

Phylogenetic origin of the endemic pigeons from Madeira (*Columba trocaz*) and Azores Islands (*Columba palumbus azorica*)

Catarina G. Dourado · Margarida Alexandra Duarte · Ana Rita Grosso ·
Cristiane Bastos-Silveira · Patricia Marrero · Paulo Oliveira ·
Octávio S. Paulo · Deodália Dias

Received: 28 November 2012 / Revised: 7 June 2013 / Accepted: 1 July 2013
© Dt. Ornithologen-Gesellschaft e.V. 2013

Abstract This study addresses for the first time the issue of pigeon (genus *Columba*) phylogeny within the archipelagos of Madeira (*Columba trocaz*) and Azores (*C. palumbus azorica*), located in the singular biogeographic area of Macaronesia. The phylogeny of these endemic pigeons was inferred based on mitochondrial (cytochrome *b* and cytochrome *c* oxidase I) and nuclear (β -fibrinogen intron 7) genetic markers through multiple approaches. Despite the non-monophyletic pattern for the insular endemic species recovered in the phylogenies, topology tests presented somewhat different results. *C. trocaz*, the Madeiran endemic species, clustered strongly with the Canarian endemic *C. bollii*, and these two are thus more closely related to each other than *C. bollii* to *C. junoniae*, the other endemic species of Canary Islands, which seems to have

diverged independently. Moreover, *C. trocaz* was found to be phylogenetically closer to *C. bollii* than to *C. palumbus* from mainland Europe and Azores Islands. No genetic differentiation was found between the continental *C. p. palumbus* and the endemic *C. p. azorica*, which suggests a relatively recent colonisation event of the Azores Islands.

Keywords *Columba* · Endemic pigeons · Macaronesia · Phylogeny · Mitochondrial DNA · β -Fibrinogen intron 7

Zusammenfassung

Phylogenetische Ursprünge endemischer Taubenarten auf Madeira (*Columba trocaz*) und den Azoren (*Columba palumbus azorica*)

Diese Studie untersucht erstmalig phylogenetische Fragen bei Tauben (Gattung *Columba*) auf den Inselgruppen

Communicated by M. Wink.

Electronic supplementary material The online version of this article (doi:10.1007/s10336-013-0988-6) contains supplementary material, which is available to authorized users.

C. G. Dourado (✉) · M. A. Duarte · A. R. Grosso · D. Dias
Departamento de Biologia Animal, Faculdade de Ciências da
Universidade de Lisboa, Campo Grande, 1749-016 Lisboa,
Portugal
e-mail: anacdourado@gmail.com

C. G. Dourado · M. A. Duarte · C. Bastos-Silveira · D. Dias
Centro de Estudos do Ambiente e do Mar (CESAM-Lisboa),
1749-016 Lisboa, Portugal

C. G. Dourado · O. S. Paulo
Computational Biology and Population Genomics Group
(CoBiG²), Centro de Biologia Ambiental (CBA), Departamento
de Biologia Animal, Faculdade de Ciências da Universidade de
Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

M. A. Duarte · C. Bastos-Silveira
Museu Nacional de História Natural e da Ciência, Rua da Escola
Politécnica, 58, 1250-102 Lisboa, Portugal

P. Marrero
Island Ecology and Evolution Research Group, IPNA-CSIC,
E 38206 La Laguna, Tenerife, Canary Islands, Spain

P. Marrero
School of Biological Sciences, University of East Anglia,
Norwich NR4 7TJ, Norfolk, England

P. Oliveira
Parque Natural da Madeira (PNM), Quinta do Bom Sucesso,
Caminho do Meio, 9050-251 Funchal, Portugal

Madeira (*Columba trocaz*) und Azoren (*C. palumbus azorica*) in der biogeografischen Region Makaronesien. Die Phylogenie dieser endemischen Taubenarten wurde mittels multipler Ansätze auf der Grundlage mitochondrialer (Cytochrom b und Cytochrom-c-Oxidase I) und nukleärer (β -Fibrinogen Intron 7) genetischer Marker rekonstruiert. Obwohl sich in den Phylogenien ein nicht-monophyletisches Muster für die endemischen Inselarten zeigte, führten die Topologie-Tests zu etwas abweichenden Ergebnissen. *C. trocaz*, die für Madeira endemische Art, war eng mit dem kanarischen Endemiten *C. bollii* gruppiert; somit sind diese beiden näher miteinander verwandt als *C. bollii* und *C. junoniae*, der anderen endemischen Art der Kanaren, welche sich unabhängig davon entwickelt zu haben scheint. Außerdem zeigte sich, dass *C. trocaz* phylogenetisch *C. bollii* näher steht als *C. palumbus* des Europäischen Festlandes und der Azoren. Es gab keine genetische Differenzierung zwischen der Festlandform *C. p. palumbus* und dem endemischen Taxon *C. p. azorica*, was für eine erst relativ kürzlich erfolgte Besiedlung der Azoren spricht.

Introduction

Oceanic islands provide model systems to study the evolution, following colonisation, dispersal and isolation, of a variety of organisms, such as birds (Gill 1995). They are frequently characterised by ecological changes resulting from geological activity and distance to the mainland, being relatively simple systems isolated by barriers that reduce genetic interchanges with the continent; thus, both patterns of dispersal and speciation can be more easily assessed than in continental systems (e.g. Keogh et al. 2008; Rubínoff 2008; Schmitz et al. 2008; Sota and Nagata 2008; Su and Saunders 2009; Meik et al. 2010; Thorpe et al. 2010).

All the islands of the Macaronesian region—which is comprised of the Azores, Madeira, Salvages, Canaries and Cape Verde archipelagos—emerged due to volcanic activity during the Tertiary and Quaternary Periods and exhibit a broad range of geological ages (Carracedo 1994; Carracedo et al. 2002; Oromí 2004; Ferreira 2005; Sá-Pinto et al. 2008; Fernández-Palacios et al. 2011). Additionally, the Macaronesian biota has been linked the fauna and flora that spread from the Mediterranean basin in the late Miocene and Pliocene (Bramwell 1972; Juan et al. 2000). These features may have contributed to a variety of dispersal patterns, colonisation and diversification, and consequently to relatively high levels of genetic variation, endemism and speciation compared to other oceanic islands (Adler et al. 1995; Francisco-Ortega et al. 2000), representing one of the world's biodiversity hotspots, the Mediterranean Basin (Myers et al. 2000).

Many authors have studied dispersion, colonisation and speciation events in many islands, moreover phylogeographic patterns and ecology of some species (e.g. Adler et al. 1995; Coyne and Price 2000; Clegg et al. 2002; Warren et al. 2003; Russell et al. 2006; Stracey and Pimm 2009), but few have analysed the endemic species of the Madeira (e.g. flora; Piñeiro et al. 2009) and Azores archipelagos (e.g. land snails; Van Riel et al. 2005; Jordaens et al. 2009), particularly the genus *Columba* (e.g. Oliveira et al. 2002, 2006; Marrero et al. 2004; Martín et al. 2000; Marrero 2009; Gonzalez et al. 2009; Nogales et al. 2009). Concerning this genus in Macaronesia, most published work has focused on ecological and conservation aspects (e.g. Hernández et al. 1999; Oliveira et al. 2002, 2006; Marrero et al. 2004).

