CONSERVATION GENETICS OF THE ENDEMIC PIGEONS OF SÃO TOMÉ AND PRÍNCIPE

Hugo José Eira Pereira

Mestrado em Biodiversidade, Genética e Evolução

CIBIO (Centro de Investigação em Biodiversidade e Recursos Genéticos), Universidade do Porto

2013

Orientador

Martim Melo, Post-doc, CIBIO (Centro de Investigação em Biodiversidade e Recursos Genéticos), Universidade do Porto

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

CIBIO
Centro de Investigação em Biodiversidade e Recursos Genéticos

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.
O Presidente do Júri,

Porto, _____/_____/_____

ACKNOWLEDGMENTS

I would like to thank all the people in CTM for the precious help during the laboratory work. A special thanks to Diana Castro and Sandra Reis my "go to" people in the lab, and of course, to the "almighty" Susana Lopes, for always giving precious advices when nothing seemed to work.

To all the researchers in CIBIO who took the time to talk with me, and were always available to clarify my doubts . A special thanks to José Melo Ferreira for losing precious time giving me a lesson on how to work with BEAST.

To Martin Stervander for always answering my questions, no matter how dumb they might be. But most of all for that advice, two years ago, to enrol in the master course in CIBIO and for putting me in contact with Martim.

To the amazing master class of 2011/2013. Sharing the last two years with you all was a pleasure. We learned a lot, we partied a lot and we panicked together while writing this thesis during the last days .

Dr. Kristof Zyskowski (Peabody Museum of Natural History), thanks for lending the *C. iriditorques* and *C. delegorguei* samples, without even asking for anything in return.

Jorge Palmeirim, Mariana Carvalho, Fábio Sequeira and Christoff Myer (the members of the "pigeons group"), thanks for the meeting in Lisbon. Having your feedback about the experience in São Tomé was really important. Thanks for all the pictures, presentations and reports that you send me. Including part of your work in this thesis really made it richer.

To Kevin Johnson, an unexpected collaboration. Having the opportunity to work and learn with you was an honour. Thanks for answering all my questions (and there we a lot of them) and for taking the time to review the first part of this thesis. Although not official, you truly were my co-supervisor.

And last to Martim Melo, for the opportunity to work with you and to work with such an amazing and unique system as the Gulf of Guinea islands. It was a pleasure and an honour. Thank you for answering to a ridiculous amount of e-mails and for always being available to have a meeting and discuss some ideas. Your commitment to this thesis during the final days was astonishing. Despite being in separate countries for half of the year, I think we made this partnership work. For all this, thank you very much, you were an amazing supervisor.

The work of this thesis was founded by the FCT (Fundação para a Ciência e Tecnologia) as part of the "Endemic pigeons of São Tomé: developing science-based conservation and sustainable use of African forest pigeons" project (PTDC/BIA-BIC/115223/2009).

PHOTO CREDITS

The credits for all photography's presented in this thesis goes to the following:

Adam Ryley

César Garcia

H. Robertson

Josep del Hoyo

João Pedro Pio

Megan Perkins

Nick Borrow

Peter van Zoest

ABSTRACT

The Gulf of Guinea islands are known for their amazing levels of endemism. Endemism levels are particularly high on São Tomé and Príncipe, with birds being one of the groups for which these levels are astonishing, by far the highest in the world relatively to surface area.

Among the endemic birds of São Tomé and Príncipe there are three forest pigeon species: the São Tomé olive pigeon *Columba thomensis* (São Tomé), the São Tomé bronze-naped pigeon *Columba malherbii* (São Tomé Príncipe and Annobón), and the São Tomé green pigeon *Treron sanctithomae* (São Tomé). The African lemon dove (*Aplopelia larvata*) occurs on the three islands and its taxonomy remains uncertain.

These species (with the exception of the lemon dove) suffer heavily from hunting pressures. While pigeon hunting is widespread in São Tomé it seems absent from Príncipe. Of particular concern is the fact that these species are among the most hunted species of São Tomé.

This thesis is part of a larger conservation project whose main objective is to gather the necessary information to manage the hunting of the endemic pigeons of São Tomé, and to develop techniques for the control of the "bushmeat trade" problem in Africa.

This thesis has two main objectives. First, molecular tools were used to clarify the evolutionary history of the pigeons of the Gulf of Guinea islands, and in particular of the endemic species affected by hunting. Secondly, population genetic data was used to determine levels of gene flow between conspecific or closely related pigeons populations of São Tomé and Príncipe islands, to estimate patterns of past demographic events and to assess current levels of neutral genetic diversity. Mitochondrial and nuclear sequences were used to infer the phylogenetic relationships and degree of divergence between the endemic pigeons and the mainland species. Additionally, the phylogenetic position of the enigmatic lemon dove and of the bronze-naped pigeons complex was investigated. Levels of genetic divergence supported the species status for the *Columba* endemics, which diverged from the mainland species from around 1.3 to 2.0 Myr ago. The same pattern was observed for *T. sanctithomae* but, due to insufficient sampling of mainland populations, no definitive conclusion could be drawn. Mainland *A. larvata*, and the island forms diverged around 2.4 Myr ago, suggesting that the São Tomé and Príncipe populations should be treated as a new endemic species. Nevertheless, as with *Treron*, incomplete sampling of mainland populations prevented definitive conclusions to be made. The results supported the placement of *Aplopelia* and the bronze-naped

pigeons in a clade comprising the Old World *Columba* and most of the *Streptopelia* species; the exact placement in relation to these two groups remained uncertain.

A population genetics approach was used for the three endemic pigeons populations with the objective of estimating genetic parameters such as nuclear diversity, effective population size, gene flow and past events of demographic changes. Results identified stable populations with weak signals of expansion, except in the case *C. thomensis*. The contrast between the current census efforts describing population declines and the large and stable population sizes inferred by genetic data supports the view that the populations may be undergoing a human driven bottleneck due mainly to hunting pressures. Nevertheless the moderate levels of genetic diversity recovered, suggests that the populations may still retain the capacity to recover in numbers. For this to be possible, current hunting levels must be curbed urgently. The *C. thomensis* case is of more concern, as its substantially lower genetic diversity is indicative of a small effective population size. Regarding the *Treron* group, the population genetics results confirmed a clear separation between *T. sanctithomae* and *T. calva* from Príncipe, supporting their distinct species status. Interestingly some rare exchanges of genetic information between the two species were identified making their evolutionary history even more intriguing.

RESUMO

As ilhas do Golfo da Guiné são conhecidas por terem níveis de endemismo fantásticos. Estes níveis são especialmente elevados nas ilhas de São Tomé e Príncipe. Para as aves, estes níveis são elevadíssimos, sendo São Tomé e Príncipe as ilhas com o maior número de aves endémicas do mundo em relação à sua área.

Entre a comunidade de aves endémicas de São Tomé e Príncipe encontram-se três pombos: o Pombo-do-mato *Columba thomensis* (São Tomé), a Rola *Columba malherbii* (São Tomé, Príncipe e Annobón), e a Céssia de São Tomé *Treron sanctithomae* (São Tomé). A Mucanha (*Aplopelia larvata*) existe nas três ilhas, mas a sua taxonomia é ainda duvidosa.

Estas espécies (com a excepção da *A. larvata*) são intensamente caçadas. A caça está fortemente implementada em São Tomé, enquanto que na ilha de Príncipe tal parece não ocorrer. Em São Tomé os pombos endémicos encontram-se entre as espécies cinegéticas mais procuradas o que constitui um motivo de preocupação.

Esta tese é parte de um projeto de conservação que tem como objetivo gerir a caça dos pombos endémicos de São Tomé e, ao mesmo tempo, desenvolver ferramentas para controlar o problema do comércio de carne de animais selvagens em África. Assenta em dois grandes objetivos: o primeiro visa perceber a história evolutiva dos pombos das ilhas do Golfo da Guiné e, em particular, dos pombos endémicos afetados pela caça, usando técnicas de genética molecular; o segundo, pretende inferir parâmetros genéticos que ajudem a clarificar o estatuto das populações do ponto de vista da sua conservação e determinar se existe fluxo genético entre populações da mesma espécie, ou de espécies próximas, entre as ilhas de São Tomé e Príncipe. Foram utilizadas sequências mitocondriais e nucleares para inferir as relações filogenéticas e o grau de diferenciação entre os pombos endémicos e as espécies continentais. Foi ainda estudada a posição filogenética dos enigmáticos M*ucanha* e pombos-de-nuca-bronzeada.

Os níveis de divergência genética reforçam o estatuto de espécie dos *Columba malherbii* e *thomensis*, tendo estes divergido das espécies continentais entre 1.3 a 2.0 milhões de anos. Resultados similares foram obtidos para o *T. sanctithomae,* mas devido à incompleta amostragem das populações do continente, não se podem tirar conclusões definitivas. Os indivíduos insulares de *A. larvata* divergiram dos indivíduos continentais há cerca de 2.4 milhões de anos, o que sugere que as populações de São Tomé e Príncipe possam ser uma nova espécie endémica. No entanto, também devido a uma amostragem incompleta da distribuição continental da espécie, não se podem

tirar conclusões definitivas. As *Aplopelia* e o complexo dos pombos-de-nuca-bronzeada estão inseridos dentro do grande clado que compreende o género *Columba* e a maior parte das espécies do género *Streptopelia*. Contudo, a exata relação filogenética entre estes grupos fica ainda por descobrir.

Um marcador mitocondrial e um nuclear foram utilizados para o estudo da genética populacional das três espécies de pombos endémicos caçados. Este estudo determinou variáveis genéticas relevantes para a conservação das espécies, incluindo a inferência de fenómenos demográficos de expansão/contração das populações. Os resultados obtidos mostram populações estáveis com alguns sinais de expansão, exceto no caso do *C. thomensis.* O contraste entre os resultados obtidos e os números observados no campo, sugerem que as populações podem estar a sofrer de um *"bottleneck"* recente devido principalmente à ação humana e tendo a caça como principal causa. Por outro lado os níveis moderados de diversidade genética sugerem que as populações ainda retêm capacidade de recuperar os seus números. No entanto, sem esforços de conservação isto parece um cenário bastante improvável. O caso do *C. thomensis* é mais preocupante, devido aos baixos níveis de diversidade genética e efetivos populacionais. Relativamente ao grupo dos *Treron*, os resultados mostraram uma clara separação entre as populações de São Tomé e do Príncipe, dando suporte ao seu estatuto atual de espécies distintas. Curiosamente foram identificadas raras trocas de informação genética entre as duas espécies o que torna a história evolutiva destas espécies particularmente interessante.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

THE GULF OF GUINEA ISLANDS

The Gulf of Guinea Islands system constitutes a unique centre of endemism, which only recently has received the attention it deserves (Gascoigne 2004). It encompasses three oceanic islands (Príncipe, São Tomé, and Annobón), one land-bridge island (Bioko), and one ecological island (Mount Cameroon); these islands are part of the Cameroon Line of Tertiary to Recent volcanoes (Fig.1.).

It is estimated that Príncipe arose c. 31 million years ago, São Tomé at least 13 million years, and Annobón at least 5 million years ago (Lee *et al*. 1994). Bioko was repeatedly connected to the mainland during glacial times and last separated 11,000 years ago (Rohling *et al*. 1998; Lambert *et al*. 2001).

One interesting feature of the oceanic islands is the similar distances between islands and between the islands and the mainland (Fig. 1).

The original habitat of the Gulf of Guinea islands was rainforest or moist tropical forest (Excell 1973; Gascoigne 2004). The forest is stratified into lowland forest (0-800m), montane forest (800-1400), and mist forest (1400-2500; Excell 1944).

SÃO TOMÉ AND PRÍNCIPE: A HOTSPOT FOR BIRD ENDEMISM

The Gulf of Guinea islands are known for the truly amazing levels of endemism of their fauna and flora. The linear arrangement of the islands and the distances between islands, and the islands and the mainland, created both the perfect scenario for multiple colonisations from the highly diverse Guineo-Congolian forests and for the occurrence of dispersal events between islands (Jones & Tye 2006, Melo 2007).

Endemism levels are particularly high on São Tomé and Príncipe, the two largest oceanic islands. Here birds are one of the groups for which the levels of endemism are astonishing – by far the highest in the world relatively to the island area.

These two islands hold 28 out of the 33 endemic birds found on the Gulf of Guinea system. Príncipe, with only 139 km^2, is home to 8 single island endemics out of 33 breeding land birds. São Tomé (857km^2) supports 17 single island endemics. Two further species are endemic to both Príncipe and São Tomé, and the a third one is shared also with Annobón. Additionally, 13 African mainland species have endemic

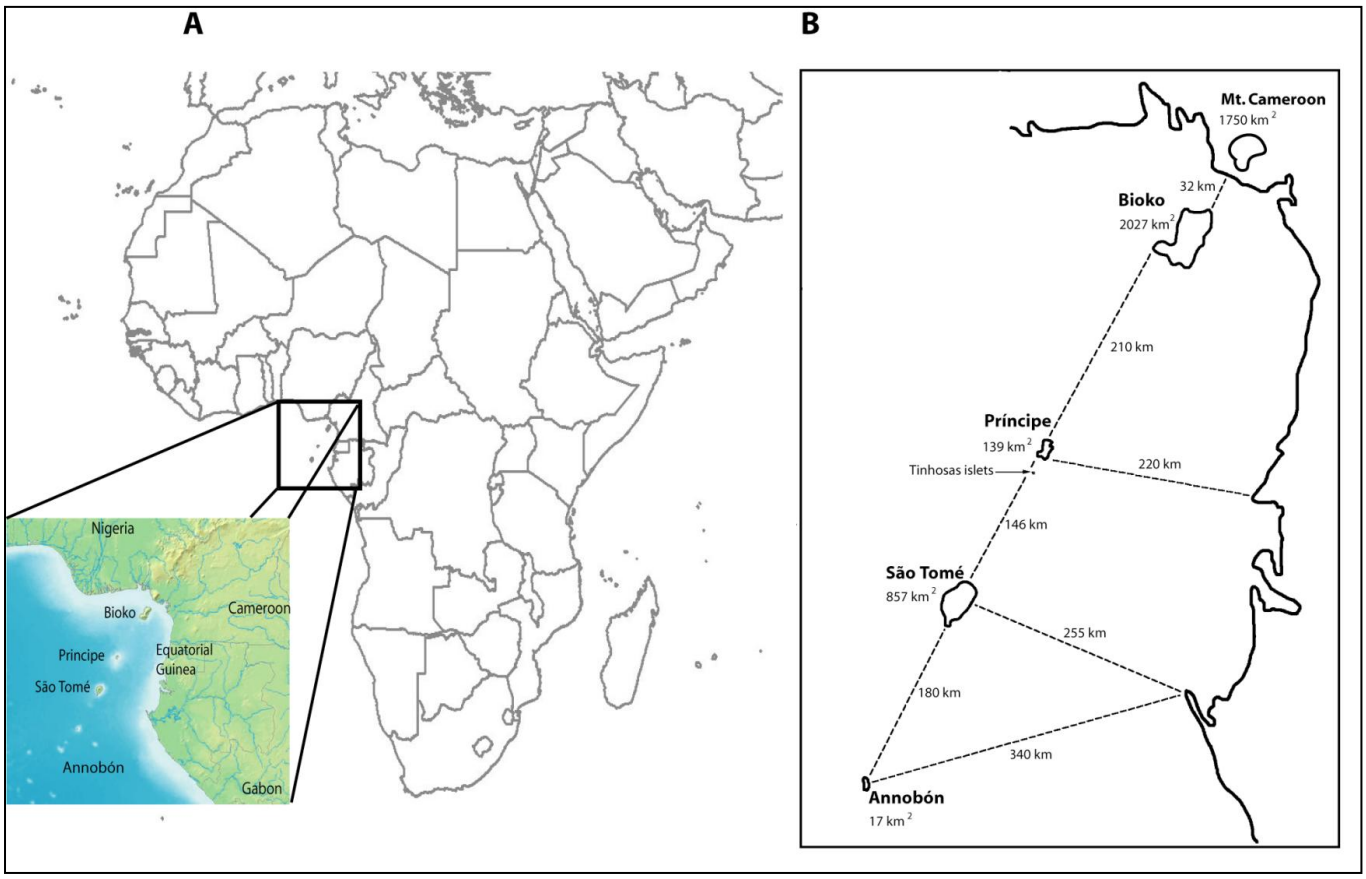


Figure 1. Map of Gulf of Guinea (A);(B) The Gulf of Guinea island system, showing island areas and distances between them and the mainland (adapted from Melo & Jones 2008).

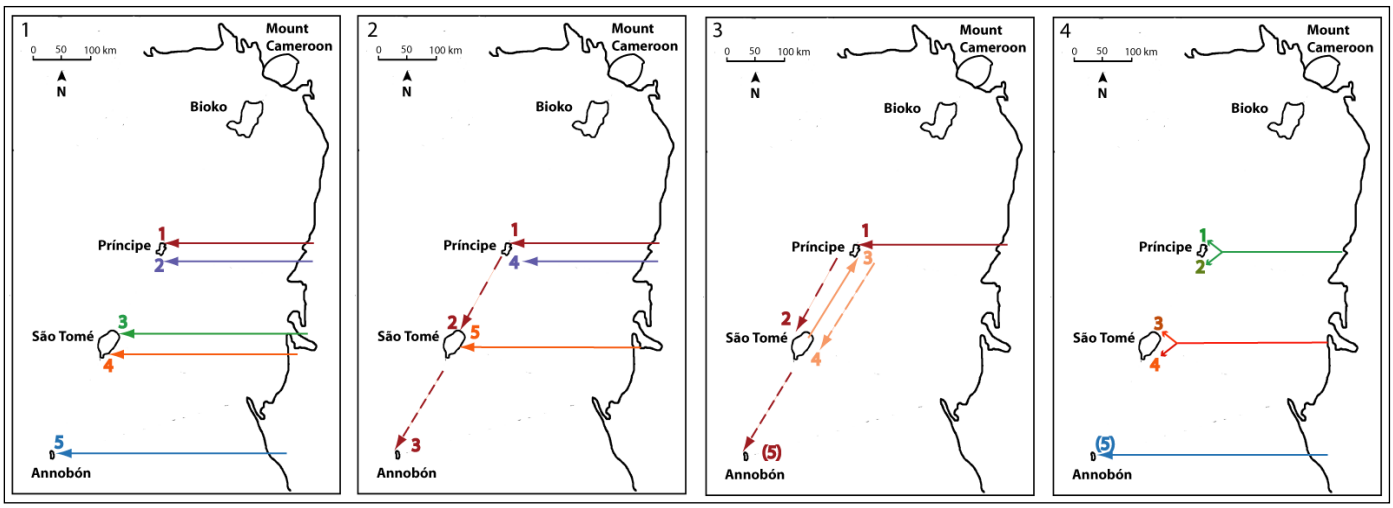


Figure 2. Models of speciation for endemic species in an oceanic archipelago. **1.** Allospeciation (speciation with isolation) with the ancestor being a colonizer from the mainland; **2.** Allospeciation with the ancestor being a colonizer from another island islands. **3.** Archipelago radiation (island-dispersal events promote contact of allopatric diverging populations, ecological interactions between them (competition) can drive further phenotypic leading to reproductive isolation); **4.** Sympatric speciation (process through which new species evolve from a single ancestral species while inhabiting the same geographic region; adapted from Melo & Jones 2008)

subspecies on the São Tomé and Príncipe islands (Jones & Tye 2006; Melo 2007; Melo & Jones 2008).

These high levels of endemism have led to the classification of São Tomé and Príncipe (and Annobón) as independent Endemic Birds Areas (EBAs) amongst the 218 worldwide (Bibby *et al.* 1992). Even more strikingly, considering their small size, São Tomé and Príncipe were rated high in the top 25%, for their species richness. More significantly, the forests of these three oceanic islands were rated the third most important forests in the world from a bird conservation perspective (Buchanan *et al*. 2011).

Most recent studies using molecular data concluded that these high endemism levels were mostly achieved by multiple colonizations from the mainland followed by speciation in allopatry (Melo 2007; model 1/2 - Fig.2). The endemics of the islands belong to 17 families present on the African mainland and therefore at least 17 colonization events took place. Inter-island dispersal events seem to account for the origin of 5 to 8 species (Melo 2007). Only two cases of small archipelago radiations (model 3 - Fig.2) took place: one leading to the five white-eyes (*Zosterops*; Melo *et al.* 2011) and the other to the two seedeaters (*Serinus rufobrunneus* and *Neospiza concolor*). This latter case is of particular interest as molecular patterns are strongly supportive of sympatric speciation (model 4 - Fig.2; Melo 2007).

Molecular studies have also shown that most speciation events in São Tomé and Príncipe are recent, from the Pliocene-Pleistocene boundary (≅ 2.5 Ma) onwards (Melo 2007). The fact that the present species are much more recent than the islands they inhabit supports the view that the islands are areas of "real time speciation" rather than just acting as a refuge for mainland species that went extinct.

Conservation issues

In isolated islands, the human-made habitat changes have a strong impact on the overall levels of species diversity. Hence, to understand the conservation issues for most species, one should first analyze the changes in the vegetation since the colonization by humans.

The lower rainforest region (sea-level to 800 m) has been almost entirely cultivated, with most clearance done shortly after human colonization, in 1470, for the plantation of sugar cane. Large extensions of the montane forest region (800 m to 1400 m) have been converted to shade plantations of coffee and cocoa by the late 18th century. The mist forest region (1400 m to 2024 m) has been little affected by human intervention (Jones & Tye 2006).

The full impact that the destruction of the lowland forest had cannot be ascertained due to lack of inventories at the time of colonisation. Nevertheless, no extinctions of endemic species have been documented (Melo 2007).

Most of the endemic bird species are primary forest specialists, but many are present in secondary forest and in shade plantations, with a few also present in heavily impacted areas, including urban centers (Atkinson *et al.* 1994; Peet *et al.* 1994, Jones & Tye 2006; de Lima *et al.* 2013).

Some conservation measures have been implemented, the most significant of which being the creation of large protected areas in each island. Since 2006, 30% of the area of São Tomé (235 km²) and almost 50% of Príncipe (65 km²) are natural parks ('Parque Natural D'Obo'). The protected areas cover all biotopes, including the lowland, montane and mist forests, mangroves and an area of savanna (in Obopark.com). In 2012, the island of Príncipe was awarded the title of Biosphere Reserve by the UNESCO Man and Biosphere (MAB) Program, the first case in the African continent (in www.unesco.org).

HUNTING AND THE ENDEMIC PIGEONS OF SÃO TOMÉ AND PRÍNCIPE

Together with habitat destruction, hunting is another major threat to the endemic birds, with pigeons being the most sought after bird group. Although a number of studies have been made concerning the impact of habitat alteration/loss in the bird communities (Dallimer 2008, 2009, de Lima *et al.* 2013), little or nothing is known about the effect of hunting pressures on the endemic birds of São Tomé and Príncipe.

The endemic bird community of the islands of São Tomé and Príncipe includes three forest pigeon species: the São Tomé olive pigeon *Columba thomensis* (São Tomé), the São Tomé bronze-naped pigeon *Columba malherbii* (São Tomé, Príncipe and Annobón), and the São Tomé green pigeon *Treron sanctithomae* (São Tomé). The African lemon dove (*Aplopelia larvata*) occurs on the three oceanic islands and its taxonomy remains uncertain, with the population on São Tomé sometimes treated as a distinct species (del Hoyo *et al.*1997). On Príncipe, there is an endemic subspecies of the African green pigeon (*Treron calva virescens*). Pigeon hunting is widespread on the island of São Tomé, but is absent from Príncipe – whereas no information is available for Annobón.

On São Tomé, pigeons are heavily hunted, being amongst the most hunted species on the island (Carvalho, M. pers. com.), except for the lemon dove that is considered unsuitable for consumption.

Pigeons are a readily available resource, making pigeon hunting an attractive economic activity in the rural communities. Pigeons are also appealing from a consumer

perspective since they are cheap (on the local market *C. malherbii* and *T. sanctithomae* costs about 1.00€ and *C. thomensis* 2.00€; Sequeira, F. pers. com.). Although the endemic pigeons are secondary and primary forest specialists, they can be observed in any habitat with dense tree cover since they seek fruiting trees to feed on. These include plantations close to human settlements. Furthermore, these species are both very gregarious and naïve: they are not afraid of human contact and birds do not fly away even when being shot at, which makes them an easy prey.

Due mainly to the hunting pressure, *C. malherbii* is classified as a “Near Threatened” species. There are no data for the population size and hunting uptake but the general trend appears to be of a population decline. *Treron sanctithomae* is classified as “Vulnerable”. Its population size is roughly estimated to be below 10,000 mature individuals and a decline of 39-49% is suspected. The most severe case is that of *C. thomensis*. Although its population size has not been quantified it is probably best placed around 250-999 mature individuals considering its scarcity, highly restricted distribution (mostly primary forest) and a suspected population decline of 30-49 %. The species is classified as “Endangered” (IUCN 2012; Bird Life International 2012).

The lack of solid data and the empirical acknowledgment that the situation of the São Tomé endemic pigeons was becoming distressing were the driving force for launching, in 2011, a study with the purpose of assessing the dimension of the problem and of developing conservation measures for managing hunting in São Tomé.

The first results support the view that hunting is indeed affecting the endemic pigeons populations. Pigeon hunting is an important source of income for the communities living close to the natural park. In recent years, hunting pigeons for sport has further increased the hunting pressure. Preliminary data alarmingly shows that the highest levels of hunting take place during the breeding season. This is expected to lead to low rates of reproductive success as appears to be corroborated by the low number of juveniles recorded. The situation of *C. thomensis* is particularly worrying. In three months and over six observation hours/day, only 30 individuals were recorded. This species, the largest, has always been regarded as a highly valued target, and its increasing rarity is now further increasing its value. (Carvalho, M. and Sequeira, F. pers. com.).

The endemic pigeons are the largest frugivorous birds of São Tomé, and are likely to play an important role as seed dispersers of the native trees in the island. Their decline may therefore have a major impact on the maintenance of its forests (McConkey *et al.* 2002; Walker 2007). Besides their unique value from a biodiversity perspective, the forests are crucial for the economy of the country, as they act as water reservoirs and provide timber for construction and fuel.

