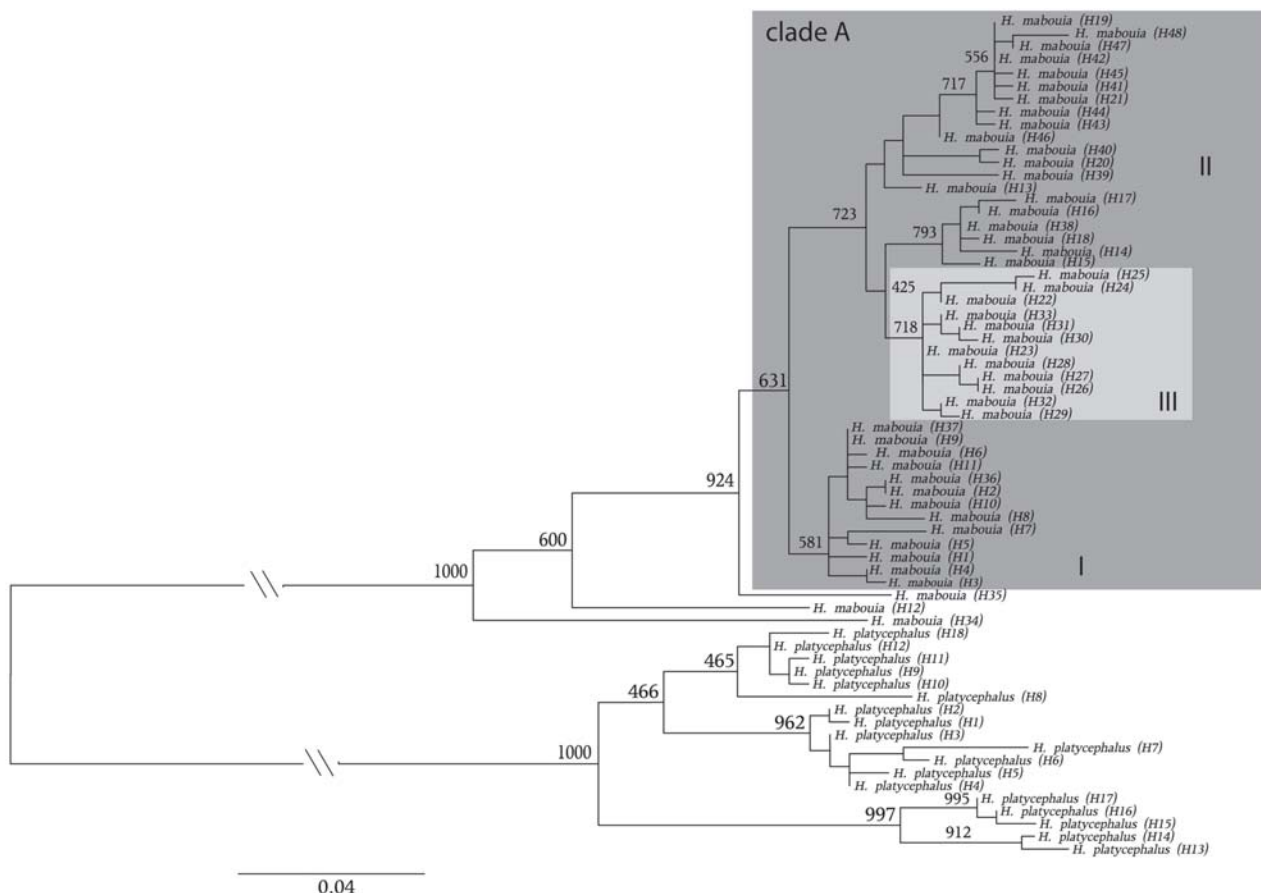


Fig. 4. Maximum likelihood tree for all haplotypes of *H. mabouia-mercatorius* and *H. platycephalus*. Relevant bootstrap values are indicated. Haplotype labelling is the same as in Figures 2 and 3. Clade A of *H. mabouia-mercatorius* and its subgroups are highlighted.