The genus *Columba* comprises 35 living species (Integrated Taxonomic Information System online database, <http://www.itis.gov>), with some of them present in Europe and the Macaronesian region. *C. palumbus* Linnaeus, 1758 is found in Europe and western Asia, *C. trocaz* Heineken, 1829 is endemic to the Madeira archipelago (Bannerman and Bannerman 1965), and *C. junoniae* Hartert, 1916 and *C. bollii* Godman, 1872 are both endemic to the Canary Islands (Bannerman 1963). Moreover, a subspecies of *C. palumbus*, *C. p. azorica* Hartert, 1905, is endemic to the Azores (Bannerman and Bannerman 1966), while *C. p. palumbus* Linnaeus, 1758 is found in mainland Europe (Cabal et al. 2005). The importance of these species also lies in their ecological significance, since they inhabit the relictual laurel forests of Macaronesia (Whittaker and Fernández-Palacios 2007). However, little is known about the phylogenetic relationships of these species, since molecular phylogenetic studies on Columbiformes have focused on the macroevolution of the entire order and did not include the Macaronesian endemic species (Johnson and Clayton 2000; Johnson et al. 2001). Only recently has Gonzalez et al. (2009) found strong support for a close relationship between *C. bollii* and *C. p. palumbus* using mitochondrial and nuclear genes. This phylogenetic relationship has long been proposed by zoologists (Stresemann 1927–1934), but also including other species, namely *C. trocaz*.

The first aim of this study was to access the phylogenetic relationship and understand the evolutionary history of the species *C. trocaz* (Madeira) and the subspecies *C. palumbus azorica* (Azores) in relation to the mainland subspecies *C. p. palumbus* and other Macaronesian species, specifically *C. junoniae* and *C. bollii*. The second goal was to reveal the genetic differentiation between the insular subspecies *C. p. azorica* and the mainland *C. p. palumbus*. In order to achieve these proposed goals, molecular phylogenetic analyses across different methodologies were performed using multi-locus molecular data: one nuclear gene region (β -fibrinogen intron 7—Fib7) and two mitochondrial

genes (cytochrome *b*—Cyt *b*, and cytochrome *c* oxidase I—COI).

Methods

Sampling and laboratory procedures

Samples from feathers, liver and muscle tissue of *C. trocaz*, *C. p. azorica*, *C. bollii*, *C. junoniae* and *C. p. palumbus*, were obtained from several locations (Fig. 1; ESM Table 1) and preserved in 96 % ethanol.

Genomic DNA was extracted from feathers using a NucleoSpin C+T Extraction Kit (Macherey-Nagel, Germany) following the manufacturer's instructions. For liver and muscle tissue samples, standard protocols requiring incubation with SDS and digestion with proteinase K, followed by a phenol–chloroform DNA extraction, were performed (Sambrook et al. 1989).

One nuclear and two mitochondrial fragments were targeted for amplification and sequencing. Polymerase

chain reactions were carried out with 250 ng of template DNA in a total volume of 25 μ l, containing a final concentration of 1 \times reaction buffer (Fermentas, Canada), 2.5 mM (Fib7 and COI) or 3 mM (Cyt *b*) of MgCl₂, 0.2 mM of each dNTP and 1 U of *Taq* DNA polymerase (Fermentas). Fib7 was amplified with the primers FIB-BI7L and FIB-BI7U (Prychitko and Moore 1997) at 0.3 μ M each, following the PCR conditions: 5 min at 92 °C, 30 cycles of 94 °C (40 s), 57 °C (45 s) and 72 °C (45 s), and a final extension at 72 °C for 10 min. COI gene was amplified using the primer pair BirdF1 (Hebert et al. 2004) and H7005 (Hafner et al. 1994) at 0.5 μ M each, under the following conditions: 5 min at 94 °C, 5 cycles of 94 °C (1 min), 47 °C (90 s) and 72 °C (90 s), 30 cycles of 94 °C (1 min), 53 °C (90 s) and 72 °C (90 s), and a final extension step during 5 min at 72 °C. Cyt *b* was amplified using the primer pair H4a (Harshman 1996) and L14841 (Kocher et al. 1989) at 0.3 μ M each, following the PCR conditions: 5 min at 94 °C, 5 cycles of 94 °C (1 min), 45 °C (90 s) and 72 °C (90 s), 30 cycles of 94 °C (1 min), 51 °C (90 s) and 72 °C (90 s), followed by a final extension at 72 °C for

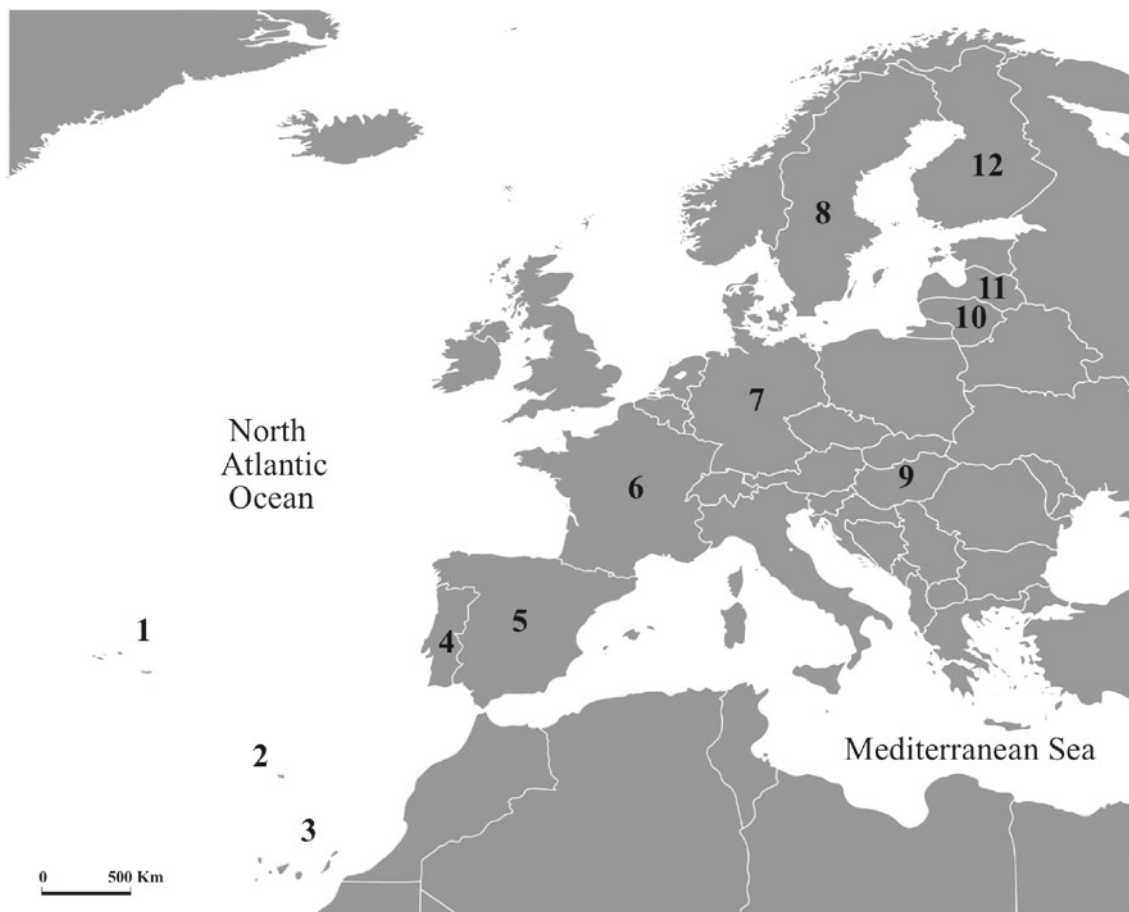


Fig. 1 The sampling locations for this study: 1 Azores Islands, 2 Madeira Island, 3 Canary Islands, 4 Portugal, 5 Spain, 6 France, 7 Germany, 8 Sweden, 9 Hungary, 10 Lithuania, 11 Latvia, 12 Finland

5 min. For *C. bollii*, the amplification of Cyt *b* was only successful with the primers L14764 (Sorenson et al. 1999) and H15646 (Sorenson et al. 1999) at 0.3 μ M each and 5 min at 92 °C, 35 cycles of 94 °C (30 s), 50 °C (45 s) and 72 °C (45 s), and a final extension for 10 min at 72 °C.