Strict conservation measures, like the interdiction of hunting, cannot be implemented both because of the conflict they would generate and because they would be impossible to enforce. Careful planning of the hunting activity, based on solid data, is the road ahead.

OBJECTIVES

This thesis is part of a larger conservation project whose main objective is to gather the information necessary to manage the hunting of the endemic pigeons of São Tomé, while providing appropriate tools to manage the "bushmeat trade" problem elsewhere in Africa.

The main aim of this thesis is to use a molecular approach to guide conservation efforts by: i) clarifying the evolutionary history of the pigeons of the Gulf of Guinea islands, and in particular of the endemic species affected by hunting, and ii) using a population genetics approach to infer the vulnerability of the populations to hunting.

The degree of genetic differentiation can be used to infer the degree of evolutionary independence from which one can assess the conservation significance of a given population. Although *C. thomensis* and *T. sanctithomae* are currently considered single-island endemics and *C. malherbii* an endemic shared with Príncipe and Annobón islands, some authors consider that they should be treated as insular populations of mainland species (Dowsett & Dowsett-Lemaire 1993). Molecular markers independent of phenotypic traits are a widely used tool to clarify such issues, and will be used, for the first time, to infer the evolutionary history of the Gulf of Guinea pigeons.

Patterns of genetic variation within populations and differentiation between populations hold signatures of demographic events of the recent past, including population expansions/contractions, and levels and pathways of gene flow between geographically separated populations. Effective population size, genetic drift, levels of gene flow and levels of inbreeding are variables that can be used as a proxy for estimating the adaptive potential of a species, which are important factors to be taken into account when designing conservation projects (Allendorf *et al.* 2007). This study will aim to retrieve this information with the main objective of guiding future conservation action.

To accomplish this, two main studies will be conducted:

1. Inference of the phylogenetic position of the São Tomé and Príncipe endemic pigeons: genetic differentiation between insular and island endemics species.

2. Inference of population genetics variables, past demographic events and patterns of genetic differentiation within the São Tomé and Príncipe endemic pigeons species.

CHAPTER 1.

PHYLOGENETIC POSITION OF THE SÃO TOMÉ AND PRÍNCIPE ENDEMIC PIGEONS AND THE ENIGMATIC LEMON DOVE AND BRONZE-NAPED PIGEONS

ABSTRACT

The islands of São Tomé and Príncipe are home to an extraordinary number of avian endemics. Among these endemics are three forest pigeons: *Columba thomensis, Columba malherbii* and *Treron sanctithomae*. Although they all are currently considered to be distinct species, endemic to these islands, this status is still a matter of controversy. In addition to these three endemic species, the African lemon dove (*Aplopelia larvata*) occurs on both islands and on Annobón Island, with the populations often treated as distinct subspecies. The form occurring on São Tomé is sometimes treated as a distinct species (*Aplopelia simplex*).

In this study, three mitochondrial genes and a nuclear intron were analysed to clarify the phylogenetic relationships and degree of genetic divergence between the three endemic pigeons and the mainland sister-species (*C. thomensis/arquatrix; C. malherbii/iriditorques* and *delegorguei*; *T. sanctithomae*/*calva*). This study also investigated the still uncertain phylogenetic position of the enigmatics *Aplopelia* and bronze-naped pigeons (*Columba malherbii* group).

Moderate levels of genetic divergence supported the species status for both of the *Columba* endemics, with island endemics diverging from the mainland species around 1.3 to 2.0 Myr ago. Similar results were found for *T. sanctithomae* although, due to

insufficient sampling from the mainland species range, no strong conclusions could be drawn. Genetic differences found between mainland *A. larvata*, and the island forms were even greater, dating back to around 2.4Myr ago. This suggests that the São Tomé and Príncipe populations should be treated as a new endemic species, although denser sampling across the mainland range is also required to confirm this result. Phylogenetic results supported the placement of *Aplopelia* and the bronze-naped pigeons in a clade together with Old World *Columba* and the bulk of *Streptopelia* species (a genus that was not monophyletic). Although support for the exact placement of these two lineages within that broader clade was not high, all phylogenetic analyses from the combined data set placed them together in a clade sister to other Old World *Columba*.

Keywords: *Columba thomensis, Columba malherbii*, *Treron sanctithomae, Aplopelia larvata,* endemics, phylogeny, genetic divergence.

INTRODUCTION

The Gulf of Guinea islands of São Tomé and Príncipe provide an excellent model for studying evolutionary divergence of island taxa. The moderate distances separating the islands from the highly diverse Guineo-Congolian rainforest along with the unique habitat conditions on the islands provided the opportunity for mainland species to colonize the island, thus starting the process of speciation (Jones & Tye 2006; Melo 2007; Melo & Jones 2008).

São Tomé and Príncipe, with 28 endemic birds, are the islands with the highest concentration of endemic birds in the world (Melo 2009). In particular, because of the high capacity for overwater dispersal of pigeons and doves, São Tomé and Príncipe are home to seven pigeon species, with three of them being endemic to the islands. Among the most important sources of wild meat in São Tomé are the three endemic species of forest pigeons: São Tomé olive pigeon (*Columba thomensis),* the São Tomé bronze-naped pigeon (*Columba malherbii)*, and the São Tomé green pigeon (*Treron sanctithomae;* Jones & Tye 2006). The taxonomic status of the subspecies of the African lemon dove (*Aplopelia larvata*) remains uncertain, with most authors recognizing at least subspecies status for the island populations (Gibbs *et al.* 2001). Del Hoyo *et al.* (1997) propose the classification of the São Tomé population as a different species. On São Tomé, pigeons are heavily hunted, whereas hunting is currently negligible on Príncipe.

The 'Endangered' *C. thomensis*, endemic to São Tomé, is considered to be part of a superspecies with the Cameroon Olive Pigeon (*C. sjostedti*), the African Olive Pigeon (*C. arquatrix*) from southern and eastern Africa and north-western Angola and the Comoro Olive-Pigeon (*C. pollenii,* Comoro Islands; del Hoyo *et al.*1997; Jones & Tye 2006). These three species have also been treated as different populations of the same species, *C. arquatrix* (Dowsett & Dowsett-Lemaire. 1993). Morphological differences between these taxa are diagnostic, with clear differences mainly on the head and neck feathers (Fig. 3). The resemblance of *C. thomensis* to *C. arquatrix* rather than to *C. sjostedti* (also present in Bioko), together with its absence from Príncipe favours an independent colonization from the mainland (Naurois 1994).

The 'Near-threatened' *Columba malherbii*, endemic to Príncipe, São Tomé and Annobón, is considered a member of a superspecies with the Bronze-naped Pigeons (*C. delegorguei* and *C. iriditorques)* from mainland Africa (del Hoyo *et al.*1997; Jones & Tye 2006; Fig.4), but the three species are, again, treated as conspecific by Dowsett & Dowsett-Lemaire (1993). Both São Tomé and Príncipe have a *Treron* green pigeon. On

São Tomé it is treated as a single-island endemic species, *Treron sanctithomae*, classified as 'Vulnerable'.

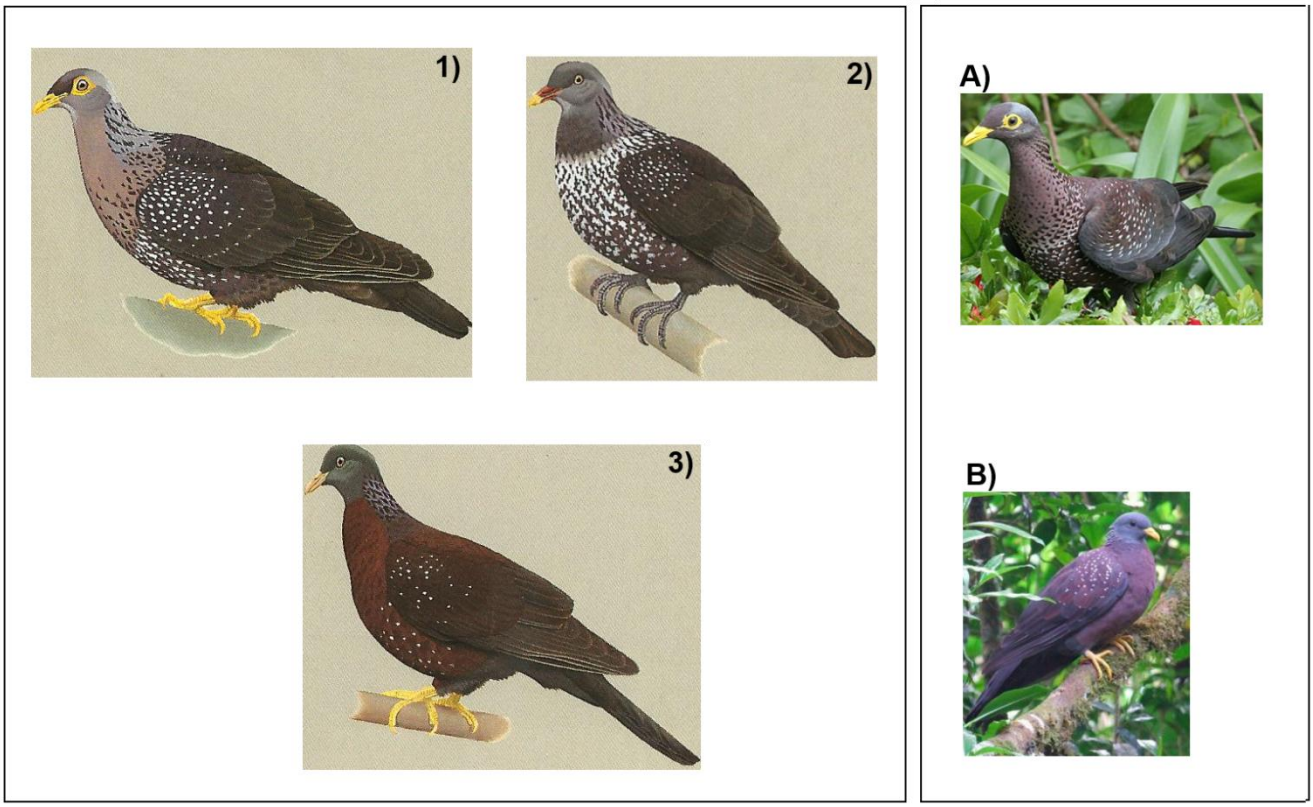


Figure 3. The São Tomé olive pigeon and the mainland closest relatives. On the left illustrations from del Hoyo *el al.* 1997. 1) *Columba arquatrix*; 2*) Columba sjostedti*; 3) *Columba thomensis*.
On the right: A) *Columba arquatrix* (photo by H. Robertson); B) *Columba thomensis* (photo by Cesar Garcia).

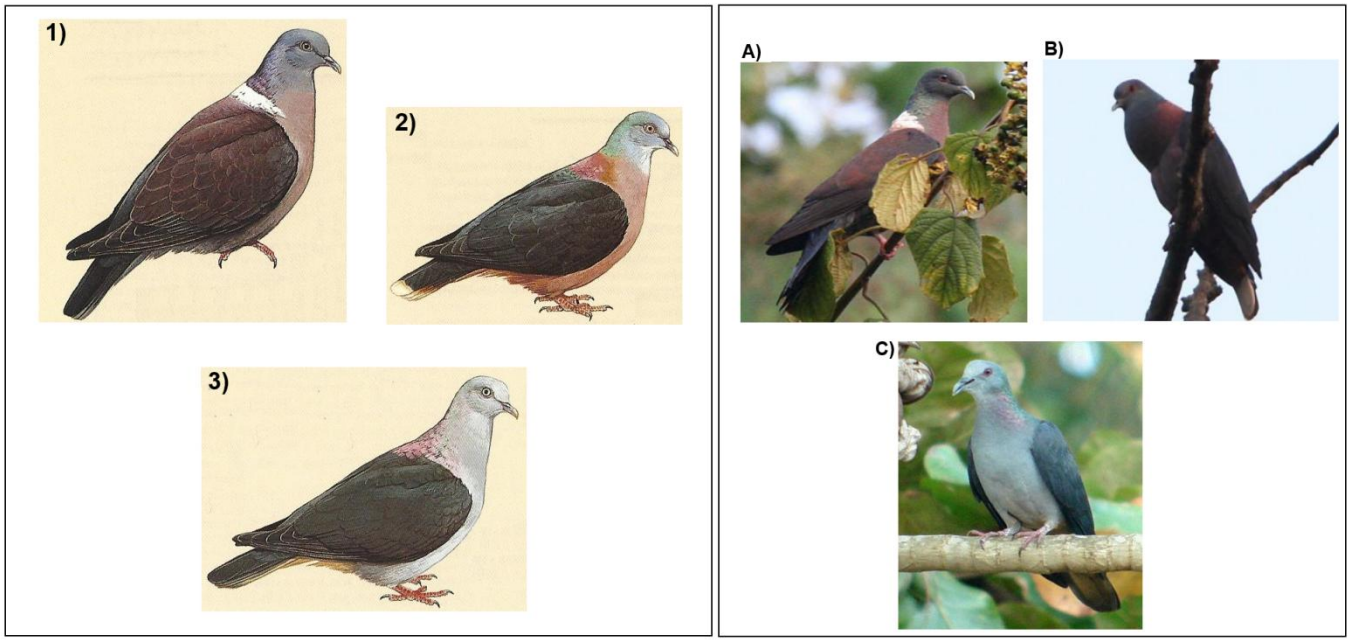


Figure 4. The bronze-napped pigeons group, islands and mainland relatives. On the left illustrations from del Hoyo *el al.* 1997. 1) *Columba delegorguei*; 2) *Columba iriditorques*; 3) *Columba malherbii*. On the right: A) *Columba delegorguei* (photo by Megan Perkins); B) *Columba iriditorques* (photo by Adam Riley); C) *Columba malherbii* (photo by Nick Borrow).

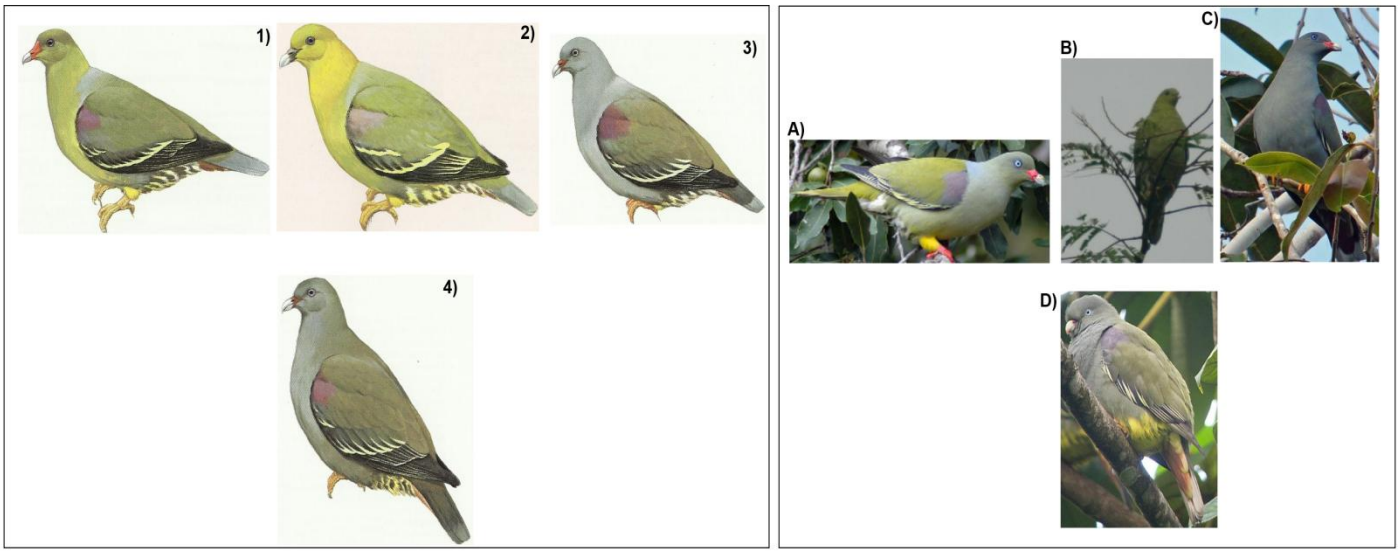


Figure 5. The São Tomé Green Pigeon and the mainland closest relatives. On the left illustrations from del Hoyo *el al.* 1997. 1) *Treron. calva*; 2) *Treron australis*; 3) *Treron pembaensis*; 4) *Treron sanctithomae.*
On the right: A) *Treron calva* (photo by Peter van Zoest*)*; B) *Treron australis*; C) *Treron pembaensis* (photo by Nick Borrow); D) *Treron sanctithomae* (photo by Nick Borrow).

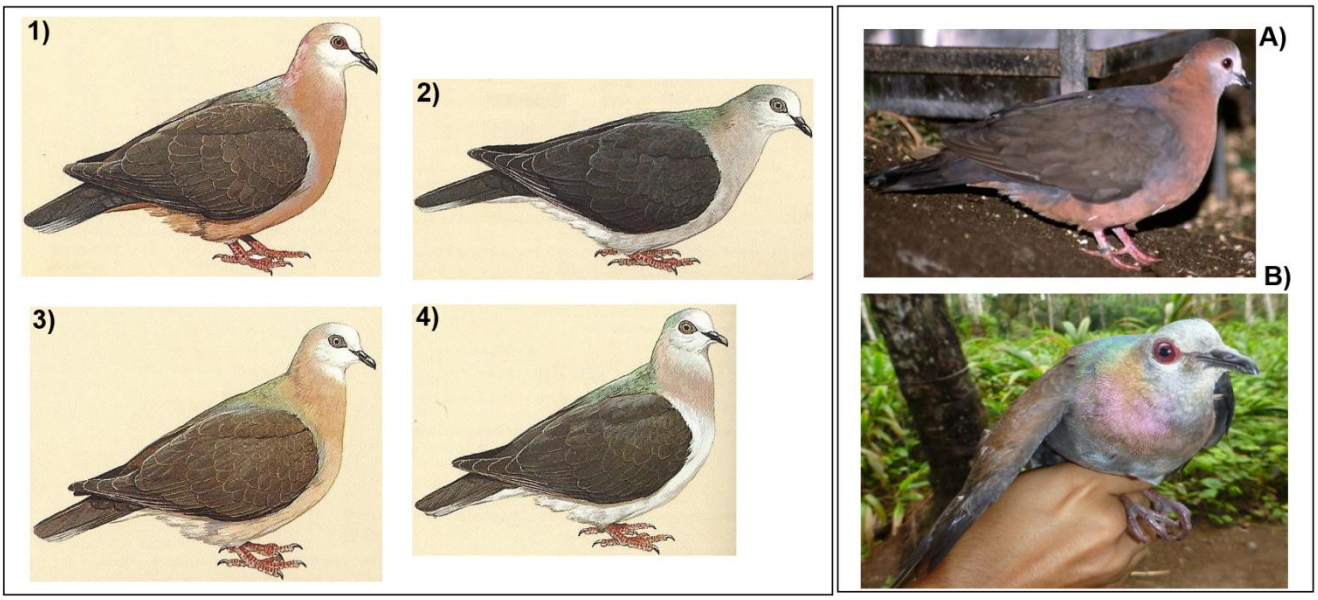


Figure 6. The Lemon Dove group, islands and mainland relatives. On the left illustrations from del Hoyo *el al.* 1997. *1) Aplopelia larvata; 2) Aplopelia larvata hypoleuca;* 3) *Aplopelia larvata principalis*; 4) *Aplopelia larvata simplex* (São Tomé)
On the right: A) *Aplopelia larvata larvata* (photo by Josep del Hoyo); B) *Aplopelia larvata simplex* (São Tomé; photo by João Pedro Pio).

On Príncipe, the green pigeon is treated as an endemic subspecies, *virescens*, of the mainland *T. calva.* With the exception of the Bruce's Green-pigeon (*T. waalia*) the relationships between and within the green pigeons from continental Africa and islands are still unresolved. They seem to form a very closely related group, but species delimitations are still much debated (del Hoyo *et al.* 1997). *T. sanctithomae* forms a superspecies with the Madagascar Green-pigeon (*T. australis*), the Pemba Green Pigeon (*T. pembaensis;* Pemba island), and the African green pigeon (*T. calva*). The population from Príncipe could have originated either from Bioko or directly from the neighbouring mainland. In the same way, *T. sanctithomae* could derive from a stepping-stone colonisation via Bioko and Príncipe or from a direct mainland colonisation (Jones & Tye 2006; del Hoyo 1997; Fig.5).

Unlike the canopy dwelling pigeons described above, the African Lemon Dove, *Aplopelia larvata,* is a moist forest under-storey specialist with a fragmented distribution across the continent. The relationships between its different populations remain poorly understood. Five mainland subspecies have been described. Populations of *Aplopelia larvata* are present on all four Gulf of Guinea islands. The populations from Bioko and Annobón have been placed together with the subspecies occurring on the neighbouring coast (*hypoleuca* from Cameroon and Gabon). The populations from Príncipe and São Tomé are treated has distinct endemic subspecies (*principalis* and *simplex*, respectively). Based on morphological and vocalization differences together with playback experiments, the population from São Tomé is sometimes treated as an endemic species (*A. simplex*; del Hoyo *et al.* 1997; Jones & Tye 2006). The extremely large and fragmented range of this species has created patterns of differentiation of difficult interpretation.

In addition to uncertainties regarding the species status of the Gulf of Guinea island pigeons, the taxonomic position of *A. larvata* and the bronze-naped pigeons within *Columbidae* is still unclear. Some previous taxonomic sequences have intriguingly placed these species, often in close proximity to the bronze-naped pigeons, after New World representatives of *Columba* (Gibbs *et al.* 2001), which were recently placed in the separate genus *Patagioenas* based on molecular phylogenetic results (Johnson et al. 2001). No previous molecular phylogenetic study included *Aplopelia* or a representative of the bronze-naped pigeons so these taxonomic arrangements have not been tested using molecular data. Other authors have used morphological similarities between the Red-eye Dove (*Streptopelia semitorquata*) and *Aplopelia* to suggest a closer relationship to the genus *Streptopelia* (Goodwin 1983). Goodwin (1983) separates *Aplopelia* from the bronze-naped pigeons, placing them in sequence after *Streptopelia* suggesting that *Aplopelia* originated in Africa, either from an ancestor to *Streptopelia* species, or from

some form ancestral to both *Streptopelia* and *Columba* (Goodwin 1983). Generally, most current treatments place *Aplopelia* near the Old-World *Columba*, particularly in association with the bronze-naped pigeons, based on the close similarities in plumage to the Western Bronze-naped pigeon (*C. iriditorques;* del Hoyo *et al.* 1997). A subgenus *Turturoena* has sometimes been proposed for the bronze-naped pigeons, emphasizing their distinctiveness from other *Columba* (del Hoyo *et al*. 1997). The creation of the separate genus *Aplopelia* for the lemon-doves has been proposed on the basis of its peculiar foraging habits (ground-foraging, mostly seed-eating pigeons). Although it is arguable that this degree of adaptive divergence might not be high enough to warrant elevation to a distinct genus, the use of the monotypic *Aplopelia* designation remains useful as a way of highlighting the uncertainty of the relationships of the lemon-doves within the *Columbidae*.

Understanding the relationships between species and accessing the degree of differentiation of the endemic pigeons is important to access the uniqueness of these species. The IUCN recognizes three primary levels of biodiversity with conservation priority: gene, species and ecosystems; hence evaluating the species status of a taxon group is crucial for the design of efficient strategies for biodiversity management and conservation. Misidentifying a threatened taxon group can lead to failure in protection and consequent extinction, whereas, on the other hand, identification of too many taxa can waste conservation efforts (Allendorf *et al.* 2007).

Thus the objectives of this study were to:

1. Determine the phylogenetic position of the endemic pigeons of São Tomé and Príncipe and the degree of differentiation from the mainland species.
2. Infer the phylogenetic position of the *A. larvata* and the bronze-naped pigeons (Subgenus: *Turturoena*) within the *Columbidae*, and determining if some of the island populations may constitute distinct, endemic, species.

METHODS

Sampling

Sampling for DNA extraction from the Gulf of Guinea islands species was done by collaborators involved in different tasks of the conservation project. Samples from the three endemics (*C. malherbii, C. thomensis, T. sanctithomae*) and from *A. larvata* came both from the biweekly collection of birds killed by hunters and from Martim Melo's collection. Toes and blood samples and, when available, muscle tissue (heart and liver) were collected and conserved in absolute ethanol.

Samples from mainland taxa proved hard to obtain. Fresh samples of *A. larvata* were obtained from the collection of the Field Museum of Natural History (Chicago); samples from *T. calva* came from the collection of Martim Melo. Samples of *C. iriditorques* and *C. delegorguei* were lent by the Yale Peabody Museum of Natural History (Table 1).

Previously sequenced species from the *Columbidae* family were also included in this study (see Appendix 1.for detailed information) adding up to a total of 42 taxa analysed. The entire phylogeny is rooted with the Namaqua sandgrouse *Pterocles namaqua* (*Pteroclidae*), which according to Hackett *et al*. (2008) is the sister group to the Columbidae.