A comprehensive clarification of *H. mabouia*-complex taxonomy is not yet possible, but, given the clear differentiation observed between the basal (South African and Mozambique) lineages and the now well sampled insular populations, we propose that *H. mercatorius* becomes the designation applied to Gulf of Guinea, Comoran, Malagasy and Seychellois populations, with the status of eastern African populations remaining unchanged pending further studies. *Hemidactylus mercatorius* has most probably an extensive distribution across East Africa, possibly sympatric with *H. mabouia* across at least part of its range, but any taxonomic assignation of continental populations should wait until the distinction and distribution of these lineages across continental Africa is properly investigated.

Hemidactylus platycephalus

Hemidactylus platycephalus was not found in the Seychelles, but was the most common species across the Comoros (Rocha et al., 2005), inhabiting both natural and humanized environments (Carretero et al., 2005), and very abundant also along the East African coast and on the islands of Zanzibar and Pemba. Again, cryptic diversity is unveiled across East Africa with haplotypes from Zanzibar, Pemba, Tanzania (mainland), Mozambique and Kenya

defining at least three, highly differentiated, clades (Figs 3 and 4). Distances between these clades vary between 2.6 and 3.7%.

Again, for this species, considering the distribution of the African haplotypes (scattered across the three lineages) and the low or lack of diversity observed respectively in the Comoros and Madagascar, it is most probable that its presence on these islands is the result of recent (possible human-aided) dispersal events. Interestingly, some geographically coherent variation is observed in East Africa, with north Tanzania and Kenya and Zanzibar and Pemba harbouring exclusive lineages. Again, it is possible that multiple species occur under the current designation of *H. platycephalus*, and this should be further investigated through extended sampling and the use of additional nuclear data.

Hemidactylus frenatus

Finally, *H. frenatus*, not found on the East African coast, was the most abundant species of this genus in the granitic Seychelles (Mahé) and was also found across some of the low coralline islands of the Seychelles (Bird, Poivre and Desroches). Several distinct haplotypes were found across the Seychelles, Madagascar, the Mascarenes and the Comoros, but without any clear phylogeographic rela-

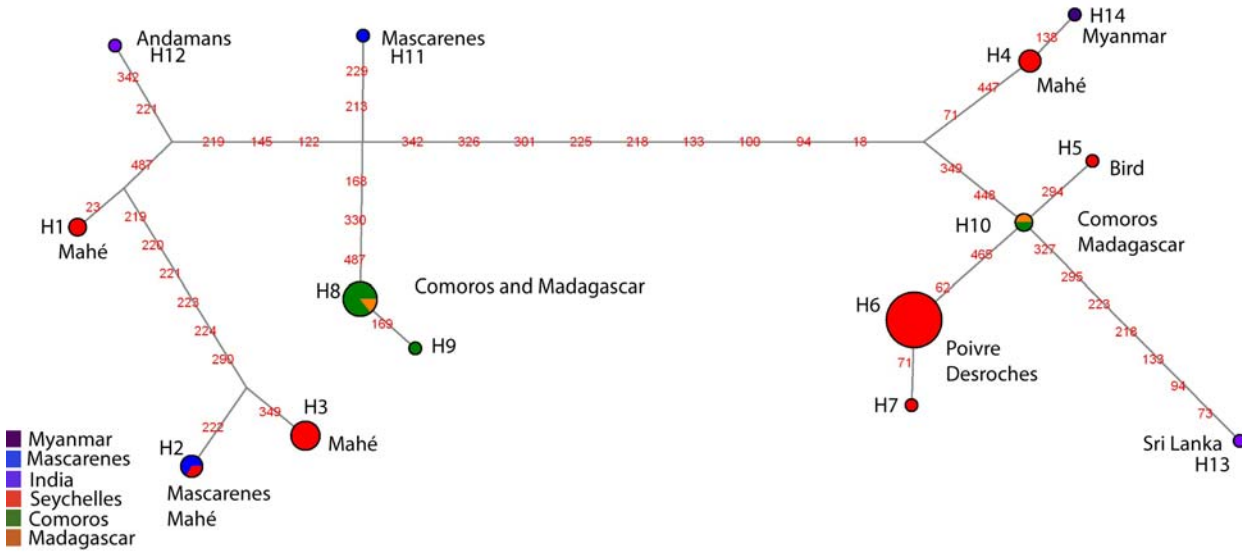


Fig. 5. MJ network for *H. frenatus* individuals. Haplotypes are colour-coded geographically and positions presenting substitutions are represented along the branches. The correspondence between haplotypes and individuals can be found in the Electronic Appendix.