PCR products were verified through 1 % agarose gels and purified by using Quick Clean (Bioline, UK). Sequencing was carried out in both directions in an ABI Prism® 310 Genetic Analyzer (Applied Biosystems, USA) with the same primers used for amplification. Sequences were deposited in GenBank (ESM Table 1).

Data analysis

In addition to the obtained sequences, 10 Cyt *b* fragments from our previous work (Grosso et al. 2006) and several sequences from GenBank (see accession numbers in ESM Table 1) were added to our datasets.

Sequences were edited using BioEdit v.7.0.5.2 (Hall 1999), aligned with ClustalX v.2.1 (Thompson et al. 1997) and manually realigned whenever necessary. Sequences identity was confirmed using the BLAST tool from the NCBI website. A published sequence from chicken mtDNA (*Gallus gallus*—accession number X52392) was used to infer the position of our mtDNA sequences in relation to the avian genome.

In the mtDNA chromatograms, no double peaks or stop codons were found. However, nucleotide ambiguities of similar peak size were detected in chromatograms of the Fib7 fragment and were considered heterozygous base positions, according to the IUPAC nucleotide ambiguity code. Due to the low number of homozygous Fib7 samples, it was not possible to accurately infer the phase of each haplotype with the available software, namely PHASE (Stephens et al. 2001; Stephens and Scheet 2005). Thus, analyses on Fib7 were performed considering the phase that maximises the differences between haplotypes, which was manually achieved by comparing each heterozygous base position of each sample with the homozygous positions in the other samples.

Phylogenetic analyses under parsimony (UP) and maximum likelihood (ML) were performed employing PAUP* v.4.0.b10 (Swofford 2002). Modeltest v.3.7 software (Posada and Crandall 1998) associated with PAUP* was used to select the most appropriate evolutionary model for the different datasets, based on the Akaike Information Criterion. The most appropriate model for each case was then used to infer the ML phylogenetic tree (Felsenstein 1988) with a heuristic search with 100 random addition sequences and tree-bisection-reconnections. The optimal UP tree was found by a heuristic search with tree-bisection-reconnection as the branch-swapping algorithm. Initial trees were obtained via stepwise addition with 100

replicates of random addition sequence. Bootstrapping with 1,000 pseudo-replicates was performed to evaluate node robustness of phylogenetic trees. Gaps were treated as missing data or removed from the analyses.

A Bayesian inference (BI) was undertaken employing MrBayes v.3.1 (Ronquist and Huelsenbeck 2003). The posterior probabilities of the phylogenetic trees were estimated by a Metropolis-Coupled, Markov Chain Monte Carlo sampling algorithm (MCMCMC). The Markov Chain Monte Carlo (MCMC) procedure ensures that trees are sampled in proportion to their probability of occurrence under the given model of gene-sequence evolution, while the MCMCMC approach ensures that the Markov chain did not become trapped in local optimum. The conditions for the Bayesian analysis were previously set up and in order that the likelihood scores of the trees reach stationarity with the generation's time, which was checked by Tracer v.1.5 (Rambaut and Drummond 2007), a total of 1.6×10^6 generations were sampled every 100 generations with a “burn-in” of the first 5 % of the trees.

For each dataset, the selection of the appropriate model was carried out with MrModeltest v.2.2 (Nylander 2004) and implemented according to the author's recommendations. A combined matrix of both mitochondrial genes was set up and the previously selected models were used to analyse the data (Nylander et al. 2004). Different partitions were allowed to evolve at different rates, with unlinked topology and parameters of the nucleotide models across partitions. For each model, four runs were carried out using different random starting seed to assess congruence of the likelihood values (Huelsenbeck et al. 2002). The model likelihood values for each analysis were compared, and then the best one was used to determine the topology branch length and clade robustness through the Bayesian posterior probability.

Several datasets were analysed with the described set of techniques. Each partition dataset was analysed independently and the mtDNA genes were concatenated as well as the nuclear and mitochondrial fragments. Analyses were carried out for separate and combined datasets after testing for incongruence. The incongruence length difference test (ILD) (Farris et al. 1995) was implemented in PAUP* with all invariant characters removed. Concatenation and data preparation according to each software requirements were carried out using Concatenator v.1.1.0 (Pina-Martins and Paulo 2008). The Philippine Cuckoo Dove *Macropygia tenuirostris* Bonaparte, 1854 and the Mackinlay's Cuckoo Dove *Macropygia mackinlayi* Ramsay, 1878 were used as outgroups for tree rooting.

Topology tests were performed comparing the recovered optimal phylogenies (see Figs. 2, 3) with alternative constraint tree topologies according to different biological hypothesis. The Shimodaira–Hasegawa test (SH test)

(Shimodaira and Hasegawa 1999) was implemented in PAUP*. Two datasets (dataset 1: Cyt *b* + COI; dataset 2: Fib7) were studied with three different constraints. Firstly, it was tested whether *C. bollii* and *C. junoniae* could be considered sister taxa. In the second approach, it was intended to test whether *C. trocaz* and *C. palumbus* could form a sister taxa. Finally, the third hypothesis demands that all four Macaronesian species form a monophyletic clade with *C. junoniae* as the basal species. One thousand replicates were carried out in each SH test.

Results

Sequence variation

A total of 53 Cyt *b*, COI and Fib7 haplotypes were generated in the present work and submitted to GenBank. Details regarding sample size and origin of the studied taxa can be checked in Fig. 1 and ESM Table 1.

Concerning Cyt *b*, 10 unique haplotypes were identified for *C. palumbus*, 4 for *C. trocaz*, 2 for *C. junoniae* and 2 for *C. bollii*. The COI sequence data revealed 12 unique haplotypes for *C. palumbus*, 5 for *C. trocaz*, 1 for *C. junoniae* and 1 for *C. bollii*. Regarding the nuclear data, 18 unique haplotypes were found for *C. palumbus*, 2 for *C. trocaz*, 2 for *C. junoniae* and 1 for *C. bollii*.

No gaps or premature stop codons were identified in the mtDNA sequences, and therefore all represented uninterrupted open reading frames, suggesting that they were functional. Some indels occurred in Fib7; the biggest one was a deletion of 165 bp, detected in the sequences of Old World pigeons. Six indels, in which the insertions were only present in a small number of sequences, were removed from the analyses, and none were heterozygotic for length variation. A 10-bp deletion was detected in *C. junoniae* and it was kept for the analysis. Almost all *C. palumbus* sequences showed potential heterozygous sites, from one to a maximum of seven; the majority C/T transitions were followed by few G/A. For *C. trocaz*, all sequences were identical and only one potential heterozygous site was detected.

Five datasets were analysed: three representing each gene individually, one with both mtDNA genes concatenated, and another concatenation concerning all fragments. For the single datasets, only the unique haplotypes were used. The mtDNA genes dataset includes 12 sequences from the present work and 19 from GenBank, and the Fib7 dataset comprises 26 and 22, respectively (Table 1; ESM Table 1).

Variability and phylogenetic model details for each gene combination analysed can be seen in Table 1.

The χ^2 test of homogeneity of base frequencies across taxa showed no significant difference for any of the

datasets analysed (mtDNA: $\chi^2 = 40.248$, $df = 90$, $p = 0.99$; Fib7 $\chi^2 = 9.599$, $df = 141$, $p = 1.00$; mtDNA and nuclear: $\chi^2 = 41.656$, $df = 96$, $p = 0.99$).