Table 1. Taxa, individuals and sampling localities obtained for this study. Country: STP (São Tomé and Príncipe); EG (Equatorial Guinea)

Taxon	Individuals	Country	Locality	Sample	Collection
Aplopelia larvata	Aplopelia_larvata_P1	STP-Príncipe	Água Petróleo	Blood	M. Melo
	Aplopelia_larvata_P26	STP-Príncipe	Boca do Inferno	Blood	M. Melo
	Aplopelia_larvata_ST14	STP-São Tomé	Lagoa Amélia	Blood	M. Melo
	Aplopelia_larvata_ST19	STP-São Tomé	Bom Sucesso	Blood	M. Melo
	FMNH444011	Malawi	Ntchisi Forest Reserve	Footpad	FMNH
Columba malherbii	Columba_malherbii_ST5	STP-São Tomé	Água Izé	Muscle	M. Melo
	Columba_malherbii_ST7	STP-São Tomé	Água Izé	Muscle	M. Melo
	Columba_malherbii_P4	STP-Príncipe	Ribeira Izé	Muscle	M. Melo
	Columba_malherbii_P5	STP-Príncipe	Ribeira Izé	Muscle	M. Melo
Columba thomensis	Columba_thomensis_1	STP-São Tomé	Lagoa Amélia	Blood	M. Melo
	Columba_thomensis_2	STP-São Tomé	Bom Sucesso	Muscle	M. Melo
Columba delegorguei	YPM88245	Tanzania	–	Footpad	YPM
Columba iriditorques	YPM50221	Angola	–	Footpad	YPM
Treron calva	Treron_calva_B1	EG-Bioko	Moka	Muscle	M. Melo
	Treron_calva_B2	EG-Bioko	Moka	Blood	M. Melo
	Treron_calva_P9	Príncipe	Ribeira Izé	Muscle	M. Melo
	Treron_calva_P10	Príncipe	Ribeira Izé	Muscle	M. Melo
	Treron_calva_P11	Príncipe	Ribeira Izé	Muscle	M. Melo
	Treron_calva_mainland	Angola	Kumbira, Conda	Tissue (Toes)	M. Melo
Treron sanctithomae	Treron_sanctithomae_1	STP-São Tomé	Alto Douro	Blood	M. Melo
	Treron_sanctithomae_3	STP-São Tomé	Lagoa Amélia	Tissue (Toes)	M. Melo

YPM: Yale Peabody Museum of Natural History; FMNH: Field Museum of Natural History

Laboratory procedures

DNA extraction

Total genomic DNA was extracted from blood, toe and muscle tissue using a JETQUICK Tissue DNA Spin Kit (Genomed), following the protocol for DNA extraction from animal tissues. Because all samples were stored in ethanol, they were left to dry in an incubator at 54ºC before the extraction.

PCR amplification

Mitochondrial DNA (mtDNA)

Three mitochondrial markers were amplified: i) the complete NADH dehydrogenase subunit 2 (ND2: 1041 bp), ii) the cytochrome *b* (cyt *b*: 1068 bp), and iii) the cytochrome oxidase I (COI: 381 bp). Overlapping sequence fragments were amplified via polymerase chain reaction (PCR; primers, references and conditions detailed in Appendix 2.).

Extraction of museum samples tends to produce low quality and low quantity DNA. Usually the samples are old and thus the DNA is degraded, and it is not possible to obtain large quantities of sampling material, as this would damage the museum specimen. The museum samples of *C. delegorguei* and *C. iriditorques* were respectively 56 and 51 years old. In order to obtain the largest sequence coverage possible, amplification of certain markers had to be made in smaller fragments.

For ND2 and COI the same procedures used for fresh samples produced good quality amplifications. The cytochrome *b* was obtained in two separate fragments (cytb-1: 366bp cytb-2: 178bp; primers, references and conditions detailed in Appendix 3.).

Nuclear DNA

Molecular information provided solely from mitochondrial markers can lead to biased estimation of species relationships, as mitochondrial genes are inherited from the maternal line as a single linkage group (Ballard & Whitlock 2004). Independent molecular information must come from the nuclear genome. Introns and exons are commonly used in phylogenetic studies. In birds, introns seem to outperform exons as they often harbour more variation and appear to give better support to deep branches (Chojnowski *et al.* 2008; Prychitko & Moore 2003; Fain & Houde 2004).

For this study the β-fibrinogen intron 7 (FIB7:833 to 1011 bp) was sequenced. Overlapping sequence fragments were amplified via PCR (primers, references and conditions detailed in Appendix 2.).

Primers used for DNA amplification from fresh samples failed to amplify the museum samples – likely because the problem of DNA degradation was compounded by the smaller quantities of nuclear DNA in the skin cells relatively to mitochondrial DNA. Amplification of the total length of the FIB7 was done in four overlapping fragments using the two published primers together with six new primers specifically designed for this study (Appendix 3). Primers were designed with the program Primer3 (Untergrasser *et al.* 2012; Koressaar *et al.* 2007) using the alignment of previously sequenced samples from *C. malherbii* and *C. thomensis* and the default parameters. Amplification success was tested using two previously sequenced individuals of *C. malherbii* and *C. thomensis* that were re-amplified and re-sequenced with the new primers (primers, references and conditions detailed in Appendix 3.).

Sequencing

The PCR products were electrophoresed in a 2% agarose gel stained with GelRed Nucleic Acid Stain (10,000x in water, BIOTIUM) and visualized using an ultraviolet transilluminator. Successful PCR amplifications were purified using ExoSAP according to the manufacturer's instructions (United States Biochemical Corporation, Cleveland, Ohio).

The amplification primers were used to sequence both DNA strands with the dye-labelled termination method using ABI BigDye v.3.1 cycle sequencing kit (following the manufacturer's protocol for 10 µl reactions). Post sequencing reaction product was purified using Sefadex (GE Healthcare). Fragments were read in ABI 3130xl Genetic Analizer.

Sequence editing and alignment

All sequences were edited using Geneious R6.1 (Biomatters, New Zealand). All sequences were visually inspected in order to detect any base miscalling. Forward and reversed sequences were assembled and trimmed, and a search for double peaks was conducted using the "Find Heterozygotes" plug-in and by eye. Base callings of too low quality were coded as missing data (assigned as "N"). Sequences were aligned using the Geneious Align and MUSCLE Align options; all alignments were inspected by eye.

For the mitochondrial markers, protein coding sequences were checked for gaps, deletions or insertions and all were translated into amino acid for making sure all translated correctly, thus reducing the possibility of including nuclear copies of mitochondrial genes (Bensasson *et al.* 2001). For this study all heterozygous bases were coded as missing data (assigned as "N").

Sequences characteristics

Pairwise genetic distances (uncorrected and model corrected) within the ingroup taxa were calculated to assess genetic distances between taxa using PAUP* version 4.0b10 (Swofford 2003), excluding the outgroup sequence. Taxa with biased base composition may affect phylogenetic inference (Sanderson & Shaffer 2002). Therefore, homogeneity of base composition across ingroup taxa was calculated for each gene and each codon using a chi-square analysis.

Phylogenetic Inference

Bayesian and maximum likelihood inference allows combining information from different data partitions or subsets evolving under different evolutionary models. This enables to analyse heterogeneous data sets consisting of different data types and to explore a wide variety of structured models mixing partition-unique and shared parameters. The hallmark of this approach is to allow phylogenetic inference under mixed models to accommodate data heterogeneity (Ronquist & Huelsenbeck 2003). Introduction of *a priori* data partition schemes was a step forward in addressing heterogeneity across sites (Ronquist & Huelsenbeck 2003; Nylander *et al.* 2004).

Model of evolution / Best partition of the data

For each marker, the best partitioning scheme and model of evolution were chosen using the Bayesian inference criterion (BIC) as implemented in Partition Finder v.0.92 (Lanfear *et al.* 2012). The "All" search algorithm was used to find an appropriate partitioning scheme, letting the branch lengths unlinked between subsets.

Bayesian test of incongruence (BTI)

Compatibility of the phylogenetic signal was assessed using a Bayesian test of incongruence (Irestedt *et al.* 2004). Two separate tests were conducted: 1. To assess the phylogenetic congruence of the mitochondrial genes; 2. To assess the phylogenetic congruence of the complete dataset (mitochondrial genes and the nuclear intron).

For each test, two Bayesian searches were run: one where all the genes were constrained to follow the same topology (but allowing each one to follow its own substitution model) and another where each gene is allowed to have an independent topology. Bayesian inference (BI) as implemented in MrBayes version 3.2.1(Ronquist *et al.* 2012) was used to obtain the trees under each model. Bayes factors approach was

used to compare both analyses, and results were interpreted following the guidelines of Kass & Raftery (1995).

Maximum Likelihood inference

Maximum-Likelihood (ML) phylogenies were estimated, for each mitochondrial gene, for the concatenated mitochondrial dataset, and for the nuclear intron, using the software Garli v2.0 (Zwickl 2006). The 'best-fit' model and partition, as determined with Partition Finder, was used without fixing the model parameters. No starting topology was defined. The program was set up to run until no significantly better scoring topology (defined by an lnL increase of 0.01) was found. For each locus, five independent search replicate runs were performed to appraise consistency of the estimates. All trees were combined in a 50% majority rule consensus tree computed in PAUP*.

Bayesian inference

Phylogenetic trees were computed for each locus and for the combined mtDNA data set using MrBayes v.3.2.1 (Ronquist & Huelsenbeck 2003). Each partition had its own set of parameters, and evolutionary rates were allowed to vary across partitions under a flat Dirichlet prior (Ronquist & Huelsenbeck 2003). For all analyses, two independent MrBayes runs were carried out simultaneously, starting from random trees and using the default priors. Each run consisted of two simultaneous Metropolis-coupled Markov chain Monte Carlo (MC^3) with chain heating parameter set to 0.5 and a length of 10 million generations. Trees were sampled every 1000 generations with the first 25% discarded as burn-in. Convergence of the MC^3 , stationarity of the runs, and effective sample size for each parameter of interest in the analysis were evaluated using the software TRACER 1.5 (Rambaut & Drummond 2007). Further assessment of the convergence was made by combining the two independent runs using LogCombiner v1.7.5 (Drummond *et al.* 2012) and assessing the same parameters for the combined runs. Bayesian posterior probabilities were obtained for each clade from the 50% majority rule consensus tree of the 30 000 sampled trees.

Species tree inference

Gene trees do not always reflect species relationships. There are many potential sources of discrepancy between gene trees and species trees: gene duplication; introgression; lineage sorting; and deep coalescence (Heled & Drummond 2010; Knowles *et al.* 2010; Edwards 2009). Two different approaches have been proposed for species tree analysis: data concatenation (Huelsenbeck *et al.*1994) and coalescent-based approaches (Heled & Drummond 2010)

In the concatenated approach the data is treated as a single locus and the genealogy is averaged across genes. The rationale behind this method is that phylogenetic accuracy improves with increasing number of variable sites (Huelsenbeck *et al.*1994). While this assumption is true for a particular locus the same does not necessarily apply across loci (different loci can have different topologies).

Coalescence-based approaches allow for incongruence across loci. In a coalescent approach, gene trees are enclosed within the species tree by following the coalescent process back in time within each branch, a process known as a multispecies coalescent (Rannala & Yang 2003; Heled & Drummond 2010)

Coalescent-based methods seem to perform better in phylogeographic studies, where incongruent lineage sorting is evident across loci (Carstens & Knowles 2007), whereas concatenation is preferred for studies at deeper taxonomic level (Wiens *et al.* 2008).

Since each method has strengths and weaknesses, both approaches were employed in this study.

The analysis of the combined dataset (mtDNA+ FIB7) was performed using MrBayes. Each partition had its own set of parameters, evolutionary rates were allowed to vary across partitions under a flat Dirichlet prior The following parameters were set as unlinked [revmat = (all), shape= (all), pinvar =(all), statefreq = (all), tratio = (all), brlens=all].

Two independent MrBayes runs were carried out simultaneously, starting from random trees and using the default priors. Each run consisted of two simultaneous MC^3 chains, with chain heating parameter set to 0.5 and a length of 100 million generations. Trees were sampled every 1000 generations with the first 25% discarded as burn-in.

Assessment of the MC^3 chain convergence was made following the same procedure described previously. Bayesian posterior probabilities were estimated for each clade from the 50% majority rule consensus tree of the 300 000 sampled trees.

The coalescent-based method *BEAST (Heled & Drummond 2010), as implemented in software BEAST v1.7.1 (Drummond *et al.* 2012), was used to infer the coalescent species tree. Posterior phylogenies were determined in *BEAST using a strict clock and the Yule tree prior. The default settings were used as priors. Three independent runs of 200 million generations sampling every 1000 generations were performed. Convergence of the MCMC runs was evaluated as before. The initial 10% of the runs were discarded as burn-in. Summary trees were generated in TreeAnnotator v1.7.1, part of the BEAST package. The resulting trees were analyzed in FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/)

All phylogenetic analyses were run through the CIPRES Science Gateway computational portal (Miller *et al.* 2010).

Divergence time estimation

To test if the trees followed a molecular clock, a likelihood-ratio test (Felsenstein 1981) was conducted in PAUP*. The ML tree generated without the clock constraint was tested against the ML tree obtained by forcing a constant rate of evolution along all its branches. The likelihood-ratio test supported a molecular clock (-lnL clock-enforced tree = 23515.319, -lnL unconstrained tree = 23482.354 χ^2=0.023 P= 0.97).

Many studies have shown that molecular data tend to violate the assumption of a strict molecular clock and that some variation will be present in the rates of substitution among lineages (Drummond *et al.* 2005; Lepage *et al.* 2007; Ronquist *et al.* 2012). New methods allow for inferring divergence times without assuming a strict clock model providing a more flexible way to model the rate variation (relaxed clock models; Drummond *et al.* 2005; Drummond *et al.* 2012). Divergence events were estimated, for the combined mtDNA data set, using BEAST v. 1.7.5 to perform a Bayesian relaxed clock analysis (Drummond *et al.* 2012). The rate of 1.9% per million years (Myr) for cyt *b* reported for *Columbiformes* (Weir & Schluter 2008) was used. A lognormal relaxed clock was set, and a uniform prior was used for the rate with a mean of 1.9, upper bound of 2.07 and lower bound of 1.86 (Weir & Schluter 2008), with the remaining priors following the default choices.Three independent runs of 10 million generations sampling every 1000 generations were performed. Tracer 1.5 was used to assess the convergence of MCMC chain.

RESULTS

Sequence characteristics

All markers amplified successfully, and sequences were of good quality. No gaps or insertions were found in the mtDNA markers, and all translated appropriately, reducing the possibility that nuclear copies of mitochondrial genes were included (Bensasson *et al.* 2001). The complete ND2 (1041bp) and cyt *b* genes were amplified for all species except for the museum skin samples (*C. delegorguei* - cyt *b*: 504 bp; ND2: 507 bp; *C. iriditorques* - ND2: all, cyt *b*: 504bp). The complete FIB7 intron was successfully amplified, with variable length between species due to the presence of insertions and deletions. Regarding the species amplified for this study, the length of the FIB7 was of

833 bp for species in the genus *Aplopelia*, 842 bp for *Columba* and of 1011bp for *Treron.* The COI gene was the most variable (informative sites: 29%), with the ND2 and cyt *b* being slightly less variable (informative sites: 17.5% and 13.7%, respectively). The FIB7 intron had similar levels of variability to these (informative sites: 15%; Table 2.). Base frequencies of all loci were homogeneous across taxa (P = 1.0 in all χ^2 tests).

The average uncorrected pairwise distances between the island endemics and their mainland counterparts were located within a narrow range (mtDNA: 2.1-3.5%, FIB7:0.12-0.75%). The island forms of *A. larvata*, generally treated as insular populations of the mainland species, had the largest genetic distances from their mainland counterparts (mtDNA: 4.3%; FIB7:0.85%; Table 3.)

Table 2. Sequence characteristics, based on all ingroup taxa. mtDNA: ND2, Cyt *b*, COI; nDNA: FIB7

	Length (bp)	Var	Info	%A	%C	%G	%T
ND2							
All	1041	212	182	32.1	33.5	10.2	24.2
1st	347	54	37	37.6	28.6	15.9	17.7
2nd	347	21	17	16.4	35.2	8.8	39.7
3rd	346	137	128	42.2	36.8	58.6	15.2
Cyt *b*							
All	1068	168	146	26.6	34.4	13.1	25.9
1st	355	29	22	25.0	29.6	21.8	23.6
2nd	355	7	4	20.3	25.9	13.4	40.3
3rd	355	132	120	34.4	47.6	40.5	13.9
COI							
All	381	120	111	27.1	27.7	17.9	27.3
1st	127	11	9	29.2	16.8	32.2	21.6
2nd	127	0	0	15.7	24.3	17.4	42.5
3rd	127	109	102	36.3	41.9	4.07	17.8
FIB7	1011*	267	155	33.8	17.5	18.3	30.5

Var: variable sites; Info: parsimony informative sites.
*Maximum length obtained.

Tabel 3. Uncorrected pairwise distances between the island and mainland species

	p' uncorrected distances (%)	
Species pairs	**mtDNA**	**FIB7**
C. malherbii - *C. delegorguei*	2.8	0.12
C. malherbii - *C. iriditorques*	2.8	0.65
C. iriditorques - *C. delegorguei*	2.1	0.55
C. thomensis - *C. arquatrix*	2.8	0.49
T. calva (mainland) - *T. calva* (islands)	2.8	0.63
T. calva (mainland) - *T. sanctithomae*	3.5	0.75
T. sanctithomae - *T. calva* (islands)	2.3	0.13
A. lavata (mainland) - *A. larvata* (islands)	4.3	0.85

Phylogenetic Inference

The analysis for the best partition and model of evolution recovered the same partition for all mitochondrial markers, with the first and second codons evolving at the same rate, distinct from the third codon (Table 4.).

Table 4. Estimation of the best partition of the data and correspondent model of evolution

Marker	Partition/DNA model		Log-likelihood
ND2	[1st, 2nd]	[3rd]	-10943.61
	GTR+I+Γ	GTR+I+Γ	
Cytb	[1st, 2nd]	[3rd]	-9480.75
	HKY+I+Γ	GTR+Γ	
COI	[1st, 2nd]	[3rd]	-3308.09
	TrN+I	TrN+Γ	
FIB7	HKY+I+Γ		-3611.68

Square brackets bound elements of a partition. 1st=first codon position;
2nd =second codon position; 3rd=third codon position

Gene trees inference

The Bayesian test of incongruence supported the congruence of the phylogenetic signal of the three mitochondrial markers, and hence the data were combined for the analysis (Bayes Factors analysis strongly supported a linked topology: 2logB12= 1435.3 >> 10). This signal of congruence was further verified by inferring individual gene trees; these did not show topological changes at supported nodes (Appendix 6.1 to 6.4)

ML and BI analysis recovered 5 major clades corresponding to the Old world *Columba*, New World *Columba* (= *Patagioenas*), the "mid-sized" New World Ground

Doves, the bulk of the species making up the *Streptopelia* genus, and the *Treron* genus. The current *Streptopelia* genus was not monophyletic. Monophyly between the Old world *Columba* and the largest *Streptopelia* clade was supported in all analyses (bootstrap:100%, posterior probability:1.0), but relationships within these two major groups were not very well resolved. These results are in agreement with previously published works on the phylogeny of Columbiformes (Johnson *et al.* 2001; Pereira *et al.* 2007; Fig.7.).

In agreement with current taxonomy, *C. arquatrix* and *C. thomensis* were sister species (all analyses: bootstrap :100%, posterior probability:1.0), belonging with strong support to the Old world *Columba* group in all analysis.

Again, in accordance with current taxonomy, the bronze-naped pigeons (*C. delegorguei*, *C. iriditorques*, and *C. malherbii*) formed a monophyletic clade, supported in all analyses. The relationships within this clade could not be confidently established. Nevertheless the analysis of the FIB7 intron recovered the mainland *C. delegorguei* and *C. iriditorques* as a monophyletic group (supported in ML and BI) separated from *C. malherbii*.

Concerning the species of the genus *Treron,* the mtDNA data recovered, with strong support, a separation between mainland and insular species (bootstrap: 100%, posterior probability: 1.0). Interestingly the individuals from Bioko were more closely related to the oceanic islands populations rather than to the closer mainland populations. Relationships within the insular species (*T. calva* and *T. sanctithomae*) were not clear, with only BI supporting the separation between the two species. Nevertheless this clade includes a very wide range of genetic distances, from 0.1% between the *T. sanctithomae* individuals and 2.5% between *T. calva* and *T. sanctithomae* (see Appendix 4.). An individual from Príncipe (P10) grouped with *T. sanctithomae* rather than with *T. calva* from Bioko and Príncipe, with a genetic distance of 0.2% (very similar to the distances found within *T. sanctithomae*). FIB7 analyses were inconclusive and the *Treron* clade was placed as a basal polytomy to the *Columbiformes* (see Appendix 6.3.).

Within *Aplopelia larvata*, an unexpected deep island-mainland divergence was recovered with strong support by all analyses (except for a ML bootstrap support of 70% for the FIB7 data set). Bayesian inference of mtDNA data also supported a split between island forms (posterior probabilities: 1.0).

The phylogenetic position of *Aplopelia* and the bronze-naped pigeon group (subgenus *Turturoena*) remained elusive. Most of the analyses were inconclusive and these two groups were recovered as two polytomies being part of the large clade comprising the Old word *Columba* and most *Streptopelia*. Only BI using the mtDNA sequences identified these two groups as being more closed related to the Old word *Columba* rather than to

Streptopelia (not strongly supported, posterior probability: 0.92). A sister relationship between *Aplopelia* and the bronze-naped pigeons was recovered by the combined mtDNA BI analysis.

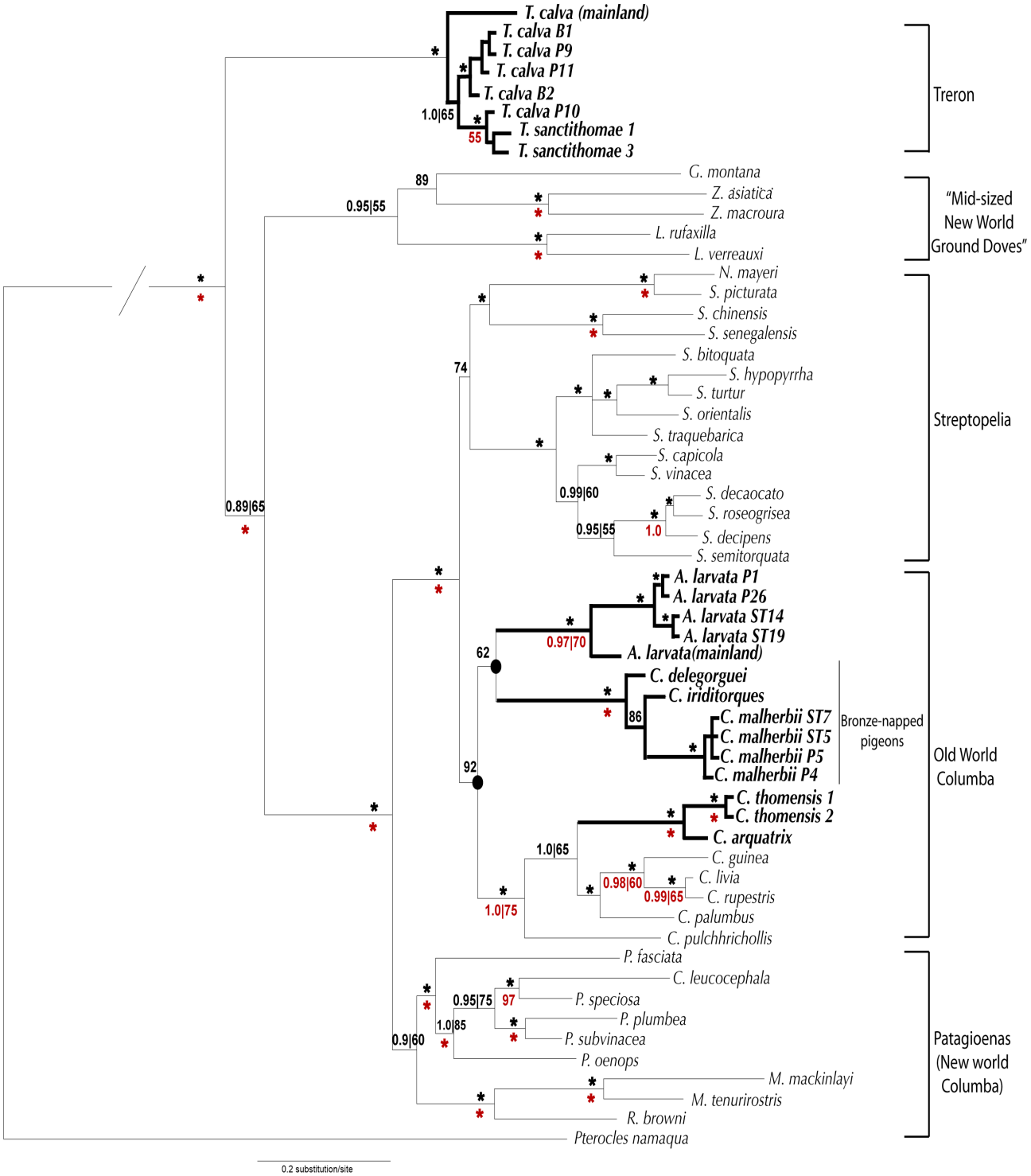


Figure 7. Bayesian inference tree obtained from the combined mitochondrial dataset, .The Maximum likelihood tree had a similar topology. Support values above branches: Bayesian posterior probabilities (≥0.5) followed by ML bootstrap values (≥50). Support values below branches (red): Bayesian posterior probabilities and bootstrap values for the tree inferred from the FIB7 intron. Nodes supported in both BI and ML inference (posterior probabilities>0.95 and bootstrap values≥75%) are coded as "*" for the mitochondrial data set and "*" for the FIB7 data set. Code "**P**" in front of species names refers to Príncipe, "**ST**" to São Tomé and "**B**" to Bioko islands (for complete name of the species see Table 1. and Appendix 1.)

Species tree inference

The Bayesian test of incongruence supported the congruence of the phylogenetic signal of the mtDNA markers and the FIB7 intron (Bayes Factors analysis strongly supported a linked topology: 2logB12= 1225.94 >> 10). Species tree inference gave similar topologies to those obtained with gene trees, as effectively only two loci were analysed (mtDNA and one intron). The tree obtained from concatenating the sequence data was better supported that the one using a coalescent approach.

The five major clades obtained with gene trees were recovered with strong support. Monophyly between the *Streptopelia* and the Old World *Columba* group was recovered in all analyses, but the relationships within this group are still not well resolved (Fig.8.).