tionship (Fig. 5), showing again, and also for the Seychelles, that its presence is most probably attributable to multiple, most likely human-aided, dispersal events. Of the two new Asian samples from this species included in this study, from Myanmar, one clearly clustered with this group of haplotypes (CAS885; H14) but another one (not included in the haplotype network) exhibited a highly divergent haplotype from these (around 10%), pointing again to cryptic diversity within this species in its native range (Carranza & Arnold, 2006).

In sum, despite the four species of *Hemidactylus* being generally widespread across the Western Indian Ocean islands, most of their range in this area may actually be the result of very recent (possibly human-aided) dispersal events. It now seems most probable that none of the *Hemidactylus* species occurring in the Comoros archipelago is actually native, the same being true for the granitic Seychelles. The only instances of natural colonization seem to be those of *H. mabouia* to Madagascar and from there to the coralline archipelago of Aldabra. Given the degree of differentiation observed, we propose that insular (Gulf of Guinea, Comoros, Madagascar, Seychelles) populations of *H. mabouia* are recognized as *H. mercatorius* (as also proposed for the Malagasy populations by Vences et al., 2004), with the status of continental populations pending further studies. Cryptic variation is depicted for the three species studied in their native range, with the cases of *H. platycephalus* and *H. mabouia* being remarkable in the sense that several highly divergent mitochondrial lineages are present along the eastern African coast. Nevertheless, extensive nuclear molecular data should also be collected and the geographical distributions of these lineages clarified before further taxonomic changes are implemented.

ACKNOWLEDGEMENTS

We are grateful to SBS of Seychelles for permits for fieldwork and tissue collection and to the Ministry of Environment and Natural Resources for logistic support during fieldwork. We thank Justin Gerlach from NPTS for support and information and to Raquel Vasconcelos and Ana Perera for help during fieldwork. Raquel Vasconcelos and Salvador Carranza kindly provided samples from Uganda and Kenya for this study, and Jens Vindum at the California Academy of Sciences (CAS) of *H. frenatus* from Myanmar. This research was supported by grants from Fundação para a Ciência e Tecnologia (FCT) SFRH/BD/1745/2004 (to SR), SFRH/BPD/27025/2006 (to MAC), POCTI/BSE/46647/2000 and PTDC/BIA-BDE/65745/2006 and by the Calouste Gulbenkian Foundation. The authors further acknowledge two anonymous reviewers for useful comments and remarks.

REFERENCES

- Asakawa, S., Kumazawa, Y., Araki, T., Himeno, H., Miura, K. & Watanabe, K. (1991). Strand-specific nucleotide composition bias in echinoderm and vertebrate mitochondrial genomes. *Journal of Molecular Evolution* 32, 511–520.
- Bandelt, H.J., Forster, P. & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16, 37–48.
- 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.
- Burridge, C.P., Craw, D., Fletcher, D. & Waters, J.M. (2008). Geological dates and molecular rates: fish DNA