The ILD test for the concatenation of the two mtDNA genes showed that the generated fragments were not incongruent ($p = 0.88$) as well as the concatenation of all fragments ($p = 0.70$).

For the concatenated mtDNA dataset, 162 equal parsimoniously trees were found with length of 1,039. For the nuclear dataset, were found 6 equal parsimonious trees with length of 159. Regarding the concatenation of all fragments, 630 equally parsimonious trees were found, with a length of 1,174.

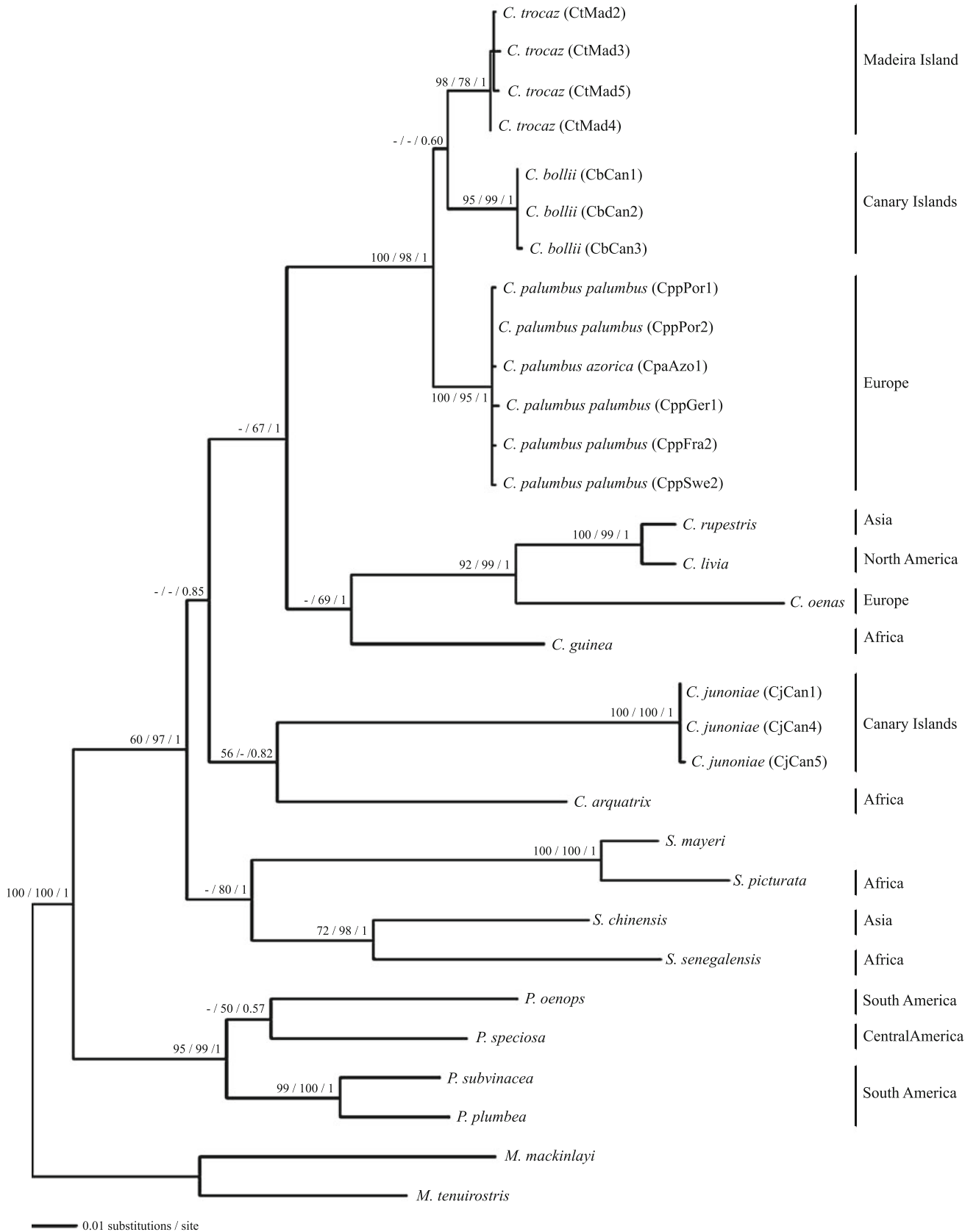
Phylogenetic relationships

Figure 2 shows the phylogram obtained by ML with the selected evolutionary model for the concatenated mtDNA dataset (Table 1). The BI method produced a similar topology but with different branch support (Fig. 2). Only UP produced a different topology, with the species *C. guinea* being basal to the *C. palumbus*, *C. trocaz* and *C. bollii* clade, and the species *C. junoniae*, *C. arquatrix*, *C. rupestris*, *C. livia* and *C. oenas* within the same group as a sister clade.

When the same analyses were carried out separately for each mtDNA gene, similar topologies to the ones with combined dataset were produced (data not shown). However, some minor differences were observed in the phylogram inferred for the COI fragment. While in the phylogeny recovered from the mtDNA dataset *Streptopelia* spp. form a basal group to *C. trocaz*, *C. bollii*, *C. palumbus*, *C. rupestris*, *C. livia*, *C. oenas*, *C. livia*, *C. junoniae* and *C. arquatrix*, in the COI phylogram, *C. arquatrix* assumes a basal position in relation to all the mentioned species. In addition to this, *C. trocaz* and *C. bollii* do not form a clade.

Figure 3 shows the ML tree inferred for the Fib7 dataset with the more adequate evolutionary model (Table 1). Trees inferred with different methods produced similar topologies. Topologies of the main clades were similar to the one generated by the mtDNA dataset (Fig. 2). The main difference is the position of *C. junoniae*, which is basal in the nuclear gene (Fig. 3) but forms a sister basal clade with *C. arquatrix* in the two other analyses (Figs. 2, 4). Data from Fib7 produced higher support values for the monophyletic relation between *C. bollii* and *C. trocaz* than both mtDNA genes (see Figs. 2, 3).

Regarding the results of the concatenated dataset (Cyt *b* + COI + Fib7), the results are similar regardless of the methodology. Figure 4 shows the ML tree resulting from the concatenation of both mtDNA genes and the nuclear fragment, which presents a similar topology to the mtDNA tree (Fig. 2), but with an overall increase of the branch support.



◀ **Fig. 2** Maximum likelihood phylogenetic tree based on mtDNA concatenated data. The geographic region from where each sample was obtained is presented. *Names* in the terminal nodes refer to the sample code, as in Table 1 of the ESM. *Numbers* above branches are the bootstrap support values (>50 %) obtained from 1,000 pseudo-replicates from maximum parsimony and maximum likelihood; the last value corresponds to the Bayesian posterior probability (>0.5)