Mainland-island splits were strongly supported for all endemic species (bootstrap:100%, posterior probability:1.0 for all analysis). Splits between conspecific or closely related island taxa were not identified by the coalescent species-tree, but were well supported under the concatenated approach. A clear split between *A. larvata* individuals from São Tomé and from Príncipe was strongly supported (bootstrap:100%, posterior probability:1.0). The same was observed between *T. calva* (Bioko and Príncipe) and the endemic *T. sanctithomae* (although the odd P10 individual, Fig.7., was not included in this data (see Appendix 6.4. for the concatenated approach).

The sister species relationship between the bronze-naped pigeons and *A. larvata* was identified by both approaches, albeit with low support (coalescent species-tree posterior probability: 0.59; BI under the concatenated approach, posterior probability: 0.75). These species were placed within the Old World *Columba*, but with low support (BI under the concatenated approach, posterior probability:0.89).

Divergence times

The Bayesian relaxed molecular clock dating methods implemented in BEAST revealed that the divergence times between the mainland and the São Tomé and Príncipe pigeons are all recent. The most recent split occurred for *T. calva* (0.91 Myr.), and the oldest for *A. larvata* (2.4 Myr.; Table 5.)

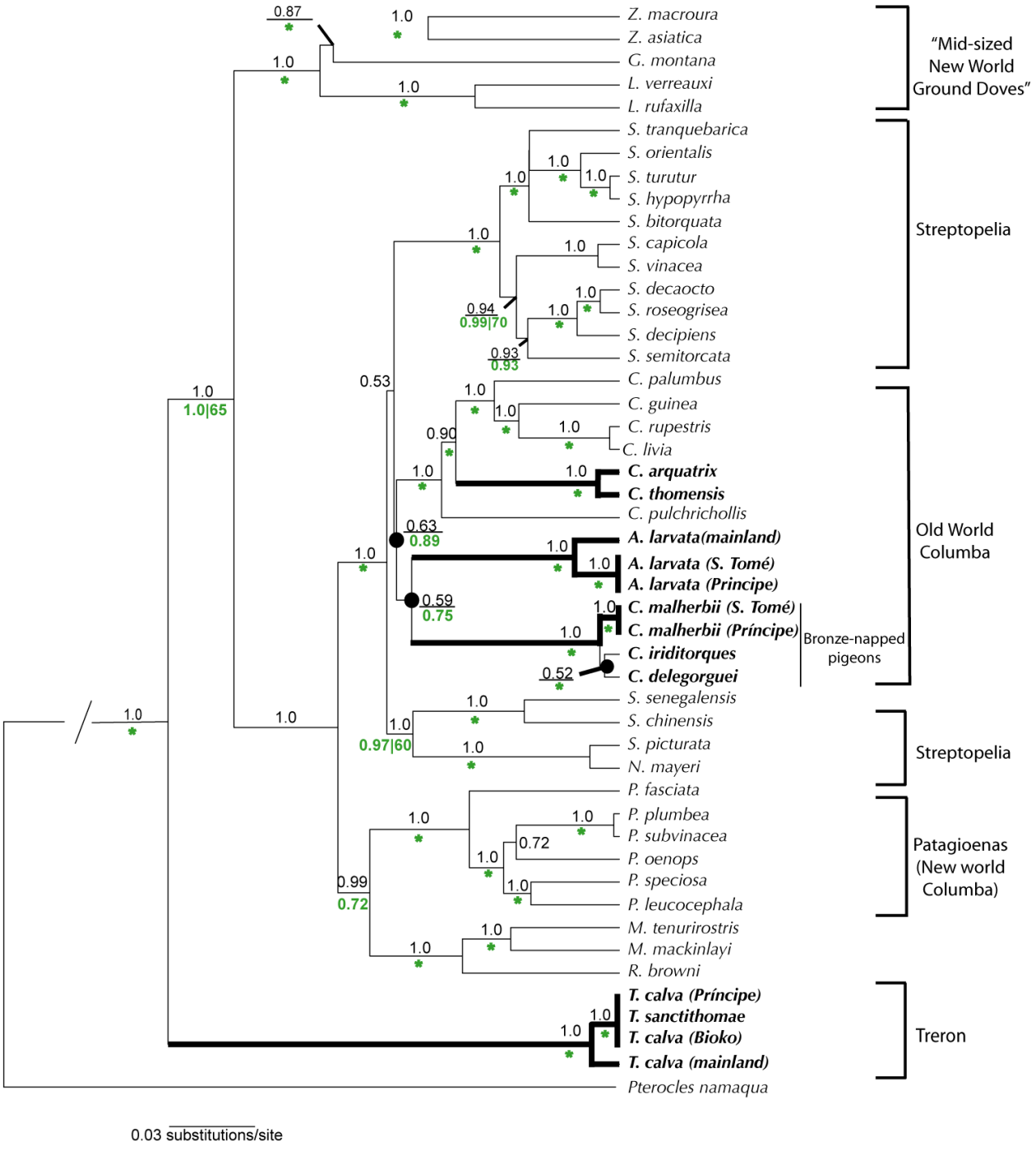


Figure 8. Species-tree (3 mtDNA markers, 1 nDNA intron) inferred with the coalescent-based approach implemented in *BEAST. The Bayesian Inference tree inferred with the concatenated dataset had an almost identical topology, with differences only at the level of different populations of the same species . Support values above branches: Bayesian posterior probabilities (≥0.5) for the *BEAST tree. Support values below branches (green): Bayesian posterior probabilities (≥0.5) followed by ML bootstrap values (≥50) for the concatenated data set; nodes supported both by BI and ML inference (posterior probabilities > 0.95 and bootstrap values ≥ 75%) are coded as "*". (for complete name of the species see Table 1. and Appendix 1.)

Table 5. Bayesian divergence time estimates for the most recent common ancestor (MRCA) of different taxa based in the mtDNA data set. Lower and upper 95% confidence intervals shown. Dates are presented in million years before present.

Clade	Time of the MRCA
A. larvata (islands) - *A. larvata* (mainland)	1,616–**2,364**–3,179
C. thomensis - *C. arquatrix*	0,952–**1,505**–2,083
C. malherbii - *C. iriditorques*	0,564–**1,279**–2,141
C. malherbii - *C. delegorguei*	0,835–**1,264**–1,738
C. iriditorques - *C. delegorguei*	0,539–**1,118**–1,752
T. calva (islands) - *T. calva* (mainland)	0,587–**0,912**–1,259
T. sanctithomae – *T. calva*(mainland)	1,938–**1,960**–2,572

DISCUSSION

This study presents the first molecular evidence on the evolutionary history of the pigeons of the Gulf of Guinea islands. We were particularly interested in determining the level of differentiation of the endemic pigeons of the islands of São Tomé and Príncipe, and in clarifying the phylogenetic position of *Aplopelia larvata* and of the bronze-naped pigeon group within Columbiformes.

From a general viewpoint, the genus *Streptopelia* was not monophyletic and must therefore be revised, whereas the majority of the Old World *Columba* did form a monophyletic group. *Aplopelia larvata* and the bronze-naped pigeons (*Columba malherbii, iriditorques, and delegorguei*) were placed with strong support in the clade comprising the Old World *Columba* and the largest *Streptopelia* clade; however, their position within this group could not be determined. All analyses that used multiple gene regions placed *Aplopelia* and the bronze-naped pigeons as sister taxa, and these formed a clade sister to the remainder of Old World *Columba*. However, this relationship was never strongly supported. This lack of resolution is most likely due to a rapid radiation in this group leading to short internal branches and hard polytomies (Gonzalez *et al.* 2009). Either *Aplopelia* should be synonymized with *Columba* (as in del Hoyo *et al.* 1997), or a new genus (i.e. *Turturoena*) should be designated for both *Aplopelia* and the bronze-naped pigeons. The latter treatment would be more reflective of the current lack of support for the precise placement of *Aplopelia* and the bronze-naped pigeons. The relationships recovered here were in agreement with earlier studies of the Columbiformes (Johnson & Clayton 2000; Johnson *et al.* 2001 Pereira *et al.* 2007).

The São Tomé and Príncipe endemic pigeons

Molecular data recovered similar relationships among these pigeon species as previous taxonomic studies based on morphological data (del Hoyo *et al.*1997; Dowsett & Dowsett-Lemaire 1993: Jones & Tye 2006). *C. thomensis* is sister to *C. arquatrix*; *C. malherbii* to the mainland bronze-naped pigeons (*C. iriditorques*, *C. delegorguei*); and *T. sanctithomae* to the continental *T. calva.* Two unexpected results did emerge. First, all green pigeons from the Gulf of Guinea islands, including Bioko, grouped together in a separate sister clade to the mainland sample. There was also very little divergence among the insular *Treron* taxa, even though only one of the islands is separated as a distinct species (*T. sanctithomae*). Secondly, the *Aplopelia larvata* populations from Príncipe and São Tomé (Annobón sample not available) were the taxa most genetically differentiated from their mainland counterparts, even though these are usually not separated as distinct species.

Interpretation of the similarities/differences between the *T. sanctithomae* and the islands form of *T. calva*, must be made with caution both because the phylogenetic analysis was inconclusive, and second because too few individuals were sampled from both species. The fact that the *T. calva* from Bioko grouped closest to the island species might be a consequence of just sampling one individual from a far away location on the mainland (Angola). The inclusion of individuals closest to the coast (Gabon and Cameroon) could change this result and it is possible that mainland *T. calva* are paraphyletic with respect to the island forms. A more comprehensive study employing techniques of phylogeography and population genetics will be needed to assess the evolutionary history of these two groups.

The degree of separation between mainland and island forms of the *Aplopelia larvata* group was the largest found among mainland-island pairs of this study (4.3% mtDNA uncorrected p' distances; 4.5% corrected under the GTR model). This split is further supported by a 3 bp insertion in the FIB7 intron in all the *A. larvata* insular individuals. This is an interesting case in the sense that the *A. larvata* populations from Príncipe and São Tomé species have been normally regarded as distinct endemic subspecies of the mainland population (Jones & Tye 2006), but the degree of genetic differentiation obtained in this study is almost twice of that obtained for the other species considered as valid endemic species. Although del Hoyo *et al.* (1997) had already suggested the separation of the São Tomé population as a different species (*Columba simplex*), results from this study suggest similar degrees of differentiation between the São Tomé and Príncipe populations and the mainland forms. This result must however be taken with caution. As with *Treron calva*, our only sample from the mainland came from a distant

population (Malawi, subspecies *larvata*) and not from the subspecies present in the Gulf of Guinea coast (*hypoleuca*). The large sequence divergence reported here could be a result of cryptic speciation in the mainland.

Very little is known about the relationships between and within the bronze-naped pigeon group. The phylogenetic inference from this study supports the superspecies view: *C. malherbii* is a distinct species, recently derived from the same common ancestor of the two mainland species.

Species delimitation

Similar divergence times were recovered between all mainland-island species pairs. In average mainland and island species differentiated around 1 to 2.4 Myr ago (Table 5.). All the splits identified in the phylogenetic analysis are amongst the most recent ones found for sister species of pigeons (Johnson & Clayton 2000; Johnson *et al.* 2001; Gonzalez *et al.* 2009), and lie close to the 2 Myr border associated, broadly, in birds with the speciation border (Price 2008). A meta-analysis showed that most avian species pairs have a genetic distance in the range of 1.5% to 4% (Johns & Avise 1998), which includes the values for the mainland-island distances found here.

Most studies done with birds from São Tomé and Príncipe have recovered similar and even lower levels of differentiation (Melo & O'Ryan 2007; Melo *et al.* 2010). Within the Columbiformes, some unambiguous *Streptopelia* and *Columba* species pairs show smaller or the same range of genetic distances/divergence times as those obtained between the São Tomé and Príncipe endemics and the mainland counterparts (*C. rupestris-livia*; *S. turtur-hypopyrrha*; *S. roseogrisea-decaocto*; *S. vinacea-capicola*; Johnson *et al.* 2001). Gonzalez *et al.* (2009) obtained similar results for the differentiation between the endemic Canarian pigeon, Bolle's pigeon (*Columba bolli*) and the mainland counterpart the wood pigeon (*Columba palumbus*). Using similar divergence time estimation methods, they dated the split at 1.7 Myr ago.

In summary, the results obtained in this study support a split between the endemics and the mainland forms around the upper Pleistocene (around 1.3 to 2Myr ago). These dates are consistent with a period of intense volcanic activity that led to the rise of the highest mountains of São Tomé and Príncipe (Caldeira *et al.* 2003). This volcanic activity may have led previous pigeons populations to extinction, while preparing the conditions for the development of extensive suitable habitat for the successful establishment of new colonisers. The fact that most endemic bird species appear to have been formed around this same period (Melo 2007) provides support for a dramatic turnover in the fauna of the islands coincident with this last period of intense volcanism. Additionally, because of the

relative proximity of the islands to the continent, the formation of high mountains (935 m in Príncipe, and 2024 m in São Tomé) might have enhanced the capability of nomadic long-distance flyers (like pigeons) to locate the islands.

CONCLUSION

Species are the fundamental units of biology, ecology and conservation. Although the classification of any taxon as a species does not change our understanding of their evolutionary history (Winker *et al.* 2007; Mallet 2008), it has important implications for conservation policies, since 'species' are the preferred unit for prioritizing and formulating laws for conservation purposes (Garnett & Christidis 2007).

The multitude of different species definitions brought some discussion about which is the better way to define a species. The biological species concept (BSC) relays on the principle of reproductive isolation that is still considered the defining feature of the speciation process (Mayr 1942; Coyne & Orr 2004). The phylogenetic species concept originally defined species as "the smallest diagnosable cluster of organisms within which there is a parental pattern of ancestry and descend"(Cracraft 1989). This concept has as its cornerstones 'diagnosability' and 'monophyly' (parental pattern of ancestry; Price 2008). Despite many different definitions and many debates on the species concept these two seem to be the most widely used. While the reproductive isolation feature (BSC) is easy to assess in sympatric species, the phylogenetic species definition deals better with the inference of separation of allopatric populations where reproductive isolation normally cannot be assessed (Hey 2001). In this study I followed the philosophy proposed by Helbig *et al.* (2002) and further developed by Tobias *et al.* (2010), which states that genetic traits can be used as proxies for the likelihood that differences accumulated during evolutionary times can lead to reproductive isolation. Comparisons with the divergence observed between sympatric species of the same group and agreement between these comparisons, provides valuable evidence for the species status

The techniques employed in this study tried to minimize the uncertainly of classifying allopatric taxa by using more than one locus, inferring the congruence/incongruence of the phylogenetic signal and evaluating relationships between taxon groups under different methods (gene and species trees). Ideally, our molecular data should have been combined with a study on phenotypic differentiation, but the lack of data for the mainland pigeons (*C. iriditorques, C. delegorguei, C. arquatrix* and *T. calva*) prevented such analysis. Sound taxonomic decisions should be made on the maximum number of

characters available including ecological, behavioural, phenotypic and genotypic data (Tobias *et al.* 2010).

To summarize, based on molecular data and the previously described morphological differences (Amadon 1953, Gibbs *et al.* 2001, Jones & Tye 2006, del Hoyo *et al.* 2007), the species status of the endemic *C. malherbii* and *C. thomensis* can be confidently accepted. The same trend was observed for *T. sanctithomae*, but the lack of a denser taxon sampling from the large mainland distribution of its closest relative, *T. calva,* does not allow to draw firm conclusions at this stage. Either no significant genetic differentiation will be found or, if current results are confirmed, the three Gulf of Guinea islands populations will be grouped in the same endemic species. The same limitation occurred in relation to the status of *Aplopelia.* The high levels of differentiation observed between the islands' populations and a geographically distant population must be confirmed in relation to the populations of the neighbouring coast. Nevertheless these high levels of differentiation suggest that the mainland subspecies of *Aplopelia* might indeed be different species and highlights the need of a more complete study of the phylogenetic relationships of the African forest pigeons. The possible existence of two isolated populations living in each island deserves further attention and studies done at a smaller scale (with fast evolving markers: microsatellites) employing techniques of population genetics should be conducted.

The possible split between the Old world *Columba* and the bronze-napped pigeons and the position of the enigmatic *A. larvata* within the bronze-napped pigeons group (subgenus *Turturoena*) opens the door for further studies and to a better understanding of the relationships within this group.

Future work to infer the species status of *T. sanctithomae* and *A. larvata* will have to be made; sampling as much as possible the entire range of the mainland species will be essential. In the same way, to better understand the relationships within the African pigeons and doves a larger sampling effort will have to be made and future analysis should include a larger array of genetic markers.

CHAPTER 2.

POPULATION GENETICS OF THE SÃO TOMÉ AND PRÍNCIPE ENDEMIC PIGEONS

ABSTRACT

The endemic pigeon species of São Tomé and Príncipe (*Columba malherbii*, *Columba thomensis* and *Treron sanctithomae*) have been suffering from severe hunting pressures. Although until now no study was conducted to assess the effects of hunting, preliminary results from a recent field study shows that hunting is being responsible for the rapid decline of the pigeons community.

Conservation genetics aims to apply genetic methods to the conservation and restoration of biodiversity. Variables such as genetic diversity and effective population size and gene flow can be used as proxies to assess the adaptive potential of a population, whereas estimates of the levels of gene flow between populations are fundamental for the management of threatened species.

In this study, one mitochondrial gene (ND2) and one nuclear intron (FIB5) were used to determine these variables, in order to provide guidelines for the development of conservation strategies.

The genetic signal was typical of stable or growing populations, fitting the expected pattern of population expansion after the recent colonisation events (Chapter 1). Levels of nucleotide diversity were intermediate ranging from 0.00187 to 0.00079 for the ND2 data set and 0.00353 to 0.00023 for the FIB5, with *C. thomensis* having the lowest levels of diversity. These diversity values translated directly in the estimates of the scaled effective population size (Θ), which showed similar values for *C. malherbii* and *T. sanctithomae* ($\Theta = N_e\mu = 0.0023$ and $\Theta = 0.0015$, respectively), with the *C. thomensis*

estimate being one order of magnitude lower (Θ =0.0006). The contrast between the genetic signal of stable populations and the recent reported decline suggests that the recent hunting activity is driving the decline and raises the concern that populations may be undergoing a bottleneck that cannot be detected by the markers used in this study.

Genetic exchange between the populations of *C. malherbii* from Príncipe and São Tomé was high and in accordance with current taxonomy (this species being also known as the Gulf of Guinea pigeon). Although preliminary, the results on the rate of migration between the two islands suggest that this is a normal behaviour in the species. By contrast, the estimates of migration rates between *T. sanctithomae* and *T. calva* were overall low (mode = 0.3 migrants per generation) and well below the theoretical level required for differentiation to take place. Differentiation of the two taxa was confirmed by AMOVA and tests of population differentiation (G-test, P<0.001). Taken together these results give further support for the current treatment of the two green pigeons populations of São Tomé and Príncipe as distinct evolutionary lineages and possibly species. The history of these lineages is nevertheless complex, but typical of archipelago settings, as the haplotype network revealed two instances of recent secondary contact – with two haplotypes from São Tomé being present in the Príncipe population.

Keywords: *Columba thomensis, Columba malherbii*, *Treron sanctithomae, Treron calva,* conservation genetics, nucleotide diversity, effective population size.

INTRODUCTION

The phylogenetic approach used in Chapter 1 showed that the endemic pigeons of São Tomé and Príncipe constitute independent lineages from their mainland counterparts that have resulted from relatively recent colonisation events. In fact, although the species status of *Columba thomensis* and *C. malherbii* was well supported, the situation for *Treron sanctithomae* was not as clear. Population genetics can be used to clarify the evolutionary dynamics of systems of recent and shallow divergence: how much lineage sorting has effectively occurred, what is the degree of connectivity between populations, etc. This information will determine how evolutionarily independent are the populations of interest and hence how significant are they for conservation.

Conservation genetics aims to study the genetic variables of a natural population that suffers from anthropogenic changes, in order to provide guidelines for reducing the probability of extinction (Allendorf *et al.* 2007). Population genetics data are a very important source of information on past demographic dynamics. Namely, using the coalescent approach (Kingman 1982; Tavare 1984), one can estimate current effective population size ($N_{e;}$ Wright 1931) and infer changes it may have suffered in the recent past – as population expansions and/or declines and bottlenecks. This is extremely relevant information as, obviously, the effective population size – the actual number of breeding individuals in a population – is the parameter of most importance for conservation. Population genetics approaches can determine, for example, if a species has had long-term small effective population sizes and may therefore be adapted to rarity; Frankham 1995) or if it is going through a human-driven bottleneck (historical large N_e but small census size) and hence may be prone to inbreeding depression.

Small and isolated populations (like island species) are more susceptible to environmental, demographic and genetic effects (inbreeding and loss of genetic diversity; Allendorf *et al.* 2007). The role of genetic factors in directly driving extinction of small populations has been controversial, although it is widely accepted that the genetic makeup of a population is extremely valuable in inferring extinction risks (Frankham *et al.* 2002).

Species from oceanic islands are renowned for their much higher risk of extinction relatively to mainland populations. In historical times, 90% of all bird extinctions have occurred on oceanic islands (Myers 1979). Human activity, mainly through habitat change and hunting are the main reasons leading island birds to extinction (Frankham 1998). Low effective population sizes, higher rates of genetic drift, and reduced gene flow in island populations all result in reduced genetic diversity that may increase

extinction by increasing the probability of inbreeding depression and by decreasing the adaptive potential of the populations (Allendorf *et al.* 2007).

This study uses a population genetics approach to infer demographic parameters that, in combination with the current census efforts on the ground, will help clarify potential impacts that hunting activity is having on the three endemic pigeons of São Tomé island. Additionally, population genetics will be used to estimate the degree of connectivity between the islands of Príncipe and São Tomé for the bronze-naped pigeons (*C. malherbii*) and the green pigeon (*T. sanctithomae* and *T. calva*) populations. In the latter case, the data can allow clarifying the evolutionary relationships between the *Treron sanctithomae* and the *Treron Calva*, which remained uncertain in the phylogenetic analyses (Chapter 1).

METHODS

Sampling

Sampling for DNA extraction was done as described in Chapter 1. Sample sizes were as follow: *C. thomensis*: 30; *C. malherbii* – Príncipe: 9; São Tomé: 28; *T. sanctithomae*: 31; *T. calva* - Príncipe: 14; Bioko: 2. Sampling locations are depicted in Figure 9.

Laboratory procedures

DNA extraction was performed as described in Chapter 1.

During the work of this thesis a set of 30 microsatellites were tested. Microsatellites, as fast evolving markers, would be the most appropriate tool for inferring relationships between populations and closely related species. Nevertheless from the microsatellites tested only eight were successfully amplified but most were non-informative . Recently a microsatellite library was created for these species and future work will involve a study using these markers.

Thus one mitochondrial and one nuclear intron were amplified: i) the complete NADH dehydrogenase subunit 2 (ND2: 1041 bp), and the ii) β-fibrinogen intron 5 (FIB5: 531 to 546 bp). Overlapping sequence fragments were amplified via polymerase chain reaction (PCR; primers, references and conditions detailed in Appendix 7.)

Individuals were sexed following a molecular protocol (Griffiths *et al.* 1998). A PCR reaction amplifies homologous segments of the two sex-linked chromo-helicase-DNA-

binding genes (CHD1). Copies of each sex chromosome (Z and W) differ in size and when tested in an agarose gel, females depict two bands (WZ) and males one band (ZZ; primers, references and conditions detailed in Appendix 7.)

DNA sequencing was done using the amplification primers and following the protocol described in Chapter 1.

Sequence editing and alignment

All sequences were handled with Geneious R6.1(Biomatters, New Zealand). All sequences were checked visually in order to detect any base miscalling. Forward and reversed sequences were assembled and trimmed, and a search for double peaks was conducted using the "Find Heterozygotes" plug-in and further inspect visualy. Base callings of too low quality were coded as missing data (assigned as “N”). Sequences were aligned using the Geneious Align and MUSCLE Align options; all alignments were inspected by eye.

For the mitochondrial marker, protein coding sequences were checked for gaps, deletions or insertions and all were translated into amino acid form making sure all translated correctly, thus reducing the possibility of including nuclear copies of mitochondrial genes (Bensasson *et al.* 2001).

For the nuclear intron, haplotype inference was carried out with the PHASE v.2.1 algorithm (Stephens & Donnelly 2003) available in DnaSP v. 5.10.1 (Librado & Rozas 2009). Haplotypes were inferred for each individual species data set. The algorithm was run with 100 burn-in iterations and 100 sampling iterations (thinning interval 1). A model that allowed for recombination was used, and the initial estimate of the recombination parameter was set to (Li & Stephens 2003) to 0.0004 (standard value). Haplotypes were retained if genotype phase probability was ≥0.6, which according to Stephens *et al.* (2001) corresponds to an estimated probability of correct phase call of 0.82. All remaining analyses were computed using the phased data set.