- sheds light on time dependency. *Molecular Biology and Evolution* 25, 624–633.
- Carranza, S. & Arnold, E.N. (2006). Systematics, biogeography, and evolution of *Hemidactylus* geckos (Reptilia: Gekkonidae) elucidated using mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* 38, 531–545.
- Carretero, M.A., Harris, D.J. & Rocha, S. (2005). Recent observations of reptiles in the Comoro islands (Western Indian Ocean). *Herpetological Bulletin* 91, 19–28.
- Debruyne, R. & Poinar, H. (2009). Time dependency of molecular rates in ancient DNA data sets, a sampling artifact? *Systematic Biology* 58, 348–360.
- Drummond, A.J. & Rambaut, A. (2007). Beast: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7, 214.
- Emerson, B.C. (2007). Alarm bells for the molecular clock? No support for Ho et al.'s model of time-dependent molecular rate estimates. *Systematic Biology* 56, 337–345.
- Gray, J.E. (1842). Description of some new species of reptiles, chiefly from the British Museum collection. *The Zoological Miscellany*, 57–59.
- Guindon, S. & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* 52, 696–704.
- Harris, D.J., Arnold, E.N. & Thomas, R.H. (1998). Relationships of lacertid lizards (Reptilia: Lacertidae) estimated from mitochondrial DNA sequences and morphology. *Proceedings of the Royal Society B* 265, 1939–1948.
- Ho, S.Y., Phillips, M.J., Cooper, A. & Drummond, A.J. (2005). Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution* 22, 1561–1568.
- Howell, N., Smejkal, C.B., Mackey, D.A., Chinnery, P.F., Turnbull, D.M. & Herrnstadt, C. (2003). The pedigree rate of sequence divergence in the human mitochondrial genome: there is a difference between phylogenetic and pedigree rates. *American Journal of Human Genetics* 72, 659–670.
- Katoh, K., Kuma, K., Miyata, T. & Toh, H. (2005). Improvement in the accuracy of multiple sequence alignment program MAFFT. *Genome Information* 16, 22–33.
- Kluge, A.G. (1969). The evolution and geographic origin of the new world *Hemidactylus mabouia*–*brookii* complex (Gekkonidae, Sauria). *Miscellaneous Publications of the Museum of Zoology, University of Michigan* 138, 1–78.
- Palumbi, S.R., Martin, A., Romano, S., McMillan, W.O., Stice, L. & Grabowski, G. (1991). *The Simple Fool's Guide to PCR, Version 2.0*. Hawaii: Privately published.
- Polzin, T. & Daneschmand, S.V. (2003). On Steiner trees and minimum spanning trees in hypergraphs. *Operations Research Letters* 31, 12–20.
- Posada, D. (2008). JModeltest: phylogenetic model averaging. *Molecular Biology and Evolution* 25, 1253–1256.
- Rambaut, A. & Drummond, A.J. (2007). *TRACER v1.4*. Available from <http://beast.Bio.Ed.Ac.Uk/tracer>
- Rocha, S., Carretero, M.A. & Harris, D.J. (2005). Diversity and phylogenetic relationships of *Hemidactylus* geckos from the Comoro islands. *Molecular Phylogenetics and Evolution* 35, 292–299.
- Rocha, S., Harris, D.J., Perera, A., Silva, A., Vasconcelos, R. & Carretero, M.A. (2010). Recent data on the distribution of lizards and snakes of the Seychelles. *Herpetological Bulletin* 110, 20–32.
- Sambrook, J., Fritsh, E.F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Press.
- Taggart, J.B., Hynes, R.A., Prodoul, P.A. & Ferguson, A. (1992). A simplified protocol for routine total DNA isolation from salmonid fishes. *Journal of Fish Biology* 40, 963–965.
- Thomson, J. & Walton, A. (1972). Redetermination of the chronology of the Aldabra atoll by th/u dating. *Nature* 240, 145–146.
- Vences, M., Stefan, W., David, R.V., William, R.B., Frank, G. & Axel, M. (2004). Natural colonization or introduction? Phylogeographical relationships and morphological differentiation of house geckos (*Hemidactylus*) from Madagascar. *Biological Journal of the Linnean Society* 83, 115–130.
- Vieites, D.R., Wollenberg, K.C., Andreone, F., Kohler, J., Glaw, F. & Vences, M. (2009). Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. *Proceedings of the National Academy of Sciences of the United States of America* 106, 8267–8272.
- Wilm, A., Mainz, I. & Steger, A. (2006) An enhanced RNA alignment benchmark for sequence alignment programs. *Algorithms for Molecular Biology* 1, 19. doi:10.1186/1748-7188-1-19.
- Wilson, A.C. (1985). The molecular basis of evolution. *Scientific American* 253, 164–173.

Accepted: 22 February 2010