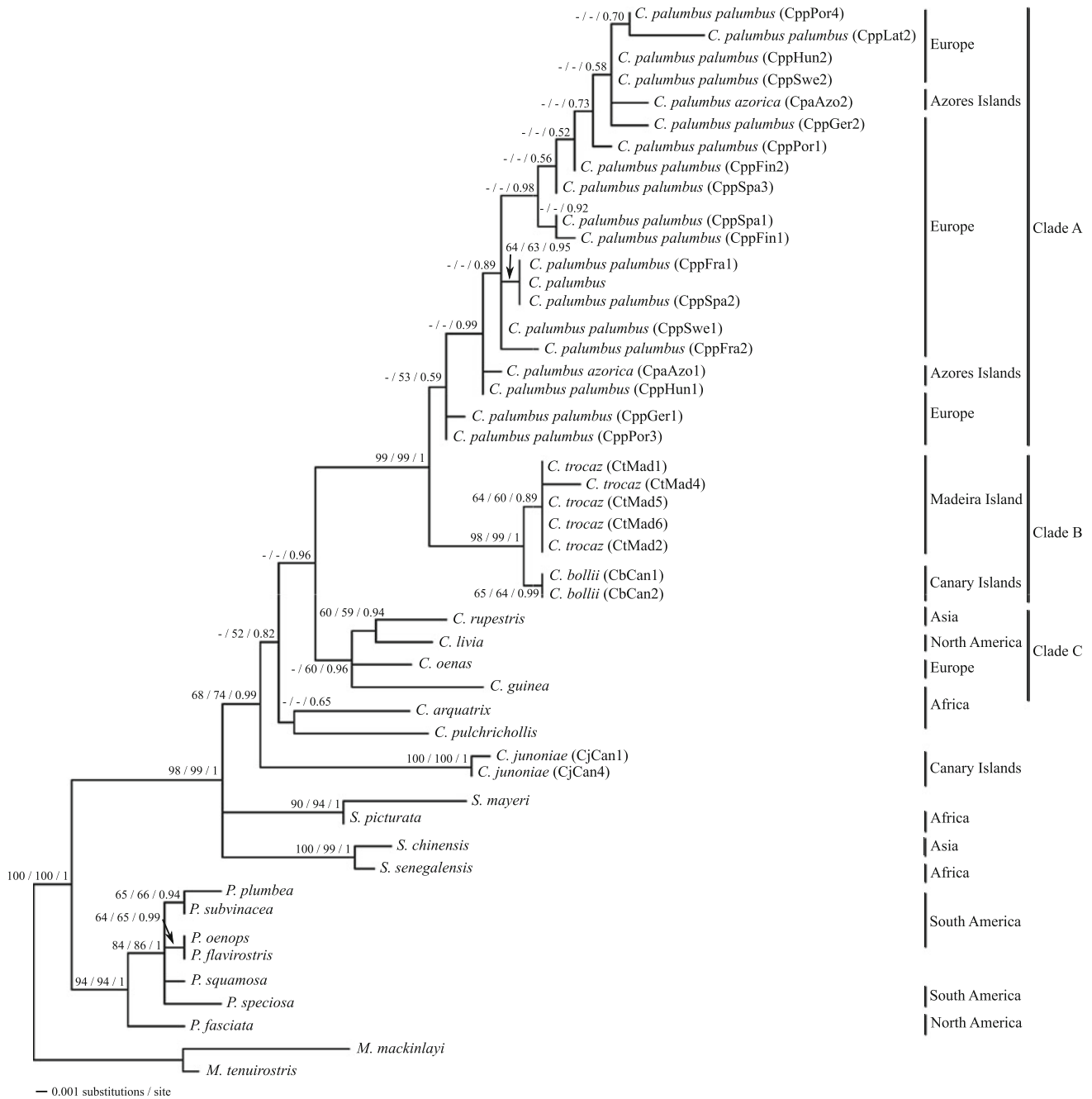


Fig. 3 Maximum likelihood phylogenetic tree based on Fib7. The geographic region from where each sample was obtained is presented. *Names* in the terminal nodes refer to the sample code, as in Table 1 of the ESM. *Numbers* above branches are the bootstrap support values

Although with minor variations, all phylogenetic results seem to be congruent, leading to two well-defined and supported clades. Patagioenas species form a basal clade (Clade F*) (see Fig. 4). Then, there is a clade (Clade E*) formed by *Streptopelia* spp. (see Fig. 4). Clade D* is composed of *C. arquatrix* and *C. junoniae*; Clade C* is

(>50 %) obtained from 1,000 pseudo-replicates from maximum parsimony and maximum likelihood; the last value corresponds to the Bayesian posterior probability (>0.5)

Table 1 Variability and nucleotide substitution model details of each dataset: number of taxa, fragment size, number of variable and parsimony informative sites, selected nucleotide substitution model, shape parameter of the gamma distribution (Γ), proportion of invariable sites (I), individual substitution rates and number and length of maximum parsimony trees

| Dataset | Cyt <i>b</i> + COI | Fib7 | All genes |
|-----------------------|--------------------|---------|-------------|
| Taxa | 31 | 48 | 33 |
| Parameter | | | |
| Size (bp) | 1,756 | 897 | 2,653 |
| Variable sites | 473 | 128 | 583 |
| Informative sites | 320 | 73 | 385 |
| Model | TVM + I + G | GTR + G | TVM + I + G |
| Γ | 2.0556 | 0.3764 | 0.5698 |
| I | 0.6244 | | 0.5397 |
| A–C | 4.1219 | 0.8410 | 1.2612 |
| G–A | 63.0602 | 3.3252 | 14.2452 |
| A–T | 4.4792 | 0.5584 | 1.0476 |
| C–G | 0.7110 | 1.5304 | 0.5261 |
| C–T | 63.0602 | 4.5933 | 14.2452 |
| No. UP trees (length) | 162 (1,039) | 6 (159) | 630 (1,174) |

composed of *C. livia*, *C. rupestris*, *C. oenas* and *C. guinea*. Clade B* is a well-supported group that includes *C. bollii* and *C. trocaz*. Lastly, Clade A* comprises the subspecies *C. palumbus palumbus* and *C. palumbus azorica*, always forming a well-supported clade. In all analyses performed, it was impossible to distinguish *C. p. azorica* from *C. p. palumbus*. Thus, besides the nature of the genetic markers used (mitochondrial and nuclear DNA), these two subspecies are genetically undifferentiated. Moreover, no differentiation was found among the Azores Islands.

Several topology tests were performed. In the first hypothesis, in which it was assumed that *C. bollii* and *C. junoniae* are sister taxa, the unconstrained tree is significantly better than the constrained tree, not only in the mtDNA dataset ($p = 0.007$) but also in the nuclear dataset ($p = 0.000$). When *C. trocaz* and *C. palumbus* are considered sister taxa (second hypothesis), different results are obtained depending on the dataset used. Regarding the mtDNA dataset, no significant differences were found between the unconstrained and the constrained tree ($p = 0.409$). However, when the nuclear dataset was tested, a significant difference ($p = 0.026$) was found. The third hypothesis implies that all Macaronesian species are a monophyletic group. For both datasets tested, non-significant values were obtained between the unconstrained and the constrained trees (mtDNA: $p = 0.133$; nuclear: $p = 0.106$).

Discussion

Consistently, species from Canary Islands (*C. junoniae* and *C. bollii*) and Africa (*C. arquatrix* and *C. guinea*) are dispersed in different clades (see Figs. 2, 3, 4), and thus Macaronesian species do not form a monophyletic clade, supporting Gonzalez et al. 2009. Nevertheless, a SH test provided a slightly different result. For instance, in the hypothesis considering that all *Columba* from Macaronesian region are monophyletic, no significant differences were detected in relation to the optimal tree. This means that it is possible that the phylogenies achieved through different approaches may not be fully correct, implying that *C. junoniae* could form a monophyletic group with the other insular species, instead of forming a clade with an African species. Thus, although a non-monophyletic relationship of the studied species was revealed by the phylogeny recovered from multiple methodologies, the alternative monophyletic pattern cannot be entirely excluded.

A splitting event may have originated two sister clades, one including *C. palumbus* from the mainland Europe and the Azores (Clade A*), and another one comprised of *C. trocaz* and *C. bollii*, endemics to Madeira and Canary Islands, respectively (Clade B*) (see Fig. 4). A SH test was performed to determine if *C. trocaz* and *C. palumbus* as sister taxa could be an alternative explanation. Since results from mitochondrial and nuclear DNA are contradictory, we cannot rule out the alternative phylogeny; nonetheless, it implies a less parsimonious biogeographic process, a higher number of independent colonisations being necessary—an initial colonisation to the Canary Islands that gave origin to *C. bollii*, and a second one to Madeira, that gave origin to *C. trocaz*. However, considering the geographic location of the Canaries, Madeira, Azores Islands and mainland Europe, and also some already known dispersal patterns in other species (e.g. Böhle et al. 1996; Kim et al. 2008), *C. trocaz* is probably more phylogenetically related to *C. bollii* than to *C. palumbus*, assuming that both archipelagos of Madeira and Canaries were colonized by an ancestor of these insular species, which then diverged due to geographic isolation.