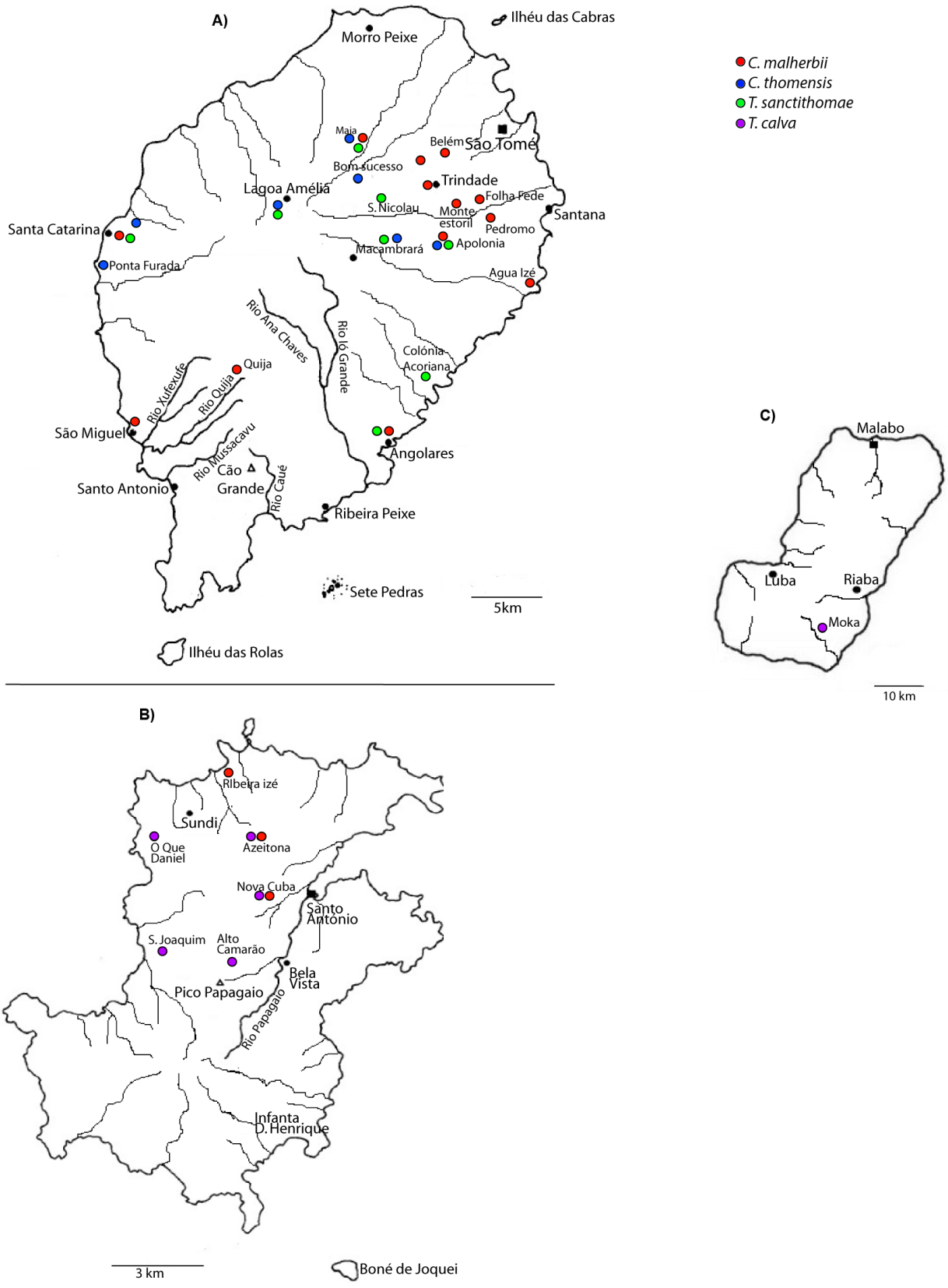


Figure 9. Sampling locations and taxa obtained for this study. A) São Tomé; B) Príncipe; C) Bioko. (see Appendix 8 to 10 for detailed information).

Sequences characteristics

Base composition, variable and parsimony informative sites were assessed using PAUP* version 4.0b10 (Swofford 2003).

The software DnaSP v. 5.10.1 was used to calculate a number of summary statistics: nucleotide diversity π (Nei 1987), haplotype diversity H_D (Nei 1987) and number of segregating sites S (Watterson 1975). Summary statistics were calculated for each individual species data set. In the case of the *C. malherbii* the same analysis was done separately for the populations of São Tomé and Príncipe.

Population Structure

To depict the genetic relationships between the populations, phylogenetic and network trees were constructed. Each analysis was conducted for the two markers and for the individual species data set with the exception of the *Treron* genus. Previous phylogenetic analysis (Chapter 1) recovered the possibility of gene flow between *T. sanctithomae* and *T. calva*, and to better assess this possibility the two species were included in the same data set.

For each marker and species data set, the best partitioning scheme and model of evolution were chosen using the Bayesian inference criterion (BIC) as implemented in Partition Finder v.0.92 (Lanfear *et al.* 2012). The "All" search algorithm was used to find an appropriate partitioning scheme, allowing the branch lengths to be unlinked between subsets (Appendix 11.).

Phylogenetic trees were constructed using BEAST v.1.7.1 (Drummond *et al.* 2012). Posterior phylogenies were determined using a strict clock and a Coalescent: Constant size prior (as suggested by the authors for population analyses). The default settings were used as priors. Three independent runs of 10 million generations sampling every 1000 generations were performed, with the first 10% discarded as burn-in. Convergence of the Markov chain Monte Carlo (MCMC), stationarity of the runs, and effective sample size were evaluated using the software TRACER 1.5 (Rambaut & Drummond 2007). LogCombiner v1.7.5 (Drummond *et al.* 2012) was used to combine the three independent runs. Convergence of the three runs was also assessed in TRACER 1.5. PAUP* was used to obtain the Bayesian posterior probabilities for each clade from the 50% majority rule consensus tree of the 27 000 sampled trees.

Phylogenetic methods normally have low performances analysing intraspecific data, mainly because of the large sample sizes and small genetic distances between

individuals. Network trees depict better the multitude of plausible trees, and allow a better understanding of the relationships between populations and recently diverged taxa (Bandlet *et al.*1999). Software Network v 4.6.1.1 (Bandelt *et al.* 1999), was used to compute the Median-Joining networks trees. The default settings were applied to the algorithm.

An analysis of molecular variance (AMOVA; Excoffier *et al.* 1992) as implemented in Alerquin v.3.5 (Excoffier *et al.* 2010) was conducted to assess the population structure of the *C. malherbii* populations of São Tomé and Príncipe and of the two green pigeons (*T. sanctithomae* and *T. calva*). Population differentiation was tested with the exact G-test of differentiation (Raymond & Rousset 1995).

Demographic analyses

Departure from a neutral model of molecular evolution was tested in DnaSP v.5.10.1 using Tajima's *D* (Tajima 1989) and Fu's Fs (Fu 1997) tests. Statistical significance associated with these tests was assessed by using 50000 coalescent simulations.

The Bayesian Skyline Plot (BSP) approach was used to explore the population history. This coalescent-based method performs an estimation of changes in effective population size through time (Drummond *et al.* 2005). The BSP method uses a piecewise-constant model of population size that can fit a wide range of demographic scenarios. It allows the effective population size to change every time there is a coalescent event in the genealogy (Pybus *et al.* 2000). BSP analysis was conducted in BEAST v.1.7. Three independent runs of 10 million generations sampling every 1000 generations were performed, with the first 10% discarded as burn-in. For the ND2 data set, the rate of 1.9% per million years (Myr) for cyt *b*, as reported for Columbiformes (Weir & Schluter 2008) was used. A strict clock was set, and a uniform prior was used for the rate with a mean of 1.9, upper bound of 2.07 and lower bound of 1.86 (from Weir & Schluter 2008), with the remaining priors following the default choices. Convergence of the MCMC was assessed using the previously described procedure.

The coalescent-based method of Beerli & Felsenstein (1999) and Beerli & Palczewski (2010), implemented in MIGRATE-N 3.6 (Beerli 2009), was used to estimate the mutation-scaled effective population size (Θ) for the three pigeon groups. Migration rates relative to the mutation rate between Príncipe and São Tomé islands were estimated for the *Treron* (Príncipe: *T. calva;* São Tomé: *T. sanctithomae*) and for the *C. malherbii* populations.

Parameters were estimated with the Bayesian approach. The ND2 data was given an inheritance scalar of 0.25, as the N_e mitochondrial DNA is four times smaller than the N_e

of nuclear DNA. Short exploratory runs were done to determine the most effective choice of analytical tools (e.g., choice between 'slice' or the 'Metropolis-Hastings' algorithms to calculate posterior distributions, most adequate prior distribution to reach convergence) and parameters (e.g., prior parameters, bin size for the time intervals used in reconstruction of past migration and coalescence events). The final search used most of the default options and parameters, with the posterior distributions being generated with the Metropolis-Hastings sampling using a uniform prior with the default parameters. A 'static' heating scheme was used, with four chains at the suggested temperatures (1.0 1.5 3.0 1,000,000.0). Ten independent searches were run simultaneously and their post-burn-in results were combined. Each search consisted in one long chain of 5,000,000 steps, with genealogies sampled 100 steps apart (50,000 sampled trees, from which the first 10,000 were discarded as burn-in). In this strategy, the parameter space was therefore explored in 50 million steps, and a total of 400,000 genealogies were used for parameter estimation. Analysis of the FIB5 dataset were not included here. It seemed that searches of the migration parameters got stuck in areas of flat likelihoods leading the program to estimate exceedingly high migration events per generation that, at least for *C. malherbii* always reached the chosen top boundary of the prior. Furthermore, in spite of the extensive search of the parameter space, the autocorrelation of the estimates from the ND2 dataset remained high and the acceptance ratios low – a clear example of the complexity of this methodology (Beerli 2009). Although the present results must be taken as preliminary, they remain informative when interpreted from a relative point of view (i.e., which population has the highest Θ, are migration rates high or low, are they symmetric or not, etc.).

Future work will include increasing the length of the long chain and the steps between sampled genealogies and using the maximum likelihood approach. This was not possible at the moment due to CPU time constraints. Likewise, an alternative coalescent-based approach to estimate N_e and migration rates will be done using the program IMa2 (Hey 2010).

RESULTS

Sequence characteristics

ND2 and FIB5 were successfully amplified, and sequences were of good quality. No gaps or insertions were found in the ND2, and it translated appropriately, reducing the possibility that nuclear copies of mitochondrial genes were included (Bensasson *et al.* 2001). The complete length of the markers was not obtained, and it varied between species (Table 6.). For haplotype inference, genotype phase probabilities were in general high (most often 1.0 or close to 1.0), thus no haplotypes were excluded.

Overall a low number of variable and informative sites were found within each species data set (Table 6). Base frequencies of both loci were homogeneous across taxa (P = 1.0 in all χ^2 tests).

ND2 estimates of polymorphism showed relatively high levels of nucleotide diversity for *T. calva* (π= 0.007; H_D= 0.4) intermediate for *T. sanctithomae* (π= 0.00187; H_D= 0.4) and *C. malherbii* (π= 0.00092; H_D= 0.6). *C. thomensis* had by far the lowest levels of diversity (π= 0.00079; H_D= 0.5). In comparison levels of nucleotide diversity calculated from the FIB5 data set were surprisingly high for an intron, with nucleotide diversity ranging from 0.00353 to 0.0002 (only *C. thomensis* showed lower levels of diversity in comparison with the ND2 data set; Table 7.)

Population Structure/Demographic analyses

Columba malherbii

The ND2 variation patterns departed from the neutral expectations both for the São Tomé population and for the two populations taken together (D=-3.348, P<0.01; Fs=-3.813; P<0.001). The large negative values, especially on São Tomé, are associated with recent population expansion (Aris-Brosou & Excoffier 1996; Fu 1997). For FIB5, only Fu's Fs detected departure from neutrality in the São Tomé population (Fs= -8.956, P=0.001; Table 7.)

Bayesian skyline plots provided similar results, with the effective population size being either stable (ND2) or depicting weak signs of expansion (Fib5; Fig.11). The mutation-scaled effective population size of the São Tomé population of *C. malherbii*, as estimated by coalescent simulations, was the largest of all populations analysed (mean Θ = 0.007; Table 8.)

Tabel 6. Sequence characteristics for the three São Tomé and Príncipe pigeons populations, based on all ingroup taxa

		Lenght (bp)	Var	Info	%A	%C	%G	%T
Columba malherbii (São Tomé)	**ND2**							
	All	897	11	1	31.5	33.5	11.1	23.9
	1st	299	3	0	37.8	29.6	15.1	17.4
	2nd	299	0	0	25.3	33.6	12.0	29.1
	3rd	299	8	1	31.1	37.4	6.2	25.3
	FIB5	531	9	7	31.7	17.5	19.9	30.8
Columba malherbii (Príncipe)	**ND2**							
	All	897	3	0	31.5	33.5	11.0	23.9
	1st	299	1	0	37.8	29.7	15.0	17.04
	2nd	299	1	0	25.4	33.6	11.9	29.1
	3rd	299	1	0	31.2	37.3	6.1	25.3
	FIB5	531	4	3	31.7	17.6	19.9	30.8
Columba thomensis	**ND2**							
	All	999	5	3	32.0	34.2	10.3	23.5
	1st	333	1	1	30.9	29.3	14.4	17.0
	2nd	333	0	0	17.1	35.3	8.4	39.0
	3rd	333	4	2	39.9	38.0	8.0	14.0
	FIB5	546	1	1	31.7	18.1	19.6	30.6
Treron sanctithomae	**ND2**							
	All	957	11	8	32.1	33.4	10.3	24.2
	1st	319	2	1	36.3	28.8	16.6	18.1
	2nd	319	1	1	16.6	35.1	8.8	39.5
	3rd	319	8	6	43.3	36.3	5.3	15.0
	FIB5	536	4	3	32.5	16.8	19.5	31.2
Treron calva	**ND2**							
	All	957	27	23	32.4	33.6	10.1	23.8
	1st	319	3	3	36.1	28.8	16.9	18.2
	2nd	319	2	2	16.3	35.4	9.0	39.2
	3rd	319	22	18	44.6	36.9	45.0	14.0
	FIB5	536	10	9	32.4	16.7	19.7	31.3

Table 7. Population genetic statistics for each species. n: number of individuals; S: segregating sites; Hd: haplotype diversity; π: nucleotide diversity; D: Tajima's D; Fs: Fu's Fs.(*p≤0.001; **p≤0.01; *** p≤0.05)

	Columba malherbii																	
	São Tomé						Príncipe						ALL					
Marker	**n**	**S**	**H_D**	**π**	**D**	**Fs**	**n**	**S**	**H_D**	**π**	**D**	**Fs**	**n**	**S**	**H_D**	**π**	**TD**	**Fs**
ND2	28	11	0,638	0,00095	-2,283**	-10.141*	9	3	0,583	0,00076	-1,513	-1.892	37	13	0,617	0,00092	-3,348**	-13.419*
FIB5	28	9	0,858	0,00353	-0,0957	-8.312*	9	4	0,824	0,00273	0,743	-1.351	37	9	0,845	0,00334	-0.0917	8.956*

	Columba thomensis					
Marker	**n**	**S**	**H_D**	**π**	**D**	**Fs**
ND2	30	5	0,538	0,00079	-1,024	-2.385***
FIB5	30	1	0,127	0,00023	-0,526	-0.246

	Treron sanctithomae						*Treron calva*						ALL					
Marker	**n**	**S**	**H_D**	**π**	**D**	**Fs**	**n**	**S**	**H_D**	**π**	**D**	**Fs**	**n**	**S**	**H_D**	**π**	**D**	**Fs**
ND2	31	9	0,389	0,00187	-0,652	1.001	14	27	0.396	0.007	-0.881	5.610	44	32	0.668	0.01108	1.5121	-14.641
FIB5	30	4	0,526	0,00135	-0,331	0.195	14	10	0.396	0.00546	0.446	-7.405*	45	14	0.755	0.00308	-1.0946	6.759***

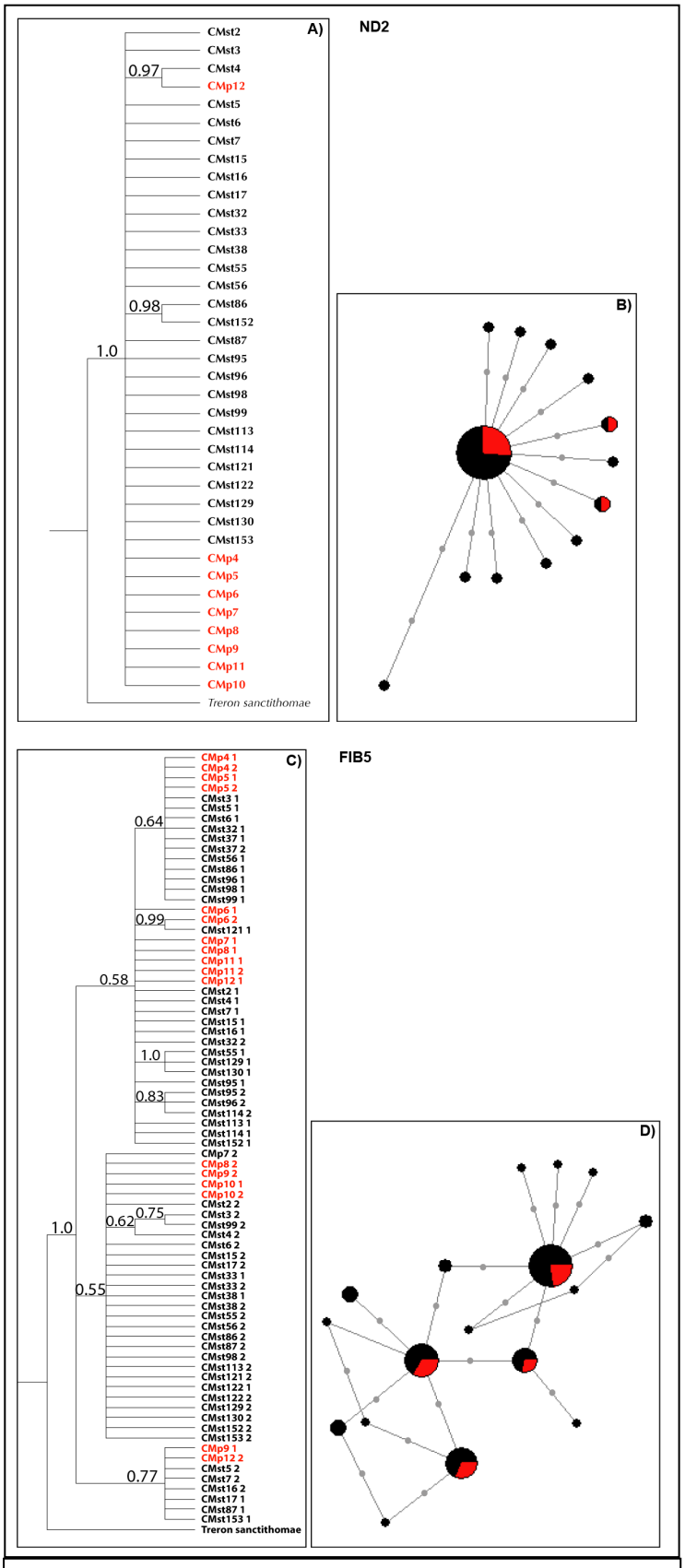


Figure 10. Phylogenetic and network trees of the *C. malherbii* population. A) ND2 based phylogenetic tree; B) ND2 based network tree; C) FIB5 phased sequence phylogenetic tree; D) FIB5 phased sequence network tree. Individuals from Príncipe coloured in red; in black individuals from São Tomé.

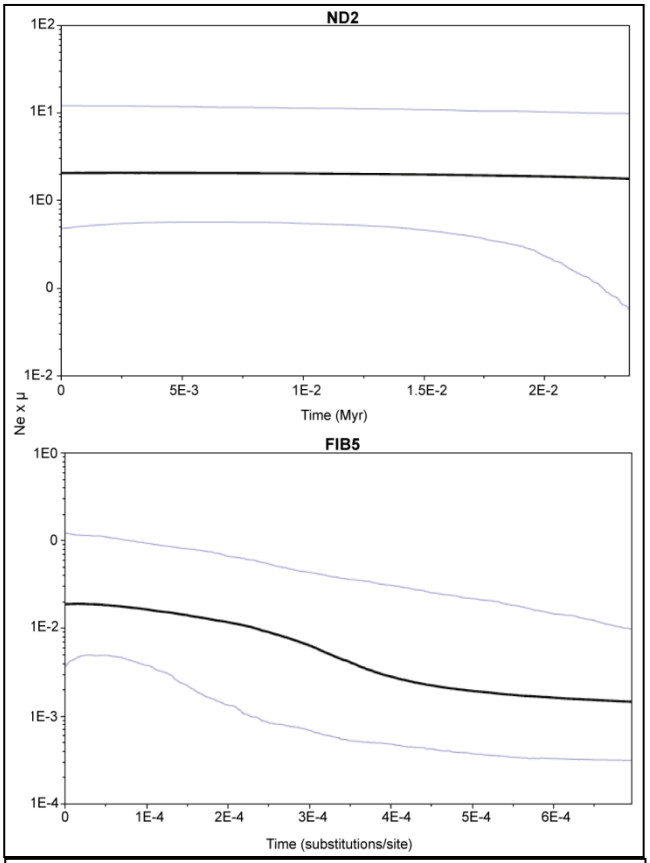


Figure 11. Bayesian skyline plot obtained for the *C. malherbii* population. The y axis is the product of the effective population size by the mutation rate ($N_e\mu$) and the x axis is in units of time (ND2: million years; FIB5: substitutions per site per million year). Blue lines correspond to the 95% HPD (highest posterior density) limits and black line corresponds to the median

No signal of population structure was found within *C. malherbii*. Phylogenetic trees and networks recovered no separation between the two island populations, with haplotypes being shared between islands (Fig.10.). AMOVA analysis was not statistically significant, with an Φ_{ST} of -0.035 for the ND2 (P=0.97) data set and -0.019 for the FIB5 (P=0.87). Consequently, the exact test of differentiation was also not significant (ND2: P=0.73; FIB5=0.94).

Lack of population differentiation translated in very high migration rates between the two islands. However as stated above, convergence of the runs was difficult to achieve and the posterior distribution had a very long tail to the right, which led to large estimates of the mean effective number of migrants per generation, N_m (c. 300-350 birds in both directions), although the most common values were much lower (mode: 6, 17). In any case, immigration rates were always above the theoretical one-migrant-per-generation considered to prevent population differentiation (Wright 1931; Spieth 1974; Table 9.)

Table 8. Estimatives of the scaled effective population size for all populations in study ($\Theta = N_E*\mu$; μ = mutation rate per site per year). ST: São Tomé; P: Príncipe.

Θ	Mean	Median	Mode
Columba thomensis	0,0006	0,0006	0,0006
***Columba malherbii* - ST**	0,0072	0,0023	0,0012
***Columba malherbii* - P**	0,0051	0,0023	0,0008
***Treron sanctithomae* - ST**	0,0022	0,0015	0,0010
***Treron calva* - P**	0,0028	0,0016	0,0009

Table 9. Estimatives of the migration rates from both directions (SãoTomé→Príncipe; Príncipe→SãoTomé) for the *C. malherbi* and *Treron* populations. Units in individual per generation.

Nm	Mean	Median	Mode
Columba malherbii			
Príncipe →São Tomé	312	122	17
São Tomé → Príncipe	354	407	6
Treron calva / sanctithomae			
Príncipe → São Tomé	61	6	0,3
São Tomé → Príncipe	71	8	0,3

Columba thomensis

C. thomensis had the lowest levels of genetic diversity of all studied species (ND2: π=0.00079; FIB5: π=0.00023 . There were only six ND2 haplotypes, two to four bp apart, and a single Fib5 haplotype was present in 27 individuals, with a second haplotype, two bp apart, present in only three individuals (Fig.12). ND2 and FIB5 data sets did not show deviation from the neutral expectation, both under Tajima's D and Fu's F (Table 7.).

Bayesian skyline plots of the ND2 data set inferred a stable population through time (Fig.13). Due to the lack of information of the FIB5 marker, the BSP could not calculate the coalescent events and thus this analysis could not be accomplished. Θ for the *C. thomensis* population had the lowest values, one order of magnitude smaller than those of the other populations (Θ = 0.0006; Table 8.).

Treron

For *T. sanctithomae* none of the markers showed deviations from the neutral model of evolution with (Tajima' D and Fu's Fs; P>0.10; Table 7.). BSP analyses identified a stable population throughout time, with a very weak signal for a recent demographic expansion (around 250 thousand years ago) being picked up by the ND2 (Fig. 15.). Regarding the *T. calva* population, for the ND2 data set Tajimas D and Fu's Fs tests identified no departures from neutral expectation (P>0.1). For FIB5, only Fu's Fs detected departure from neutrality (Fs=-7.405; P≤0.001). Θ estimates of the *T. calva* and *T. sanctithomae* population were similar and in the same order of magnitude as that of *C. malherbii* (Θ = 0.0015 and 0.0016 respectively; Table 8.).

Interesting results came out of the phylogenetic and network analyses of the *Treron* taxa of São Tomé and Príncipe. The ND2 data set recovered a split between the two taxa although this was not supported (posterior probability:0.68; Dxy= 2.2%; Fig. 14^{A,B}.). As in the phylogenetic study, two individuals of *T. calva* grouped together with *T. sanctithomae*. Analysis of the nuclear intron recovered no clear signs of a split between the two taxa with some haplotypes being shared among them (Fig. 14^{C,D}.). Nevertheless, even with haplotype sharing present, AMOVA results from both markers came out significant with an Φ_{ST} of 0.589 for the ND2 (P<0.001) data set and 0.219 for the FIB5 (P<0.001). The exact test of differentiation between *T. sanctithomae* and *T. calva* was also significant (ND2 and FIB5: P<0.001).