No differentiation was observed between the insular endemic subspecies *C. p. azorica* and the continental *C. p. palumbus*, despite the large geographic distance separating the Azores Islands and mainland Europe (about 1,500 km). This suggests a relatively recent colonisation of the Azores by the continental *C. palumbus*. The lack of molecular support for the taxonomic separation, based on morphology and geographical range (Cramp et al. 1985), seems to be related to the nature of the markers used in the present study.

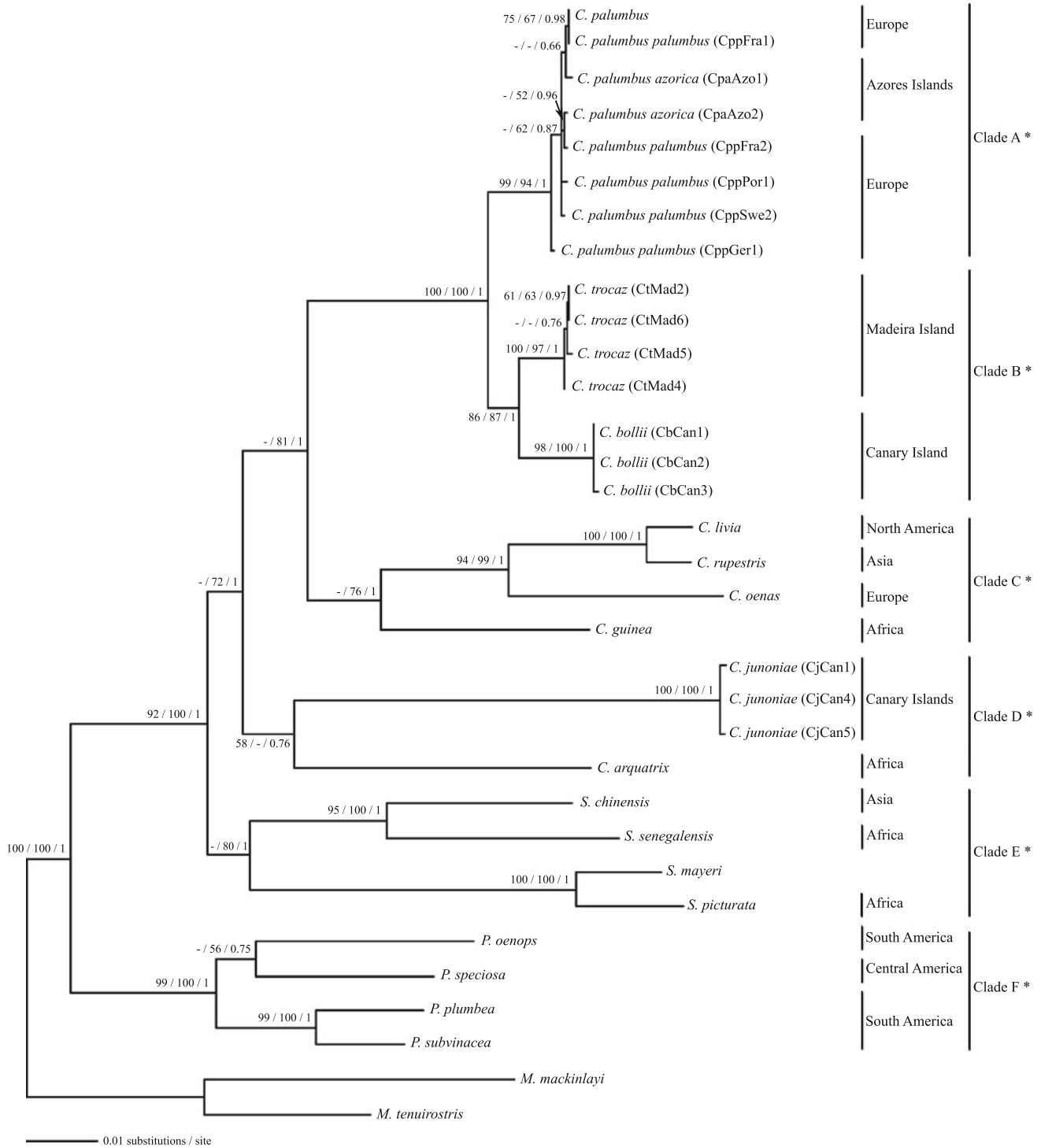


Fig. 4 Maximum likelihood phylogenetic tree based on all fragments concatenated. The geographic region from where each sample was obtained is presented. *Names* in the terminal nodes refer to the sample code, as in Table 1 of the ESM. *Numbers* above branches are the

bootstrap support values (>50 %) obtained from 1,000 pseudo-replicates from maximum parsimony and maximum likelihood; the last value corresponds to the Bayesian posterior probability (>0.5)

Additionally, we found that *C. p. palumbus* individuals from different European locations sustain our previous findings based only on cytochrome *b* sequences (Grosso

et al. 2006), which did not support the existence of a geographically-based divergence among *C. p. palumbus* populations. This may be related to the migratory

behaviour of this widespread species, with the northern populations being migratory, while the southern ones are essentially sedentary (Elias et al. 1999).

Gonzalez et al. (2009) found that the two species from the Canary Islands are not closely related. They observed that *C. bollii* clusters with *C. palumbus*, while *C. junoniae* was placed at the base of the clade that included other *Columba* species; however, they have not included *C. trocaz* in their analyses. We observed that *C. junoniae* seems to have evolved from an African stock while *C. bollii* appears to have evolved from a European stock, with *C. trocaz* as its closest relative and with *C. palumbus* as a sister clade. Conversely, we performed a topology test in which it was considered that *C. bollii* and *C. junoniae* are sister taxa, and significant differences between the unconstrained and the constrained trees were obtained between the two datasets. This means that the optimal trees are more accurate in explaining the phylogenetic relationships between these two insular endemic species, and they are probably not sister taxa.

Although these frugivorous pigeons from the Canaries currently inhabit the relict laurel forests in sympatry, they might have distinct genetic origins. In this case, these results could be explained by an independent colonisation event from mainland Africa or Europe to the Canary Islands, originating *C. junoniae*, and a second colonisation of the Canary Islands from a common ancestor with *C. palumbus* followed by dispersal to Madeira Islands (or vice versa), which could explain the clustering of *C. bollii* and *C. trocaz* and the species *C. palumbus* as the sister clade. These hypotheses are partially supported by the previous divergence time estimates that point to an old colonisation event to the Canary Islands around 20 Mya, originating *C. junoniae*, and a more recent episode around 5 Mya that originated the *C. bollii* pigeons (Gonzalez et al. 2009).

Similar colonisation events have been observed in a diversity of Macaronesian organisms, such as plants (e.g. Böhle et al. 1996; Francisco-Ortega et al. 1996; Vargas et al. 1999; Carine et al. 2004; Kim et al. 2008), reptiles (e.g. Carranza et al. 2000, 2001) and arthropods (e.g. Brunton and Hurst 1998; Ribera et al. 2003; Jordall and Hewitt 2004; Dimitrov et al. 2008; Villacorta et al. 2008).

Further investigation using more nuclear markers and a more robust sampling will help us to elucidate the origin and direction of the colonisations among the Azores, Madeira and Canary Islands, and this is currently underway.