Within the *T. sanctithomae* population the ND2 phylogenetic and network analyses recovered two divergent haplotype groups (Fig.14^{A,B}.). As no geographical pattern was found between these groups, no further tests on within island population structuring

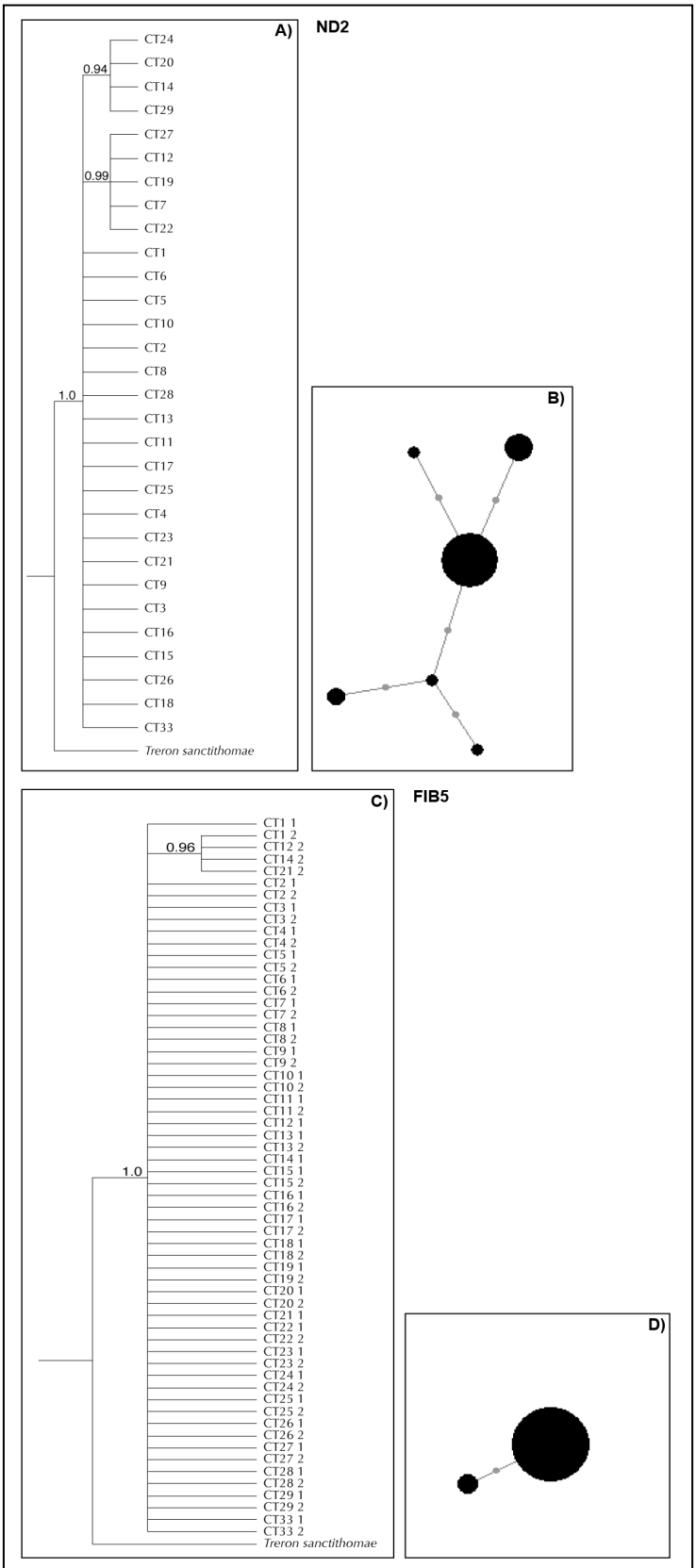


Figure 12. Phylogenetic and network trees of the *C. thomensis* population. A) ND2 based phylogenetic tree; B) ND2 based network tree; C) FIB5 phased sequence phylogenetic tree; D) FIB5 phased sequence network tree.

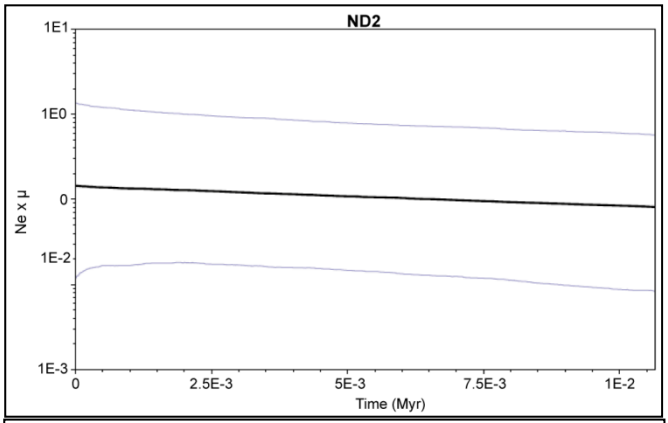


Figure 13. Bayesian skyline plot obtained for the *C. thomensis* population. The y axis is the product of the effective population size by the mutation rate ($N_e\mu$) and the x axis is in units of time (million years). Blue lines correspond to the 95% HPD (highest posterior density) limits and black line corresponds to the median

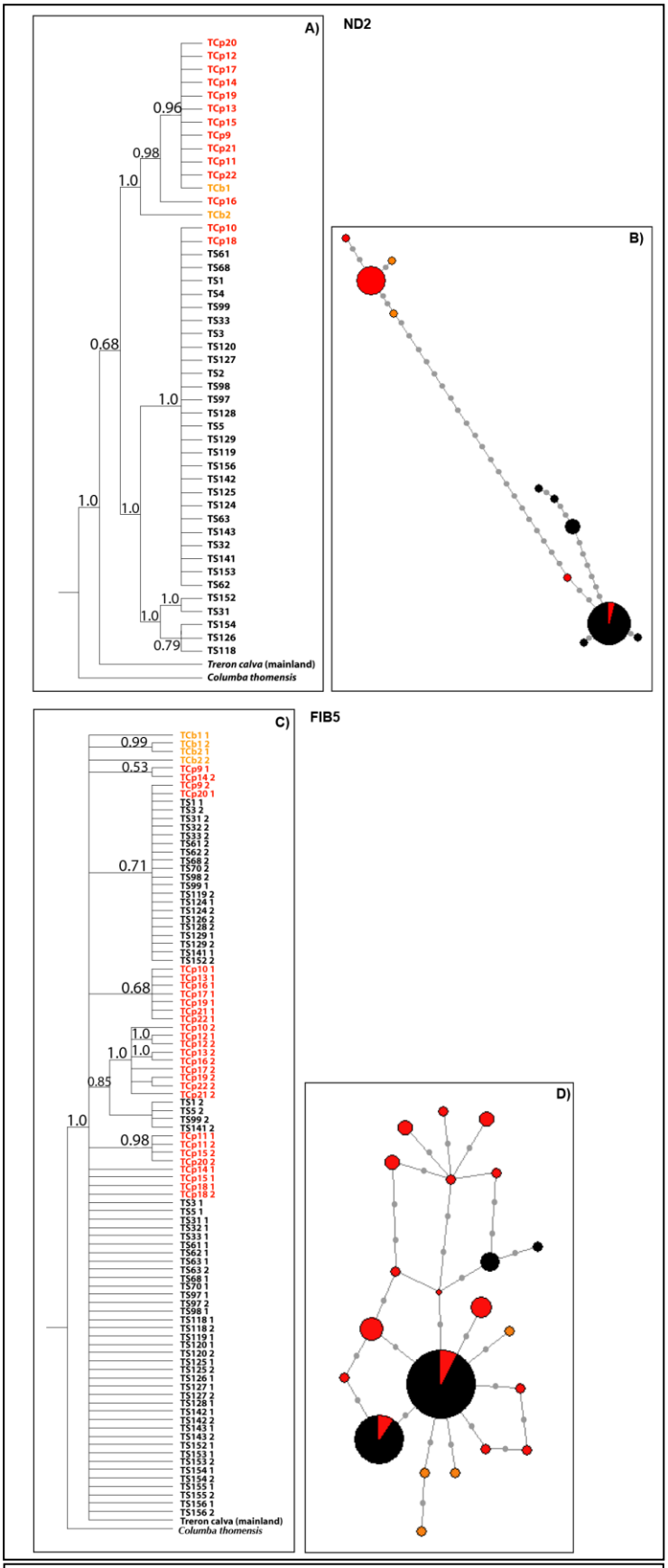


Figure 14. Phylogenetic and network trees of the *Treron* population. A) ND2 based phylogenetic tree; B) ND2 based network tree; C) FIB5 phased sequence phylogenetic tree; D) FIB5 phased sequence network tree. Individuals from *T. calva* (Príncipe) coloured in red; in orange individuals from *T. calva* (Bioko); in black individuals from *T. sanctihomae*.

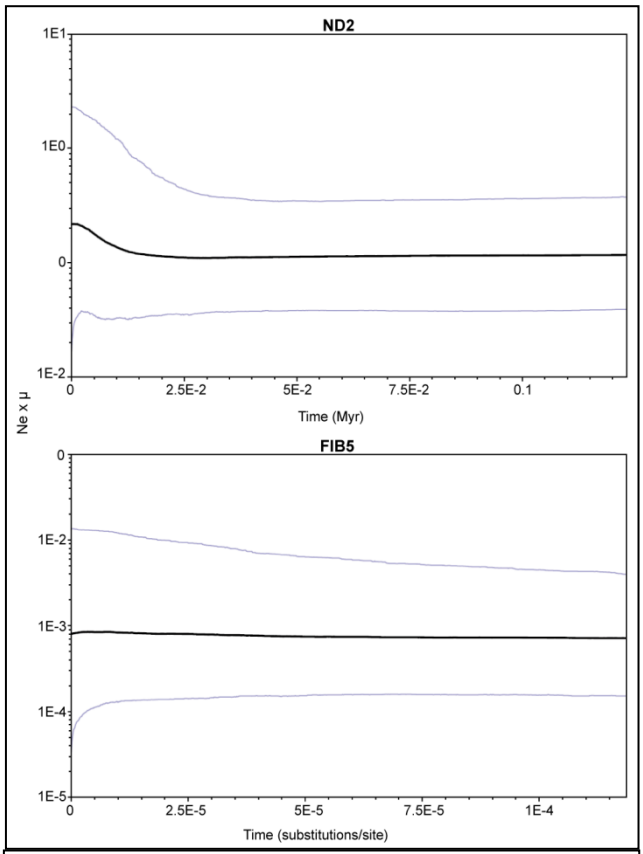


Figure 15. Bayesian skyline plot obtained for the *T. sanctihomae* population. The y axis is the product of the effective population size by the mutation rate ($N_e\mu$) and the x axis is in units of time (ND2: million years; FIB5: substitutions per site per million year). Blue lines correspond to the 95% HPD (highest posterior density) limits and black line

were conducted. The FIB5 analysis recovered no major haplotype differences with three very similar clusters (differing on average in one or two mutations; Fig.14[C,D].).

Migration rates were considerable for the *Treron* group (mean São Tomé→Príncipe: 71; mean Príncipe→São Tomé = 71). As with *C. malherbii*, the mean might be overestimated due the long tail of the posterior distribution. This odd distribution could be the result of the haplotypes of the two Príncipe birds that were most likely descendent from recent immigrants from São Tomé (TCp10, TCp18, Fig. 14[A,B].).The most common values were in fact quite low (mode:0.33), and below the one-migrant-per-generation rate that allows for population divergence to occur (Wright 1931; Spieth 1974; Tabel 9.).

DISCUSSION

Population structure

As expected for frugivorous birds confined to small areas, no population structure was identified within the islands. Fruiting trees are a resource that is patchily and unpredictably distributed which has led pigeons to adopt a nomadic or exploratory behaviour (del Hoyo *et al.* 1997). More surprisingly were the results showing that this nomadic behaviour may span an area including both Príncipe and São Tomé islands. This was particularly evident for *C. malherbii*, where the lack of population structure between islands was corroborated by the very high migration rates between islands. Considering these levels, one wonders if some regular, migratory-like, movements occur between São Tomé and Príncipe. Although rather far-fetched, this hypothesis could nevertheless explain the long-standing mystery of the virtual disappearance of *C. malherbii* from São Tomé after the breeding season.

In the case of the green pigeons, gene flow and migration events between islands appear to be rarer or accidental events. The ND2 haplotype network showed two very distinct ND2 haplotype groups, almost exclusively distributed by island (and hence by taxa). Migration between islands was low but, nevertheless, two individuals from *T. calva* grouped together with the *T. sanctithomae* cluster. This pattern is typical of rare secondary contact events between populations evolving otherwise in isolation. Although the phylogenetic and network analyses obtained from the nuclear marker did not depict such a clear differentiation, the test of differentiation still separated these two taxa. The fact that only two haplotypes were shared between the two island taxa (evolutionary distinct lineages) supports the idea that these two populations evolved separately in

isolation. The shallow differences found between haplotypes might be just a result of the low mutation rate of the nuclear marker (in comparison with the mitochondrial) and thus complete lineage sorting would take longer to occur (nDNA N_e 4x greater than mtDNA).

T. sanctithomae diversity was split into two distinct haplotype groups in the ND2 analysis. These had no geographic structure, being scattered throughout the island. This pattern is therefore likely to be a consequence of two separate, earlier, colonisations events. Alternatively, this could reflect incomplete sampling of intermediate haplotypes – either because they were missed or because they went extinct, or became very rare, due to the hunting pressure.

Population demography

Genetic tools, in particular when combined with coalescent methods, are useful for inferring the demographic history of a taxon group. Nevertheless, the time span captured by these methods will depict the history from thousands to millions of years ago. Thus the molecular markers used here will not be able to pick up the effect of the very recent hunting activity. Nevertheless, they can provide information on the demography of the populations before hunting started.

No bottleneck events could be detected by both markers. This can mean that either the endemic pigeons populations did not undergo any historical contraction, or can be just an artefact of the lack of resolution of the data allied with the difficulty in assessing such events in islands populations that can suffer and recover from multiple bottlenecks during their evolutionary history.

Results showed that the *C. thomensis* and *T. sanctithomae* populations have been stable during evolutionary times. The *C. malherbii* populations showed signs of a population expansion (in FIB5). Considering the rate of mutation of 0.135% substitution/site/myr proposed by Ellegren (2007) for intronic divergence in birds, this expansion can be dated around 3700 thousand years ago. The ND2 also identified a very weak sign of expansion of the *T. sanctithomae* population date at around 250 thousand years ago, although the large confidence intervals are more likely indicative of a stable population. In both cases, the putative expansions detected occurred a long time after colonisation (c. 2 Myr ago, Chapter 1), which could indicate that the populations only fully established on the island recently.

Levels of genetic diversity and estimates of the scaled effective population sizes were similar in *C. malherbii* and *T. sanctithomae*, with *C. thomensis* having considerably lower values. Island populations are mostly characterized by low genetic diversity if compared with their mainland counterparts. This is often explained as a consequence of founder

effects and occurrences of bottlenecks events (Habel & Zachos 2012). Nevertheless, it is also accepted that bottlenecks resulting from founder events are quickly overcome and should no longer be detectable, and anthropogenic impact may be the cause underlying current low levels of genetic diversity (Stuessy *et al.* 2012). Curiously the *Treron* populations of São Tomé and Príncipe showed similar Θ estimates. Being that the area of Príncipe is six times smaller than São Tomé, one would expect that Θ would be larger on São Tomé. This result might not depict any influence of hunting in the *T. sanctithomae* population, but rather it may still reflect population sizes from the end of the last glaciation (c. 11,0000 ago) when, during glacial times, the surface area of the Príncipe was about 10 times larger than today (Jones & Tye 2006).

CONCLUSION

Human colonization of the islands is a recent event (around 500 years ago) with hunting being a common practice probably from just around 100 to 200 years ago. Genetic data provided a view on the status of the populations before the novel impact of hunting. The genetic picture that was recovered for the endemic pigeons is typical of stable populations. After colonisation, these populations did not undergo any events of population contraction. On one hand, this could indicate that if hunting is having a strong impact on population decline, the populations may be undergoing a bottleneck from which they may have difficulty in recovering if not halted in time. On the other hand, the moderate levels of genetic diversity still present indicate that if measures controlling hunting are successfully implemented in the near future, these populations are likely to still retain the genetic capacity to recover their numbers.. In this regard, the situation of the *C. thomensis* populations showed some signals of concern due to its substantially lower genetic diversity and mutation-scaled effective population size.

The *Treron* group from the São Tomé and Príncipe islands remains an interesting case. This study showed a clear separation between the two populations, which seems to support their status as distinct species, and at the same time revealed that rare genetic exchanges add some complexity to their evolutionary history. Such rare genetic exchanges in oceanic archipelago settings have been shown to play an important role in adding genetic diversity to small, genetically depauperate, island populations (Arnold & Emms 1998; Grant & Grant 1998).

As a final remark, this study should be seen as a first step for our understanding of how the endemic pigeons of São Tomé are affected by hunting and how they may

recover from it. Further studies should focus on a stronger sampling effort and, in particular, on the use of faster evolving markers (microsatellites) that are able to recover a more recent history of these populations and provide more detailed information on demographic parameters that are required to guide conservation action.

CHAPTER 3.

FINAL REMARKS / CONSERVATION MEASURES

This study tried to shed some light on the evolutionary history of the Gulf of Guinea islands pigeons and recover genetic information relevant for guiding conservation action of the endemic pigeons populations of the São Tomé and Príncipe.

In the first chapter, the relationships between island and mainland taxa were analysed. Genetic data supported the species status for the *Columba malherbii* and *Columba thomensis*. Although the results showed a similar case for *Treron sanctithomae*, incomplete sampling from the mainland distribution prevented further considerations. Interesting results were also found regarding the *Aplopelia larvata* populations, with the São Tomé and Príncipe populations diverging at a species level from the mainland sample from Malawi. Nevertheless, as with *T. sanctithomae*, incomplete sampling of the mainland distribution prevented definitive conclusions. The unknown phylogenetic position of the *Aplopelia* and the bronze-naped pigeons was investigated, with these species grouping together with Old World *Columba* and the major clade of *Streptopelia* species, although the exact placement in relation to these two groups remained uncertain.

In the second chapter, a population genetics approach was used to infer past population dynamics of the endemic pigeons. Results showed no signs of population declines, with the *C. malherbii* population depicting apparent signs of population expansion. Basically, three stable populations were identified, which means that if the numbers observed in the field turn out to be accurate at least these populations still retain the genetic capacity to recover their numbers. This capacity is essential to make these species viable candidates for conservation projects.

Conservation efforts should focus first on hunting management. A rotational hunting system should be implemented to let the pigeons populations recover, and regular monitoring of the forests should be made to prevent the high levels of hunting during breeding season. Measures such as creating new laws on hunting (e.g., hunting licenses and quotas), the establishment of wildlife management areas, conservation awareness campaigns, should all be part of a strategy for managing the bushmeat trade and increasing the number of the endemic pigeons of São Tomé and Príncipe. The low number of juveniles recorded in the field suggest that reproductive success is low and, call for the need of a more intensive effort on nest monitoring and a better understanding of mortality rates. Hunting should be forbidden during the breeding season – which is currently the season of highest hunting activity.

As a last resource, conservation measures should focus on increasing genetic diversity and breeding success. Captive breeding and reintroduction of individuals into the wild may be an option at a future stage. A captive breeding program should aim to select healthy individuals that represent the genetic diversity of the population, and minimize effects of genetic drift. The second step in a genetic conservation project regards the choice of individuals to be introduced into the wild. Genetic data gives useful information on the origin, number and sex of the individuals. Candidates for reintroduction should maximise the reproductive success (and reduce inbreeding) and adaptive potential of the species (Frankham *et al.* 2002; Cruz 2011). Hopefully, a successfully implemented hunting management on the islands will prevent such drastic measures to be taken.

REFERENCES

Allendorf FW, Luikart G, Fourier J (2007) *Conservation and the Genetics of Populations*. Blackwell Publishing. Oxford, UK.

Amadon D, (1953) *Avian systematics and evolution in the Gulf of Guinea. Bulletin of the American Museum of Natural History*, **100**, 394-451.

Aris-Brosou S, Excoffier L (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. *Molecular Biology and Evolution*, **13**, 494-504.

Arnold ML, Emms SK (1998) Paradigm lost: natural hybridization and evolutionary innovations*. In: Endless Forms: Species and Speciation*. Howard DJ, Berlocher SH (eds). Oxford University Press, New York, USA.

Atkinson P, Dutton J, Peet N (1994) A study of the birds, small mammals, turtles and medicinal plants of Sao Tome with notes on Principe. *BirdLife International Study Report*, **56,** 1-106.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729-744.

Bandelt HJ, Forster P, Röhl A (1999) Median-Joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37-48.

Bensasson D, Zhang DX, Hartl DL, Hewitt GM (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution*, **16**, 314-321.

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, **152**, 763-773.

Beerli P (2009) How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, and C. Vernesi (eds). *Population Genetics for Animal Conservation*, volume 17 of Conservation Biology. Cambridge University Press, Cambridge, UK.

Beerli P, Palczewski M (2010) Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics*, **185**, 313-326.

Bibby CJ, Collar NJ, Crosby MJ, Heath MF, Imboden C, Johnson TH, Long AJ, Sattersfield AJ, Thirgood SJ (1992) *Putting Biodiversity on the Map: Priority Areas for Global Conservation.* International Council for Bird Preservation, Cambridge, UK.

BirdLife International. (2012) Species factsheet: *Columba thomensis*. Downloaded from http://www.birdlife.org.

BirdLife International. (2012) Species factsheet: *Columba malherbii*. Downloaded from http://www.birdlife.org.

BirdLife International. (2012) Species factsheet: *Treron sanctithomae.* Downloaded from http://www.birdlife.org.

BirdLife International. (2012) Species factsheet: *Aplopelia larvata*. Downloaded from http://www.birdlife.org.

BirdLife International (2012) *Columba thomensis*. In: IUCN (2012) IUCN Red List of Threatened Species. Version (2012)1.

BirdLife International (2012) *Columba malherbii*. In: IUCN (2012) IUCN Red List of Threatened Species. Version (2012)1.

BirdLife International (2012) *Treron sanctithomae*. In: IUCN (2012) IUCN Red List of Threatened Species. Version (2012)1.

Bowie RCK, Fjeldså J (2005) Genetic and morphological evidence for two species in the Udzungwa forest partridge *Xenoperdix udzungwensis*., *Journal of East African Natural History,* **94**, 191-201.

Buchanan GM, Donald PF, Butchart SHM (2011) Identifying priority areas for conservation: a global assessment for forest-dependent birds. *PLOSONE*, **12**, 316-320.

Caldeira R, Madeira J, Munhá JM, Afonso RS, Mata J, Tassinari CC, Nascimento E (2003) *Caracterização das principais unidades vulcano-estratigráficas da ilha de São Tomé, Golfo da Guiné*. Ciências da Terra (UNL), Lisboa, nº esp. V, CD-ROM, pp. A15-A18.

Carstens BC, Knowles LL (2007) Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. *Systematic Biology*, **56**, 400-411.

Chojnowski J, Kimball R, Braun E, (2008) Introns outperform exons in analyses of basal avian phylogeny using clathrin heavy chain genes. *Gene*, **410**, 89-96.

Chun-I J, Tsai LL, Hwa P, Chan C, Huang A, Chin S, Wang L, Lin J, Linacre A, Hsieh H, (2010) A novel strategy for avian species and gender identification using the CHD gene. *Molecular and Cellular Probes*, **24**, 27-31.

Coyne JA, Orr HA (2004) *Speciation.* Sinauer Associates, Sunderland, USA.

Cracraft J, (1989) Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. Otte D, Endler JA (eds) *Speciation and its Consequences*: 28–59. Sinauer Associates, Sunderland, USA.

Cruz BM, (2011) Conservation genetics of Iberian raptors. *Animal Biodiversity and Conservation*, **34**, 341-353.

de Lima RF, Dallimer M, Atkinson P, Barlow J (2013) Biodiversity and land-use change: understanding the complex responses of an endemic-rich bird assemblage. *Diversity and Distributions*, **19**, 411-422.

Dallimer M, King T (2008) Habitat preferences of the forest birds on the island of Príncipe, Gulf of Guinea. *African Journal of Ecology*, **46**, 258-266.

Dallimer M, King T (2009) Pervasive threats within a protected area: conserving the endemic birds of São Tomé, West Africa. *Animal Conservation*, **12**, 209-219.

del Hoyo J, Elliott A, Sargatal J (1997) *Handbook of the birds of the world. Vol.4. Sandgrouse to Cuckoos.* Lynx Editions, Barcelona, Spain.

Desjardins P, Morais R (1990) Sequence and gene organization of the chicken mitochondrial genome: a novel gene order in higher vertebrates. *Journal of Molecular Evolution,* **212**, 599-634.

Dowsett RJ, Dowsett-Lemaire F (1993) *A Contribution to the Distribution and Taxonomy of Afrotropical and Malagasy Birds.* Tauraco Press, Liège, Belgium.

Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution,* **22**, 1185-92.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969-1973.

Edwards SV (2009) Is a new and general theory of molecular systematics emerging? *Evolution*, **63**, 1-19.

Ellegren H (2007) Molecular evolutionary genomics of birds. *Cytogenetic Genome Resources*, **117**, 120-130.

Exell AW (1944) *Catalogue of the Vascular Plants of S. Tomé (with Príncipe and Annobon).* British Museum (Natural History), London, UK.

Exell, AW (1973) Angiosperms of the islands of the Gulf of Guinea (Fernando Pó, Príncipe, S. Tomé and Annobón). *Bulletin of the British Museum (Natural History) Botany*, **4**, 325-411.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application. *Genetics Society of America*, **131**, 479-491.

Excoffier L, Lischer HEL (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.

Fain MG, Houde P (2004) Parallel radiations in the primary clades of birds. *Evolution*, **58**, 2558-2573.

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368-376.

Frankham, R (1995) Conservation genetics. *Annual Review of Genetics*, **29**, 305-327.

Frankham R (1998) Inbreeding and extinction: island populations. *Conservation Biology*, **12**. 665-675.

Frankham R (2002) Genetics and conservation biology. *Comptes Rendus Biologies*, **326**, S22-S29.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915-925.

Garnett ST, Christidis L (2007) Implications of changing species definitions for conservation purposes. *Bird Conservation International*, **17**, 187-195.

Gascoigne A, (2004) São Tomé, Príncipe, and Annobon Moist Lowland Forests. *In: Terrestrial Ecoregions of Africa and Madagascar: a Conservation Assessment*. Burgess N, D'Amico Hales J, Underwood E et al. (eds), pp. 236-238. Island Press, Washington, USA.

Geneious version R6 created by Biomatters. Available from http://www.geneious.com/.

Gibbs D, Barnes E, Cox J (2001) *Pigeons and Doves: A Guide to the Pigeons and Doves of the World*. Sussex: Pica Press, London, UK.

Goodwin D (1983) *Pigeons and Doves of the World*. Cornell University Press, New York, USA.

Gonzalez J, Castro GD, Garcia-del-Rey E, Berger C, Wink M (2009) Use of mitochondrial and nuclear genes to infer the origin of two endemic pigeons from the Canary Islands. *Journal of Ornithology*, **150**, 357-367.

Grant PR, Grant BR (1998) Speciation and hybridization of birds on islands. *In: Evolution on Islands.* Grant PR (eds). Oxford University Press, Oxford, USA.

Griffiths R, Double MC, Orr K, Dawson RJG (1998) A DNA test to sex most birds. *Molecular Ecology*, **7**, 1071-1075.

Habel JC, Zachos FE, (2012) Past population history versus recent population decline – founder effects in island species and their genetic signatures. *Journal of Biogeography*, **40**, 206-207.