Acknowledgments We are grateful to Dr. J. Faustino (Direcção Geral dos Recursos Florestais, Portugal), Prof. José Madeira (Faculdade de Ciências da Universidade de Lisboa, Portugal), Conseil Régional de La Chasse d'Aquitaine (France) and Federación Guipuzcoana de Caza (Spain) for providing biological samples for this study. We would like to thank the two anonymous reviewers for

providing useful comments on the manuscript. Margarida Alexandra Duarte participation was supported by a PhD grant FCT (SFRH/BD/70646/2010), Patricia Marrero was supported by a postdoctoral grant of the Spanish Minister of Education, Cristiane Bastos-Silveira was supported by a postdoctoral grant FCT (SFRH/BPD/26612/2005) and Octávio S. Paulo was supported by the grant FCT and FEDER (POCTI/BSE/48365/2002).

References

- Adler GH, Austin CC, Dudley R (1995) Dispersal and speciation of skinks in the tropical Pacific Ocean among archipelagos. *Evol Ecol* 9:529–541
- Bannerman DA (1963) Birds of the Atlantic Islands, vol 1. A history of the birds of the Canary Islands and of the Salvages. Oliver & Boyd, Edinburgh
- Bannerman DA, Bannerman WM (1965) Birds of the Atlantic Islands, vol 2. A history of the birds of Madeira, the Desertas, and the Porto Santo Islands. Oliver & Boyd, Edinburgh
- Bannerman DA, Bannerman WM (1966) Birds of the Atlantic Islands, vol 3. A history of the birds of the Azores. Oliver & Boyd, Edinburgh
- Böhle UR, Hilger HH, Martin WF (1996) Island colonization and evolution of the insular woody habit in *Echium* L. (Boraginaceae). *Proc Natl Acad Sci USA* 93:11740–11745
- Bramwell D (1972) Endemism in the flora of the Canary Islands. In: Valentine DH (ed) Taxonomy, phytogeography and evolution. Academic, London, pp 141–159
- Brunton CFA, Hurst GDD (1998) Mitochondrial DNA phylogeny of Brimstone butterflies (genus *Gonepteryx*) from the Canary Islands and Madeira. *Biol J Linn Soc* 63:69–79
- Cabral MJ, Almeida J, Almeida PR, Dellinger T, Almeida NF, Oliveira ME, Palmeirim J, Queiroz AI, Rogado L, Reis MS (2005) Livro Vermelho dos Vertebrados de Portugal. Instituto de Conservação da Natureza, Lisboa
- Carine MA, Russell SJ, Santos-Guerra A, Francisco-Ortega J (2004) Relationships of the Macaronesian and Mediterranean floras: molecular evidence for multiple colonizations into Macaronesia and back-colonization of the continent in *Convolvulus* (Convolvulaceae). *Am J Bot* 91:1070–1085
- Carracedo JC (1994) The Canary Islands: an example of structural control on the growth of large oceanic-island volcanoes. *J Volcanol Geotherm Res* 60:225–241
- Carracedo JC, Pérez FJ, Ancochea E, Meco J, Hernán F, Cubas CR, Casillas R, Rodríguez E, Ahijado A (2002) Cenozoic volcanism II: the Canary Islands. In: Gibbons W, Moreno T (eds) The geology of Spain. Geological Society of London, London, pp 439–472
- Carranza S, Arnold EN, Mateo JA, López-Jurado LF (2000) Long-distance colonization and radiation in gekkonid lizards, *Tarentola* (Reptilia: Gekkonidae), revealed by mitochondrial DNA sequences. *Proc R Soc Lond B* 267:637–649
- Carranza S, Arnold EN, Mateo JA, López-Jurado LF (2001) Parallel gigantism and complex colonization patterns in the Cape Verde scincid lizards *Mabuya* and *Macroscoincus* (Reptilia: Scincidae) revealed by mitochondrial DNA sequences. *Proc R Soc Lond B* 268:1595–1603
- Clegg SM, Degnan SM, Kikkawa J, Moritz C, Estoup A, Owens IPF (2002) Genetic consequences of sequential founder events by an island-colonizing bird. *Proc Natl Acad Sci USA* 99:8127–8132
- Coyne JA, Price TD (2000) Little evidence for sympatric speciation in islands birds. *Evolution* 54:2166–2171
- Cramp S (ed) (1985) The Birds of the Western Palearctic, vol IV—Terns to Woodpeckers. Oxford University Press, New York

- Dimitrov D, Arnedo MA, Ribera C (2008) Colonization and diversification of the spider genus *Pholcus* Walckenaer, 1805 (Araneae, Pholcidae) in the Macaronesian archipelagos: evidence for long-term occupancy yet rapid recent speciation. *Mol Phylogenet Evol* 48:596–614
- Elias GL, Reino LM, Silva T, Tomé R, Gerales P (1999) Atlas das Aves Invernaes do Baixo Alentejo. Sociedade Portuguesa para o Estudo das Aves, Lisboa, pp 234–235
- Farris JS, Källersjö M, Kluge AG, Bult C (1995) Testing significance of incongruence. *Cladistics* 10:315–319
- Felsenstein J (1988) Phylogenies from molecular sequences: inference and reliability. *Annu Rev Genet* 22:521–565
- Fernández-Palacios JM, de Nascimento L, Otto R, Delgado JD, García-del-Rey E, Arévalo JR, Whittaker RJ (2011) A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *J Biogeogr* 38:226–246
- Ferreira AB (2005) Geodinâmica e perigosidade natural das ilhas dos Açores. *Finisterra* XL 79:103–120
- Francisco-Ortega J, Jansen RK, Santos-Guerra A (1996) Chloroplast DNA evidence of colonization, adaptive radiation, and hybridization in the evolution of the Macaronesian flora. *Proc Natl Acad Sci USA* 93:4085–4090
- Francisco-Ortega J, Santos-Guerra A, Kim S-C, Crawford DJ (2000) Plant genetic diversity in the Canary Islands: a conservation perspective. *Am J Bot* 87:909–919
- Gill FB (1995) Ornithology, 2nd edn. Freeman, New York
- Gonzalez J, Castro GD, Garcia-del-Rey E, Bergerand C, Wink M (2009) Use of mitochondrial and nuclear genes to infer the origin of two endemic pigeons from the Canary Islands. *J Ornithol* 150:357–367
- Grosso AR, Bastos-Silveira C, Coelho MM, Dias D (2006) *Columba palumbus* Cyt *b*-like *Numt* sequence: comparison with functional homologue and the use of universal primers. *Folia Zool* 55:131–144
- Hafner MS, Sudman PD, Villablanca FX, Spradling TA, Demastes JW, Nadler SA (1994) Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science* 2625:1087–1090
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
- Harshman J (1996) Phylogeny, evolutionary rates, and ducks. PhD dissertation, University of Chicago, Chicago
- Hebert PDN, Stroeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. *PLoS Biol* 2:1657–1663
- Hernández MA, Martín A, Nogales M (1999) Breeding success and predation on artificial nest of the endemic pigeons Bolle's Laurel Pigeon *Columba bollii* and White-tailed Laurel Pigeon *Columba junoniae* in the laurel forest of Tenerife (Canary Islands). *Ibis* 141:52–59
- Huelsenbeck JP, Larget B, Miller RE, Ronquist F (2002) Potential applications and pitfalls of Bayesian inference of phylogeny. *Syst Biol* 51:673–688
- Johnson KP, Clayton DH (2000) Nuclear and mitochondrial genes contain similar phylogenetic signal for pigeons and doves (Aves: Columbiformes). *Mol Phylogenet Evol* 1:141–151
- Johnson KP, de Kort S, Dinwoodey K, Mateman AC, ten Cate C, Lessells CM, Clayton DH (2001) A molecular phylogeny of the dove genera *Streptopelia* and *Columba*. *Auk* 118:874–887
- Jordaens K, Van Riel P, Martins AMF, Backeljau T (2009) Speciation on the Azores islands: congruent patterns in shell morphology, genital anatomy, and molecular markers in endemic land snails (Gastropoda, Leptaxinae). *Biol J Linn Soc* 97:166–176
- Jordall BH, Hewitt GM (2004) The origin and radiation of Macaronesian beetles breeding in *Euphorbia*: the relative importance of multiple data partitions and population sampling. *Syst Biol* 53:711–734
- Juan C, Emerson BC, Oromí P, Hewitt GM (2000) Colonization and diversification: towards a phylogenetic synthesis for the Canary Islands. *Trends Ecol Evol* 15:104–109
- Keogh JS, Edwards DL, Fisher RN, Harlow PS (2008) Molecular and morphological analysis of the critically endangered Fijian iguanas reveals cryptic diversity and a complex biogeographic history. *Philos Trans R Soc Lond B* 363:3413–3426
- Kim S-C, McGowen MR, Lubinsky P, Barber JC, Mort ME, Santos-Guerra A (2008) Timing and tempo of early and successive adaptive radiations in Macaronesia. *PLoS ONE* 3:e2139
- Kocher TD, Thomas WK, Meyer K, Edwards SV, Pääbo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc Natl Acad Sci* 86:6196–6200
- Marrero P (2009) Ecología trófica de las palomas endémicas de las Islas Canarias. PhD dissertation, University of La Laguna, Tenerife
- Marrero P, Oliveira P, Nogales M (2004) Diet of the endemic Madeira Laurel Pigeon *Columba trocaz* in agricultural and forest areas: implications for conservation. *Bird Conserv Int* 14:165–172
- Martín A, Hernández MA, Lorenzo JA, Nogales M, González C (2000) Las palomas endémicas de Canarias. Consejería de Política Territorial y Medio Ambiente del Gobierno de Canarias & SEO/BirdLife. Santa Cruz de Tenerife
- Meik JM, Lawing AM, Pires-daSilva A (2010) Body Size Evolution in Insular Speckled Rattlesnakes (Viperidae: *Crotalus mitchelli*). *PLoS ONE* 5:e9524
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403:853–858
- Nogales M, Marrero P, Padilla DP, Medina FM (2009) The study of habitat use by censuses and molecular methods in birds: the case of two sympatric pigeons. *Bird Study* 56:414–418
- Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Sweden
- Nylander JAA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey JL (2004) Bayesian phylogenetic analysis of combined data. *Syst Biol* 53:47–67
- Oliveira P, Marrero P, Nogales M (2002) Diet of the endemic Madeira laurel pigeon and fruit resource availability: a study using microhistological analyses. *Condor* 104:811–822
- Oliveira P, Menezes D, Jones M, Nogales M (2006) The influence of fruit abundance on the use of forest and cultivated field habitats by the endemic Madeira laurel pigeon *Columba trocaz*: Implications for conservation. *Biol Conserv* 130:538–548
- Oromí P (2004) Biospeleology in Macaronesia. *AMCS Bulletin* 19/SMES Boletín 7:98–104
- Pina-Martins F, Paulo OS (2008) Concatenator: sequence data matrices handling made easy. *Mol Ecol Resour* 8:1254–1255
- Piñeiro R, Aguilar JF, de Sequeira MM, Feliner GN (2009) Low genetic diversity in the rare Madeiran endemic *Armeria madeirensis* (Plumbaginaceae). *Folia Geobot* 44:65–81
- Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817–818
- Prychitko TM, Moore WS (1997) The utility of DNA sequences of an intron from the beta-fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: Picidae). *Mol Phylogenet Evol* 8:193–204
- Rambaut A, Drummond AJ (2007) Tracer v1.4. <http://beast.bio.ed.ac.uk/Tracer>
- Ribera I, Bilton DT, Vogler AP (2003) Mitochondrial DNA phylogeography and population history of *Meladema* diving beetles on the Atlantic Islands and in the Mediterranean basin (Coleoptera, Dytiscidae). *Mol Ecol* 12:153–167

- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574
- Rubinoff D (2008) Phylogeography and ecology of an endemic radiation of Hawaiian aquatic case-bearing moths (*Hyposmocoma*: Cosmopterigidae). *Philos Trans R Soc Lond B* 363:3459–3465
- Russell GJ, Diamond JM, Reed TM, Pimm SL (2006) Breeding birds on small islands: island biogeography or optimal foraging? *J Anim Ecol* 75:324–339
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory Press, New York
- Sá-Pinto A, Branco M, Sayanda D, Alexandrino P (2008) Patterns of colonization, evolution and gene flow in species of the genus *Patella* in the Macaronesian Islands. *Mol Ecol* 17:519–532
- Schmitz P, Cibois A, Landry B (2008) Cryptic differentiation in the endemic micromoth *Galagete darwini* (Lepidoptera, Autostichidae) on Galápagos volcanoes. *Philos Trans R Soc Lond B* 363:3453–3458
- Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16:1114–1116
- Sorenson MD, Ast JC, Dimcheff DE, Yuri T, Mindell DP (1999) Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Mol Phylogenet Evol* 12:105–114
- Sota T, Nagata N (2008) Diversification in a fluctuating island setting: rapid radiation of *Ohomopterus* ground beetles in the Japanese Islands. *Philos Trans R Soc Lond B* 363:3377–3390
- Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *Am J Hum Genet* 76:449–462
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68:978–989
- Tracey CM, Pimm SL (2009) Testing island biogeography theory with visitation rates of birds to British islands. *J Biogeogr* 36:1532–1539
- Stresemann E (1927–1934) Aves. In: Kükenthal W, Krumbach T (eds) *Handbuch der Zoologie*. Walter de Gruyter, Berlin
- Su YCF, Saunders RMK (2009) Evolutionary divergence times in the Annonaceae: evidence of a late Miocene origin of *Pseuduvaria* in Sundaland with subsequent diversification in New Guinea. *BMC Evol Biol* 9:153
- Swofford DL (2002) PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4.0b10. Sinauer Associates, Sunderland
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876–4882
- Thorpe RS, Surget-Groba YS, Johansson H (2010) Genetic tests for ecological and allopatric speciation in anoles on an island archipelago. *PLoS Genet* 6:e1000929
- Van Riel P, Jordaens K, Van Houtte N, Martins AMF, Verhagen R, Backeljau T (2005) Molecular systematics of the endemic Leptaxini (Gastropoda: Pulmonata) on the Azores islands. *Mol Phylogenet Evol* 37:132–143
- Vargas P, Morton CM, Jury SL (1999) Biogeographic patterns in Mediterranean and Macaronesian species of *Saxifraga* (Saxifragaceae) inferred from phylogenetic analyses of ITS sequences. *Am J Bot* 86:724–734
- Villacorta C, Jaume D, Oromí P, Juan C (2008) Under the volcano: phylogeography and evolution of the cave-dwelling *Palmorchestia hypogaea* (Amphipoda, Crustacea) at La Palma (Canary Islands). *BMC Biol* 6:7
- Warren BH, Bermingham E, Bowie RCK, Prys-Jones RP, Thébaude C (2003) Molecular phylogeography reveals island colonization history and diversification of western Indian Ocean sunbirds (Nectarinia: Nectariniidae). *Mol Phylogenet Evol* 29:67–85
- Whittaker RJ, Fernández-Palacios JM (2007) *Island biogeography: ecology, evolution, and conservation*, 2nd edn. Oxford University Press, Oxford