Hackett S, Kimball R, Reddy S, Bowie R, Braun E, Braun M, Chojnowski J, Cox A, Han K (2008) Phylogenomic Study of Birds Reveals Their Evolutionary History. *Science*, **320**, 1763-1768.

Hafner MS, Sudman PD, Villablanca FX, Sprading TA (1994) Disparate Rates of Molecular Evolution in Cospeciating Hosts and Parasites. *Science*, **265**, 1087-1090.

Helbig AJ, Knox AG, Parkin DT, Sangster G, Collinson M (2002) Guidelines for assigning species rank. *Ibis*, **144**, 518-525.

Heled J, Drummond AJ, (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570-580.

Hey J, (2001) The mind of the species problem. *Trends in Ecology and Evolution*, **16**, 326-329.

Hey J (2010) With migration models for more than two populations. *Molecular Biology and Evolution*, **27**, 905-920.

Irestedt M, Fjeldså J, Nylander, JAA, Ericson PGP (2004) Phylogenetic relationships of typical antbirds (*Thamnophilidae*) and test of incongruence based on Bayes factors. *BMC Evolutionary Biology*, **4**, 1-16.

Johns GC, Avise JC, (1998) A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome b gene. *Molecular Biology and Evolution*, **15**, 1481-1490.

Johnson KP, Sorenson MD (1998) Comparing Molecular Evolution in Two Mitochondrial Protein Coding Genes (Cytochrome b and ND2) in the Dabbling Ducks (Tribe: *Anatini*). *Molecular Phylogenetics and Evolution*, **1**, 82-94.

Johnson KP, Clayton DH (2000) Nuclear and mitochondrial genes contain similar phylogenetic signal for pigeons and doves (Aves: *Columbiformes*). *Molecular Phylogenetics and Evolution*, **14**, 141-151.

Johnson KP, De Kort S, Dinwoodey K, Mateman AC, Ten CC, Lessells CM, Clayton DH (2001) A molecular phylogeny of the dove genera *Streptopelia* and *Columba*. *Auk*, **118**, 874-887.

Jones PJ, Tye A (2006) *The Birds of São Tomé and Príncipe, with Annobón: Islands of the Gulf of Guinea*. British Ornithologists' Union, Oxford, UK.

Kass R, Raftery AE (1995) Bayes Factors. *Journal of the American Statistical Association*, **430**, 773-795.

Kingman JFC (1982) The coalescent. *Stochastic Processes and their Applications,* **13**, 235-248.

Knowles LL, Kubatko LS (2010) *Estimating species trees: an introduction to concepts and models*. Knowles LL, Kubatko LS (eds). Wiley-Blackwell. pp 1-14.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Procedings of the National Academy of Science of the United States of America*. **86**, 6196-6200.

Koressaar T, Remm M, (2007) Enhancements and modifications of primer design program Primer3. *Bioinformatics*, **23**, 1289-1291.

Lambert K, Chappel J (2001) Sea level change through the last glacial cycle. *Science*, **292**, 679-686.

Lanfear R, Calcott B, Ho S, Guindon S (2012) PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Molecular Biology and Evolution*, **26**, 1695-1701.

Lee DC, Halliday AN, Fitton JG, Poli G (1994) Isotopic variations with distance and time in the volcanic islands of the Cameroon line: evidence for a mantle plume origin. *Earth and Planetary Science Letters*, **123**, 119-138.

Lepage T, Bryant D, Philippe H, Lartillot N (2007) A general comparison of relaxed molecular clock models. *Molecular Biology and Evolution*, **24**, 26692680.

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

Li N, Stephens M (2003) Modeling Linkage Disequilibrium and Identifying Recombination Hotspots Using Single-Nucleotide Polymorphism Data. *Genetics,* **165**, 2213-2233.

Mallet J (2008) Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B-Biological Sciences,* **363**, 2971-2986.

Mayr E (1942) *Systematics and the Origin of Species.* Columbia University Press, New York, USA.

McConkey KR, Drake DR (2002) *Extinct Pigeons and Declining Bat Populations: Are Large Seeds Still Being Dispersed in the Tropical Pacific?* CABI Publications, Wallingford, UK.

Melo M (2007) *Bird Speciation in the Gulf of Guinea*. PhD Thesis. Institute of Evolutionary Biology, University of Edinburgh, UK.

Melo M, O'Ryan C (2007) Genetic differentiation between Príncipe Island and mainland populations of the grey parrot (*Psittacus erithacus*), and implications for conservation. *Molecular Ecology,* **16**, 1673-1685.

Melo M, Jones PJ (2008) *Bird speciation in the Gulf of Guinea island system*. Craig DM, Anderson AJFK, Rakotomanana MD, Muchai HM (eds). Proceedings of the 12th Pan-African Ornithological Congress, Cape Town, South Africa.

Melo M (2009) As aves endémicas de São Tomé e Príncipe. *Quercus Ambiental*, **36**, 16-18.

Melo M, Bowie RCK, Voelker G, Dallimer M, Collar NJC, Jones PJ (2010) Multiple lines of evidence support the recognition of a very rare bird species: the Príncipe thrush. *Journal of Zoology*, **282**, 120-129.

Melo M, Warren B, Jones P (2011) Rapid parallel evolution of aberrant traits in the diversification of the Gulf of Guinea white-eyes (Aves, *Zosteropidae*). *Molecular Ecology*, **20**, 4953-4967.

OboPark.com. (2012) Obo Natural Parks. http://www.obopark.com/pt/saotomeprincipe/nationalparks.html. Accessed June 2013.

Miller MA, Pfeiffer W, Schwartz T (2010) *Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees*. Proceedings of the Gateway Computing Environments Workshop (GCE), pp1-8, New Orleans, USA.

Myers N, (1979) Islands of conservation. *New Scientist*, **83**, 600-602.

Naurois Rd (1994) *Les Oiseaux des Iles du Golfe de Guinée / As Aves das Ilhas do Golfo da Guiné*. Instituto de Investigação Científica Tropical, Lisboa, Portugal.

Nei M (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York, USA.

Nylander J, Ronquist F, Huelsenbeck JP, Nieves-Aldrey JL (2004) Bayesian phylogenetic analysis of combined data. *Systematic Biology,* **53**, 47-67.

Palumbi SR (1996) *Nucleic acids II: the polymerase chain reaction.* Hillis DM, Moritz C, Mable BK (eds). Molecular Systematics, Sinauer Associates, Sunderland, USA.

Peet N, Atkinson P (1994) The biodiversity and conservation of the birds of São Tomé and Príncipe. *Biodiversity and Conservation*, **3**, 851-867.

Pereira SL, Johnson KP, Clayton DH, Baker AJ (2007) Mitochondrial and nuclear DNA sequences support a Cretaceous origin of Columbiformes and a dispersal-driven radiation in the Paleogene. *Systematic Biology*, **56**, 656-672.

Price T (2008) *Speciation in Birds*. Roberts and Company, Greenwood Village, Colorado, USA.

Prychitko TM, Moore WS (1997) The utility of DNA sequences of an intron from the β-fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: *Picidae*). *Molecular Phylogenetics and Evolution*, **8**, 193-204.

Prychitko TM, Moore WS (2003) Alignment and phylogenetic analysis of beta-fibrinogen intron 7 sequences among avian orders reveal conserved regions within the intron. *Molecular Biology and Evolution*, **20**, 762-771.

Pybus G, Rambaut A, Harvey PH (2000) An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics*, **155**, 1429-1437.

Rambaut A, Drummond AJ (2007) Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer

Rannala B, Yang Z (2003) Bayes Estimation of Species Divergence Times and Ancestral Population Sizes Using DNA Sequences From Multiple Loci. *Genetics*, **164**, 1645-1656.

Raymond M, Rousset F (1995) An exact test of population differentiation. *Evolution*, **49**, 1280-1283.

Rohling EJ, Fenton M, Jorissen FJ, Bertrand P, Ganssen G, Caulet JP (1998) Magnitudes of sea level lowstands of the past 500,000 years. *Nature*, **394**, 162-165.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572-1574.

Ronquist F, Klopfstein S, Vilhelmsen L, Schulmeister S, Murray DL, Rasnitsyn AP (2012) A total-evidence approach to dating with fossils, applied to the early radiation of the *Hymenoptera*. *Systematic Biology*, **6**, 973-999.

Sanderson MJ, Shaffer HB (2002) Troubleshooting molecular phylogenetic analyses. *Annual Review of Ecology and Systematics*, **30**, 49-72.

Spieth PT (1974) Gene flow and genetic differentiation. *Genetics*, **78**, 961-965.

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. *Molecular Phylogenetics and Evolution*, **12**, 105-114.

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

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *The American Journal of Human Genetics,* **73**, 1162-1169.

Stuessy TF, Takayama K, Sepulveda PL (2012) Founder effects are invisible in endemic species of oceanic islands. *Journal of Biogeography*, **39**, 1565–1566.

Swofford DL (2003) *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer Associates, Sunderland, USA.

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

Tavare S (1984) Line-of-descent and genealogical processes, and their applications in population genetics models. *Theoretical Population Biology*, **26**, 119-164.

Tobias JA, Seddon N, Spottiswoode CN, Pilgrim GD, Fishpool LDC, Collar NJ (2010) Quantitative criteria for species delimitation. *Ibis*, **152**, 724-746.

Unesco.org. 2013. http://www.unesco.org/new/en/media-services/multimedia/photos/mab-2012/sao-tome-e-principe/Accessed June 2013

Untergrasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3 - new capabilities and interfaces. *Nucleic Acids Research*, **40**, 1-12.

Walker J (2007) Geographical patterns of threat among pigeons and doves (*Columbidae*). *Oryx*, **41**, 289-299.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256-276

Weir JT, Schluter D (2008) Calibrating the avian molecular clock. *Molecular Ecology*, **17**, 2321-2328.

Wiens JJ, Kuczynski CA, Smith SA, Mulcahy DG, Sites JW (2008) Branch lengths, support, and congruence: testing the phylogenomic approach with 20 nuclear loci in snakes. *Systematic Biology*, **57**, 420-431.

Winker K, Rocque DA, Braile TM, Pruett CL (2007) Vainly beating the air: species-concept debates need not impede progress in science or conservation. *Ornithology Monographs*, **63**, 30-44.

Wright S (1931) Evolution in mendelian populations. *Genetics*, **16**, 97-159.

Yuri T, Mindell DP (2002) Molecular phylogenetic analysis of *Fringillidae*, “New World nine-primaried oscines” (Aves: *Passeriformes*). *Molecular Phylogenetics and Evolution*, **29**, 229-243.

Zwickl DJ (2006) *Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets Under the Maximum Likelihood Criterion*. PhD Thesis. Austin, University of Texas.

APPENDICES

Appendix 1. Previously published sequences obtained from Genbank

		Genbank assession number				
Taxon	**Locality**	**ND2**	**Cytb**	**COI**	**FIB7**	**Source**
Columba livia	Utah	AF353433	AF182694	AF279733	AF182661	
Columba rupestris	Mongolia	AF353434	AF353410	AF353482	AF353461	
Columba guinea	South Africa	AF353435	AF279706	AF279732	AF279718	
Columba palumbus	Captive	AF353436	AF353411	AF353483	AF353462	
Columba arquatrix	South Africa	AF353437	AF353412	AF353484	AF353463	
Columba pulchrichollis	Captive	AF353438	AF353413	AF353485	AF353464	
Columba leucocephala	Florida	AF353441	AF182689	AF353488	AF182656	
Streoptopelia decaocto	Netherlands	AF353418	AF353398	AF353469	AF353449	
Streptopelia roseogrisea	Cameroon	AF353419	AF353399	AF353470	AF353450	
Streptopeliadecipiens	Cameroon	AF353420	AF353400	AF353471	AF353451	
Streptopelia semitorquata	South Africa	AF353421	AF353401	AF353472	AF353452	
Streptopelia capicola	South Africa	AF353422	AF279709	AF279734	AF279719	
Streptopelia vinacea	Cent. Afr. Rep	AF353423	AF353402	AF353473	AF353453	
Streptopelia hypopyrrha	Cameroon	AF353424	AF353403	AF353474	AF353454	
Streptopelia turtur	Kazakhastan	AF353425	AF353404	AF353475	AF353455	Johnson *et al.* 2001
Streptopelia orientalis	Russia	AF353426	AF353405	AF353476	AF353456	
Streptopelia bitorquata	Captive	AF353427	AF353406	AF353477	AF353457	
Streptopelia tranquebarica	Captive	AF353428	AF353407	AF353478	AF353458	
Streptopelia picturata	Madagascar	AF353430	AF353409	AF353480	AF353460	
Streptopelia chinensis	Philippines	AF353431	AF182695	AF353481	AF182662	
Streptopelia senegalensis	South Africa	AF353432	AF279710	AF279735	AF279720	
Nesoenas mayeri	Captive	AF353429	AF353408	AF353479	AF353459	
Patagioenas plumbea	Brazil	AF251547	AF182691	AF279736	AF182658	
Patagioenas subvinacea	Peru	AF353439	AF182692	AF353486	AF182659	
Patagioenas oenops	Peru	AF353440	AF182690	AF353487	AF182657	
Patagioenas speciosa	Mexico	AF353442	AF279711	AF279737	AF279721	
Patagieonas fasciata	Utah	AF353443	AF353414	AF353489	AF353465	
Macropygia mackinlayi	Solomon Islands	AF353444	AF353415	AF353490	AF353466	
Macropygia tenurirostris	Philippines	AF353445	AF353416	AF353491	AF353467	
Reinwardtoena browni	Captive	AF353446	AF353417	AF353492	AF353468	
Zenaida asiatica	Arizona	AF251543	AF251533	AF279731	AF258324	
Zenaida macroura	Arizona	AF251535	AF251530	AF353493	AF258321	
Geotrygon montana	Peru	AF353447	AF182696	AF279728	AF182663	
Leptotila rufaxilla	Peru	AF251546	AF182698	AF353494	AF182665	
Leptotila verreauxi	Texas	AF353448	AF279705	AF279725	AF279715	
Pterocles namaqua	–	DQ385080	DQ385216	DQ385165	EU739477	*

*Hackett *et al.* (2008)

Appendix 2. Primers and PCR conditions for all mitochondrial markers and the nuclear intron (FIB7) used in chapter 1.

					PCR-mix				PCR-cycle					
Gene	Primer name[1]	Primer sequence 5'-3'	Product (bp)	Source	Primer (µM)	Mg (mM)	dNTP (mM)	Taq (U)	[PREMIT]	[DENAT	ANNEAL	EXT]	CYCLES	FINAL
ND2-1	L5219	**F:** CCCATACCCCGAAAATGATG	547	1	0.2	2	0.2	1.6	95°-15'	95°-30''	58/55°-30''	72°-90''	11+26	72°-10'
	H5766	**R:** GGATGAGAAGGCTAGGATTTTKCG		1										
ND2-2	L5758	**F:** GGCTGAATRGGMCTNAAYCARAC	555	1										
	H6313	**R:** CTCTTATTTAAGGCTTTGAAGGC		1										
Cytb	L14764	**F:** TGRTACAAAAAAATAGGMCCMGAAGG	1302	2	0.2	2	0.2	1.6	95°-15'	95°-30''	56°-30''	72°-90''	37	72°-10'
	H4a	**R:** AAGTGGTAAGTCTTCAGTCTTTGGTTTACAAGACC		3										
COI	L6625	**F:** CCGGATCCTTYTGRTTYTTYGGNCAYCC	381	4	0.2	2	0.2	1.6	95°-15'	95°-30''	57°-30''	72°-75''	37	72°-10'
	H7005	**R:** CCGGATCCACNACRTARTANGTRTCRTG		4										
FIB7	Fib7U	**F:** GGAGAAAACAGGACAATGACAATTCAC	963	5	0.2	2	0.2	1.6	95°-15'	95°-30''	56°-30''	72°-75''	37	72°-10'
	Fib7L	**R:** TCCCCAGTAGTATCTGCCATTAGGGTT		5										

[1]:Primer name given as the position of the 3' end base on the chicken mtDNA genome (Desjardins & Morais 1990) - L: light strand, H: heavy strand;
Primer name as in source - F: forward, R: reverse. PREMIT: Initial denaturation; DENAT: Denaturation; ANNEAL: Annealing; EXT: Extension; FINAL: Final extension.
SOURCE: 1 - Sorenson *et al.* (1999); 2 - Kocher (1989); 3 - Hafner *et al.* (1994); 4 - Prychitko & Moore (1997).

Appendix 3. Primers and PCR conditions for all markers used in the amplification of the museum samples

					PCR-mix				PCR-cycle					
Gene	Primer name[1]	Primer sequence 5'-3'	Product (bp)	Source	Primer (µM)	Mg (mM)	dNTP (mM)	Taq (U)	[PREMIT]	[DENAT	ANNEAL	EXT]	CYCLES	FINAL
ND2-1	L5219	**F:** CCCATACCCCGAAAATGATG	547	1	0.2	2	0.2	1.6	95°-15'	95°-30''	58/55°-40''	72°-90''	11+34	72°-10'
	H5766	**R:** GGATGAGAAGGCTAGGATTTTKCG		1										
ND2-2	L5758	**F:** GGCTGAATRGGMCTNAAYCARAC	555	1										
	H6313	**R:** CTCTTATTTAAGGCTTTGAAGGC		1										
Cytb-1	L14841	**F:** AAAAAGCTTCCATCCAACATCTCAGCATGATGAAA	366	2	0.2	2	0.2	1.6	95°-15'	95°-30''	56°-30''	72°-90''	45	72°-10'
	H15298	**R:**CCCTCAGAATGATATTTGTCCTCA		3										
Cytb-2	L5517	**F:**CACGAATCAGGCTCAAACAACC	178	4										
	H5649	**R:** GCTGGTGTGTGAAGTTTTCTGGGTC		5										
COI	L6625	**F:** CCGGATCCTTYTGRTTYTTYGGNCAYCC	381	6	0.2	2	0.2	1.6	95°-15'	95°-30''	57°-35''	72°-75''	45	72°-10'
	H7005	**R:** CCGGATCCACNACRTARTANGTRTCRTG		6										
FIB7-1	Fib7L	**R:** TCCCCAGTAGTATCTGCCATTAGGGTT	289	7	0.2	2	0.2	1.6	95°-15'	95°-30''	50°-30''	72°-75''	45	72°-10'
	FIB7_R1	**R:** GACATTCCTAAAGAGATGC		Designed for this study										
FIB7-2	FIB7_F2	**F:** GTTTAGCTGCATCTCTTTAGG	279											
	FIB7_R2	**R:** CTCACTTCTGTAAGACATGGG												
FIB7-3	FIB7_F3	**F:** GAGCTGTCTTCTAAGTAGGC	199											
	FIB7_R3	**R:** CGGTCTTTCTGTTAGCTTTCACAGC												
FIB7-4	FIB7_F4	**F:** GGATATTTGGCTGGTAACTTGC	268											
	FIB7U	**F:** GGAGAAAACAGGACAATGACAATTCAC		7										

[1]:Primer name given as the position of the 3' end base on the chicken mtDNA genome (Desjardins & Morais 1990) - L: light strand, H: heavy strand;
Primer name as in source - F: forward, R: reverse. PREMIT: Initial denaturation; DENAT: Denaturation; ANNEAL: Annealing; EXT: Extension; FINAL: Final extension.
SOURCE: 1 - Sorenson *et al.* (1999); 2 - Kocher (1989); 3- Palumbi (1996); 4 - Johnson (1998); 5 - Yuri & Mindel (2002); 6 - Hafner *et al.* (1994); 7 - Prychitko & Moore (1997).

Appendix 4. Pairwise distances (%) calculated from the mtDNA concatenated dataset (p' uncorrected and corrected with GTR model respectively).. Initials for genera names: A: *Aplopelia*; C:*Columba*; T:*Treron*.

	A_larvata (mainland)	A_larvata P1	A_larvata P26	A_larvata ST14
A_larvata (mainland)				
A_larvata P1	4.3-4.5			
A_larvata P26	4.3-4.5	0.0-0.0		
A_larvata ST14	4.3-4.5	0.3-0.3	0.3-0.3	
A_larvata ST19	4.3-4.5	0.3-0.3	0.3-0.3	0.0-0.0

	C_delegorguei	C_iriditorques	C_malherbii_ST7	C_malherbii_ST5	C_malherbii_P4
C_delegorguei					
C_iriditorques	2.3-2.4				
C_malherbii_ST7	2.8-2.8	2.2-2.2			
C_malherbii_ST5	2.8-2.8	2.1-2.1	0.1-0.1		
C_malherbii_P4	2.7-2.8	2.1-2.1	0.1-0.1	0.1-0.1	
C_malherbii_P5	2.8-2.8	2.2-2.2	0.1-0.1	0.1-0.1	0.1-0.1

	C_arquatrix	C_thomensis_1
C_arquatrix		
C_thomensis_1	2.7-2.8	
C_thomensis_2	2.8-2.8	0.2-0.2

	T_calva (mainland)	T_calva_B1	T_calva_B2	T_calva_P9	T_calva_P10	T_calva_P11	T_sanctithomae_1
T_calva (mainland)							
T_calva_B1	3.1-3.2						
T_calva_B2	2.5-2.5	1.1-1.1					
T_calva_P9	2.6-2.6	0.5-0.5	0.2-0.2				
T_calva_P10	3.5-3.6	2.6-2.6	1.6-1.6	1.9-1.9			
T_calva_P11	2.8-2.9	0.3-0.3	0.8-0.8	0.2-0.2	2.3-2.4		
T_sanctithomae_1	3.5-3.6	2.7-2.7	1.7-1.7	2.0-2.0	0.2-0.2	2.5-2.5	
T_sanctithomae_3	3.5-3.6	2.7-2.7	1.7-1.7	2.0-2.0	0.1-0.1	2.5-2.5	0.1-0.1

Appendix 5. Pairwise distances (%) calculated from FIB7 intron (p' uncorrected and corrected with GTR model respectively). Initials for genera names: A: *Aplopelia*; C:*Columba*; T:*Treron*.

	A_larvata (mainland)	A_larvata_P1	A_larvata_P26	A_larvata_ST14
A_larvata (mainland)				
A_larvata_P1	0.85-0.86			
A_larvata_P26	0.85-0.86	0.60-0.60		
A_larvata_ST14	0.85-0.86	0.60-0.60	0.00-0.00	
A_larvata_ST19	0.85-0.86	0.72-0.73	0.84-0.85	0.84-0.85

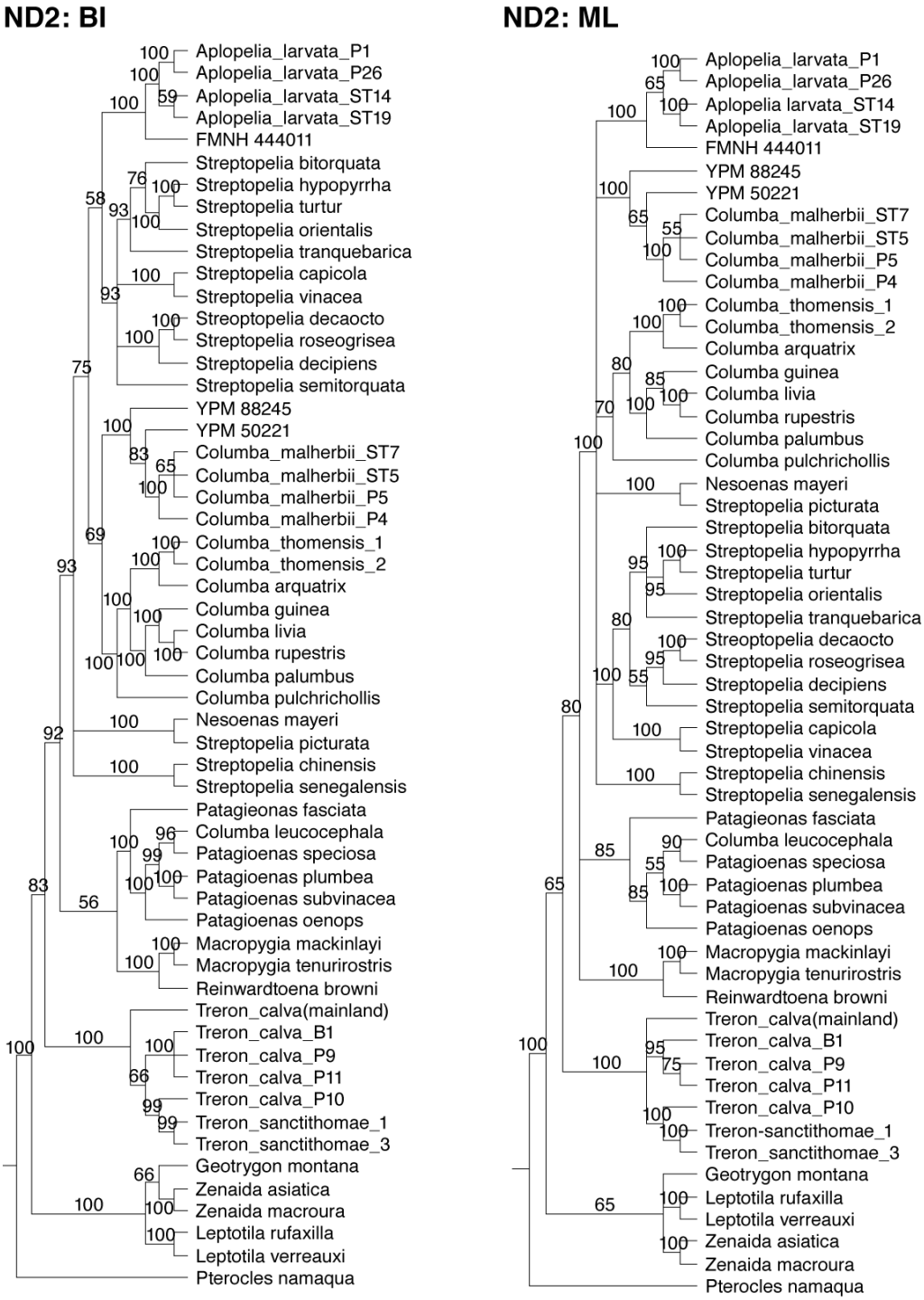
	C_delelgorguei	C_iriditorques	C_malherbii_P4	C_malherbii_P5	C_malherbii_ST5
C_delelgorguei					
C_iriditorques	0.12-0.12				
C_malherbii_P4	0.71-0.72	0.59-0.60			
C_malherbii_P5	0.71-0.72	0.59-0.60	0.00-0.00		
C_malherbii_ST5	0.59-0.60	0.48-0.48	0.12-0.12	0.12-0.12	
C_malherbii_ST7	0.59-0.60	0.48-0.48	0.12-0.12	0.12-0.12	0.00-0.00

	C_arquatrix	C_thomensis_1
C_arquatrix		
C_thomensis_1	0.49-0.49	
C_thomensis_2	0.49-0.49	0.00-0.00

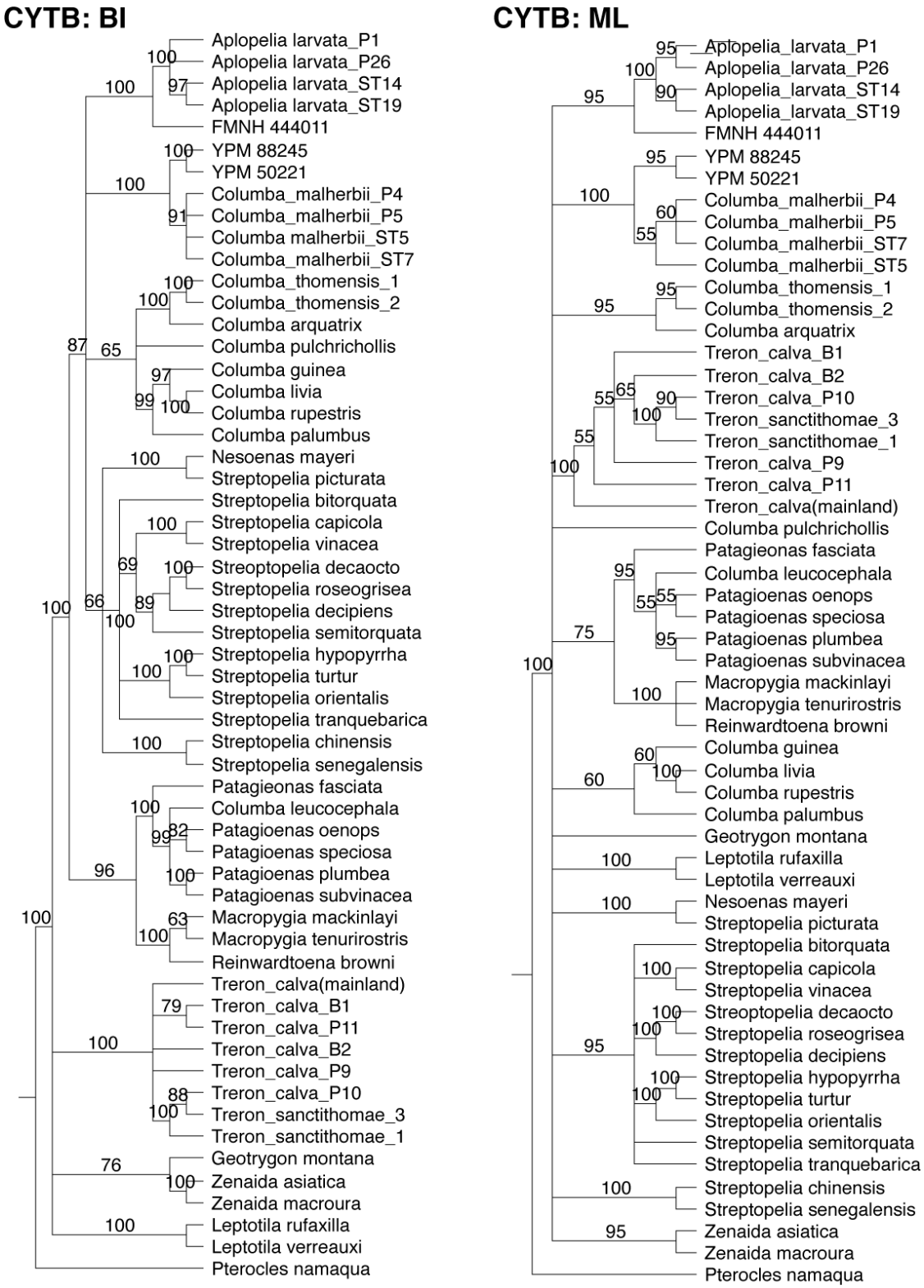
	T_calva(mainland)	T_calva_B1	T_calva_B2	T_calva_P9	T_calva_P10	T_calva_P11	T_sanctithomae_1
T_calva (mainland)							
T_calva_B1	0.69-0.70						
T_calva_B2	0.59-0.60	0.10-0.10					
T_calva_P9	0.59-0.60	0.10-0.10	0.00-0.00				
T_calva_P10	0.89-0.90	0.40-0.40	0.30-0.30	0.30-0.30			
T_calva_P11	0.59-0.60	0.10-0.10	0.00-0.00	0.00-0.00	0.30-0.30		
T_sanctithomae_1	0.79-0.80	0.30-0.30	0.20-0.20	0.20-0.20	0.10-0.10	0.20-0.20	
T_sanctithomae_3	0.69-0.70	0.20-0.20	0.10-0.10	0.10-0.10	0.20-0.20	0.10-0.10	0.10-0.10

Appendix 6. Phylogenies based on individual genetic markers analysed with Bayesian inference (BI) and maximum likelihood (ML). Bayesian posterior probabilities and ML bootstrap values ≥0.5 are presented as percentage at nodes.

Appendix 6.1

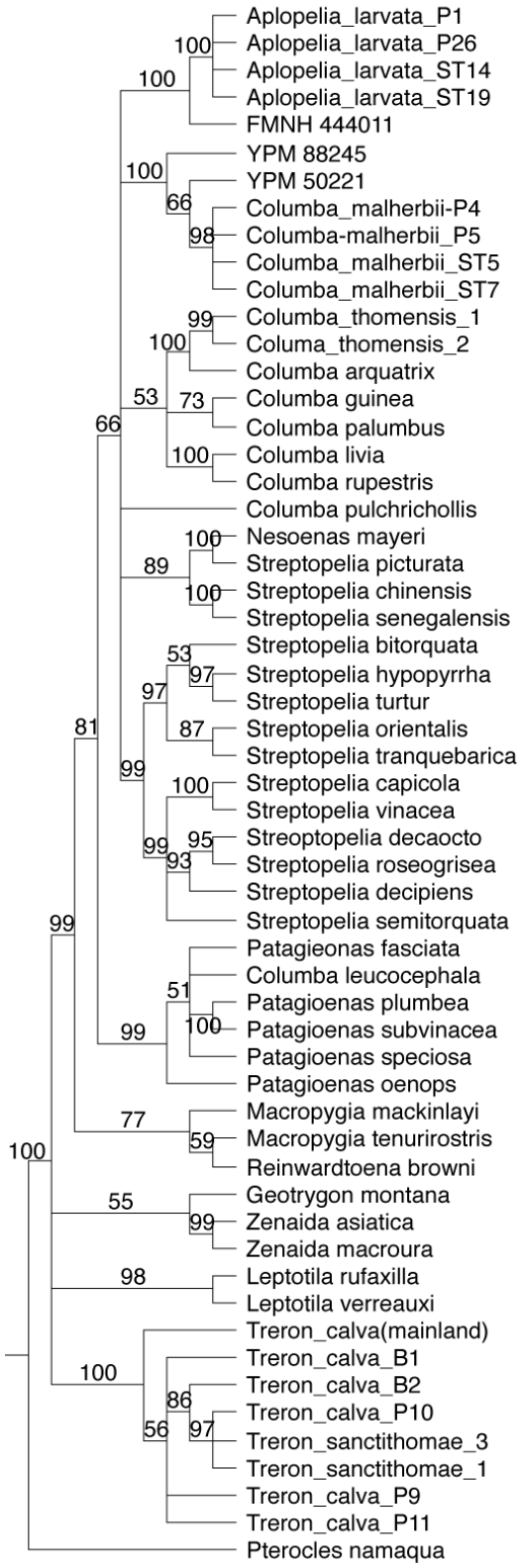


Appendix 6.2

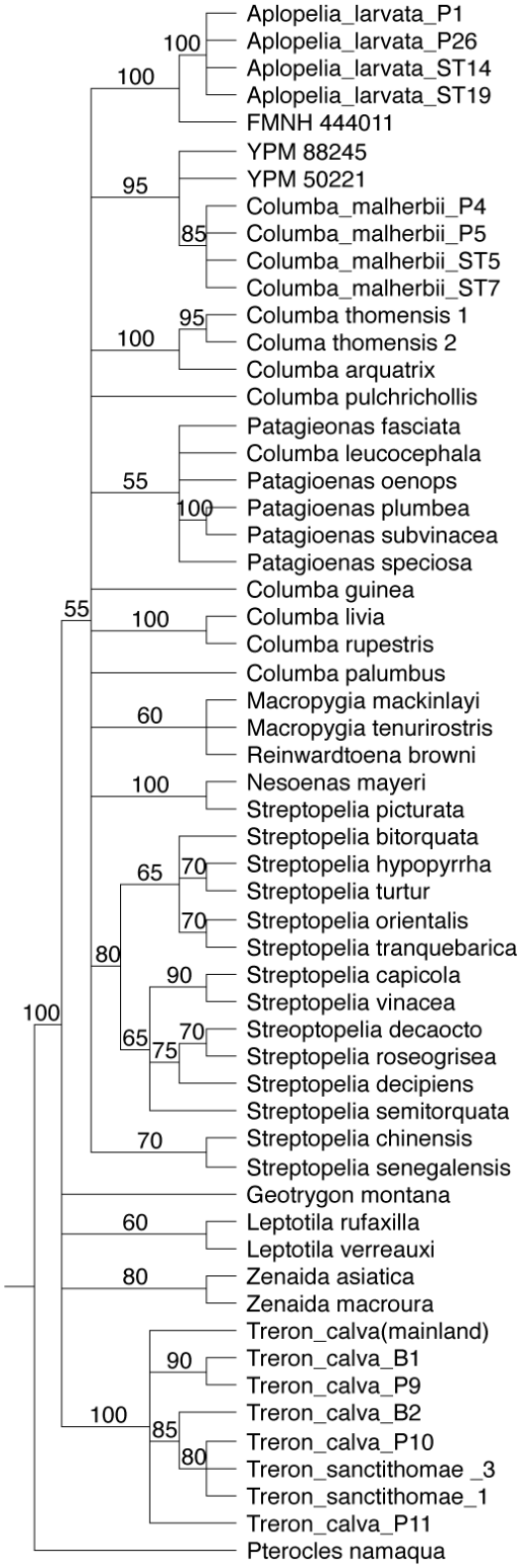

CYTB: BI
Aplopelia larvata_P1
Aplopelia larvata_P26
Aplopelia larvata_ST14
Aplopelia larvata_ST19
FMNH 444011
YPM 88245
YPM 50221
Columba_malherbii_P4
Columba_malherbii_P5
Columba malherbii_ST5
Columba_malherbii_ST7
Columba_thomensis_1
Columba_thomensis_2
Columba arquatrix
Columba pulchrichollis
Columba guinea
Columba livia
Columba rupestris
Columba palumbus
Nesoenas mayeri
Streptopelia picturata
Streptopelia bitorquata
Streptopelia capicola
Streptopelia vinacea
Streoptopelia decaocto
Streptopelia roseogrisea
Streptopelia decipiens
Streptopelia semitorquata
Streptopelia hypopyrrha
Streptopelia turtur
Streptopelia orientalis
Streptopelia tranquebarica
Streptopelia chinensis
Streptopelia senegalensis
Patagieonas fasciata
Columba leucocephala
Patagioenas oenops
Patagioenas speciosa
Patagioenas plumbea
Patagioenas subvinacea
Macropygia mackinlayi
Macropygia tenurirostris
Reinwardtoena browni
Treron_calva(mainland)
Treron_calva_B1
Treron_calva_P11
Treron_calva_B2
Treron_calva_P9
Treron_calva_P10
Treron_sanctithomae_3
Treron_sanctithomae_1
Geotrygon montana
Zenaida asiatica
Zenaida macroura
Leptotila rufaxilla
Leptotila verreauxi
Pterocles namaqua
CYTB: ML
Aplopelia_larvata_P1
Aplopelia_larvata_P26
Aplopelia_larvata_ST14
Aplopelia_larvata_ST19
FMNH 444011
YPM 88245
YPM 50221
Columba_malherbii_P4
Columba_malherbii_P5
Columba_malherbii_ST7
Columba_malherbii_ST5
Columba_thomensis_1
Columba_thomensis_2
Columba arquatrix
Treron_calva_B1
Treron_calva_B2
Treron_calva_P10
Treron_sanctithomae_3
Treron_sanctithomae_1
Treron_calva_P9
Treron_calva_P11
Treron_calva(mainland)
Columba pulchrichollis
Patagieonas fasciata
Columba leucocephala
Patagioenas oenops
Patagioenas speciosa
Patagioenas plumbea
Patagioenas subvinacea
Macropygia mackinlayi
Macropygia tenurirostris
Reinwardtoena browni
Columba guinea
Columba livia
Columba rupestris
Columba palumbus
Geotrygon montana
Leptotila rufaxilla
Leptotila verreauxi
Nesoenas mayeri
Streptopelia picturata
Streptopelia bitorquata
Streptopelia capicola
Streptopelia vinacea
Streoptopelia decaocto
Streptopelia roseogrisea
Streptopelia decipiens
Streptopelia hypopyrrha
Streptopelia turtur
Streptopelia orientalis
Streptopelia semitorquata
Streptopelia tranquebarica
Streptopelia chinensis
Streptopelia senegalensis
Zenaida asiatica
Zenaida macroura
Pterocles namaqua

Appendix 6.3

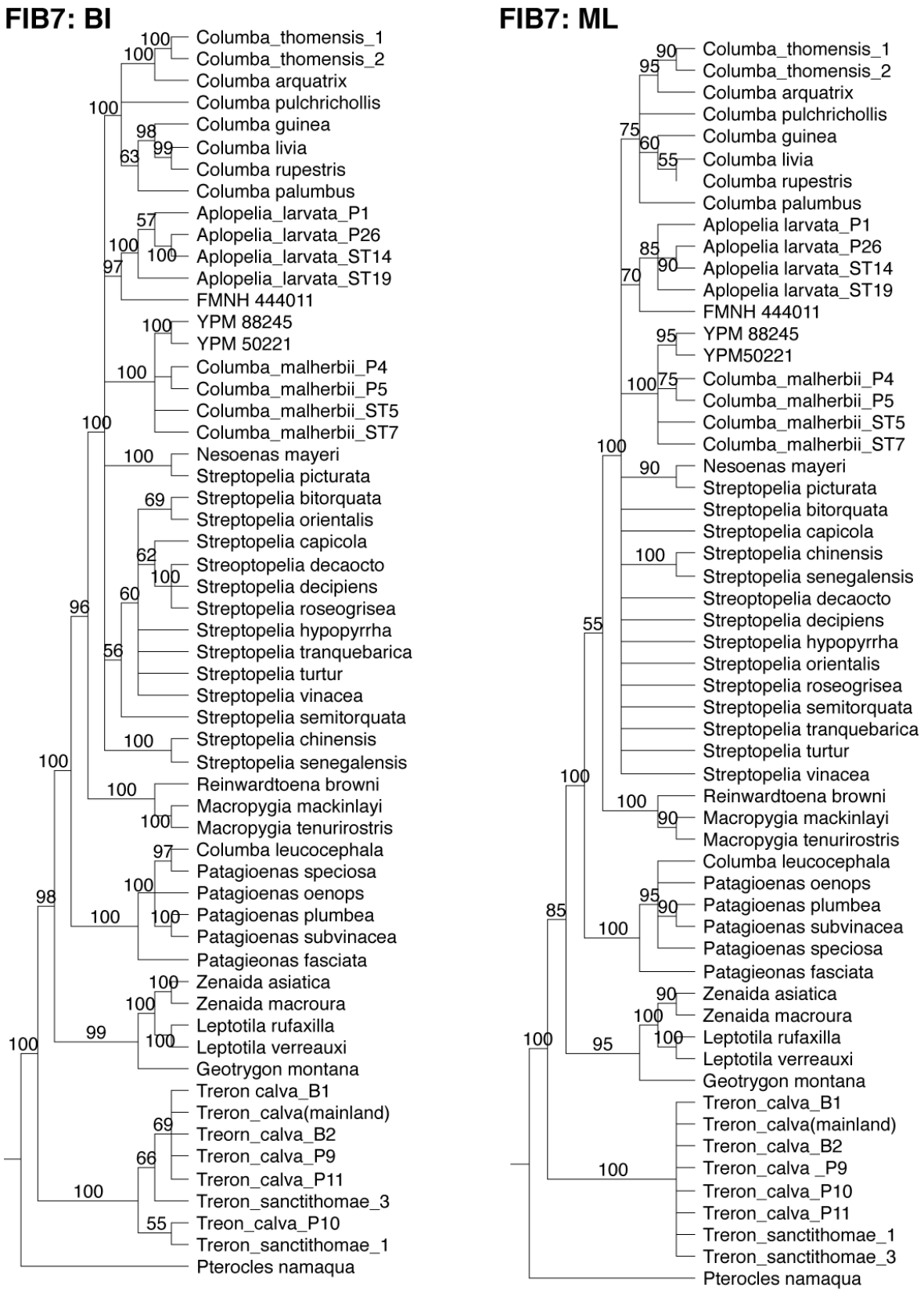
COI: BI



COI: ML

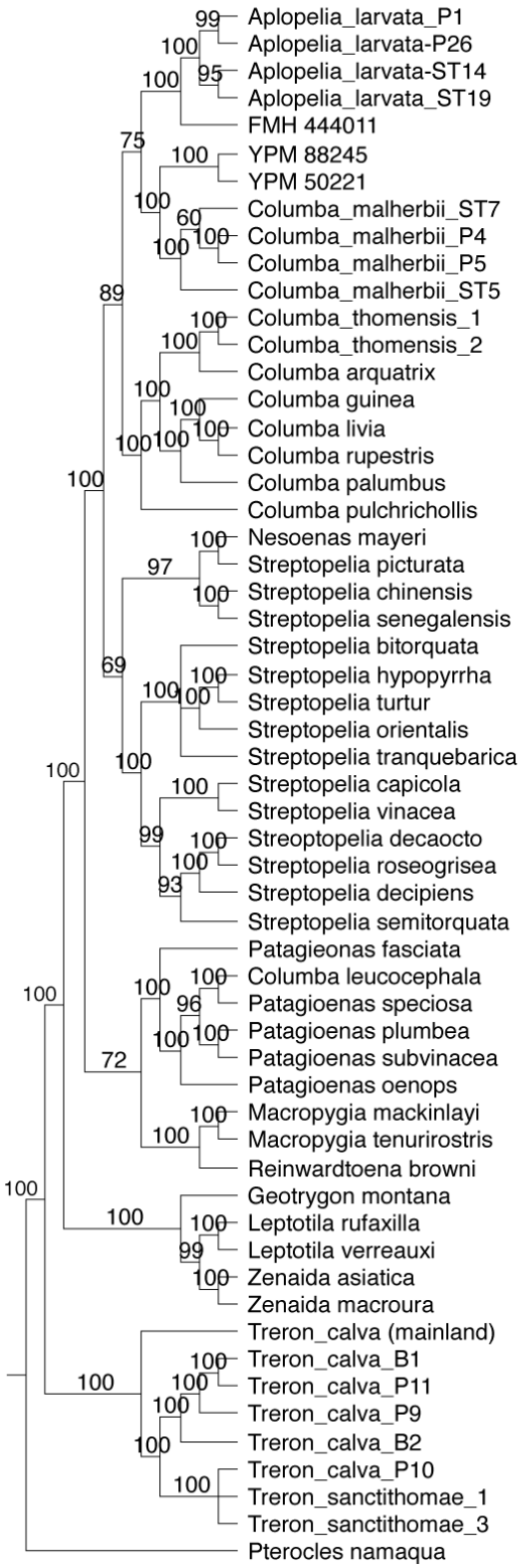


Appendix 6.4

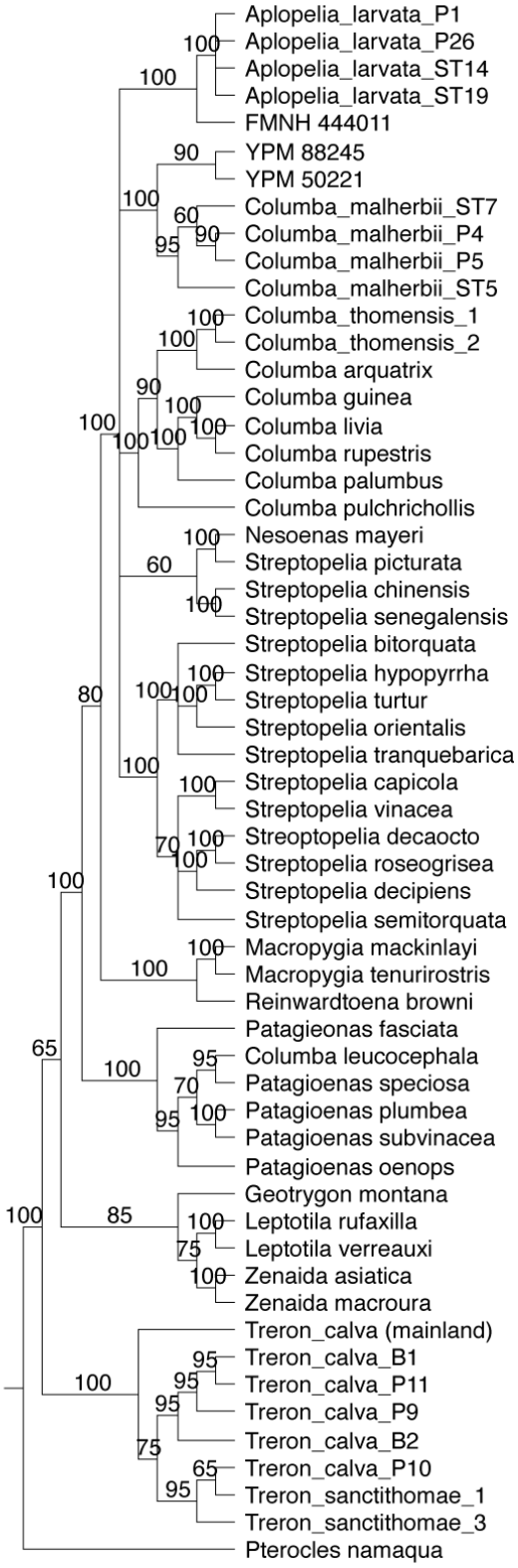

FIB7: BI
Columba_thomensis_1
Columba_thomensis_2
Columba arquatrix
Columba pulchrichollis
Columba guinea
Columba livia
Columba rupestris
Columba palumbus
Aplopelia_larvata_P1
Aplopelia_larvata_P26
Aplopelia_larvata_ST14
Aplopelia_larvata_ST19
FMNH 444011
YPM 88245
YPM 50221
Columba_malherbii_P4
Columba_malherbii_P5
Columba_malherbii_ST5
Columba_malherbii_ST7
Nesoenas mayeri
Streptopelia picturata
Streptopelia bitorquata
Streptopelia orientalis
Streptopelia capicola
Streoptopelia decaocto
Streptopelia decipiens
Streptopelia roseogrisea
Streptopelia hypopyrrha
Streptopelia tranquebarica
Streptopelia turtur
Streptopelia vinacea
Streptopelia semitorquata
Streptopelia chinensis
Streptopelia senegalensis
Reinwardtoena browni
Macropygia mackinlayi
Macropygia tenurirostris
Columba leucocephala
Patagioenas speciosa
Patagioenas oenops
Patagioenas plumbea
Patagioenas subvinacea
Patagieonas fasciata
Zenaida asiatica
Zenaida macroura
Leptotila rufaxilla
Leptotila verreauxi
Geotrygon montana
Treron calva_B1
Treron_calva(mainland)
Treorn_calva_B2
Treron_calva_P9
Treron_calva_P11
Treron_sanctithomae_3
Treon_calva_P10
Treron_sanctithomae_1
Pterocles namaqua
FIB7: ML
Columba_thomensis_1
Columba_thomensis_2
Columba arquatrix
Columba pulchrichollis
Columba guinea
Columba livia
Columba rupestris
Columba palumbus
Aplopelia larvata_P1
Aplopelia larvata_P26
Aplopelia larvata_ST14
Aplopelia larvata_ST19
FMNH 444011
YPM 88245
YPM50221
Columba_malherbii_P4
Columba_malherbii_P5
Columba_malherbii_ST5
Columba_malherbii_ST7
Nesoenas mayeri
Streptopelia picturata
Streptopelia bitorquata
Streptopelia capicola
Streptopelia chinensis
Streptopelia senegalensis
Streoptopelia decaocto
Streptopelia decipiens
Streptopelia hypopyrrha
Streptopelia orientalis
Streptopelia roseogrisea
Streptopelia semitorquata
Streptopelia tranquebarica
Streptopelia turtur
Streptopelia vinacea
Reinwardtoena browni
Macropygia mackinlayi
Macropygia tenurirostris
Columba leucocephala
Patagioenas oenops
Patagioenas plumbea
Patagioenas subvinacea
Patagioenas speciosa
Patagieonas fasciata
Zenaida asiatica
Zenaida macroura
Leptotila rufaxilla
Leptotila verreauxi
Geotrygon montana
Treron_calva_B1
Treron_calva(mainland)
Treron_calva_B2
Treron_calva _P9
Treron_calva_P10
Treron_calva_P11
Treron_sanctithomae_1
Treron_sanctithomae_3
Pterocles namaqua

Appendix 6.5

mtDNA + FIB7: BI



mtDNA + FIB7: ML



Appendix 7. Primers and PCR conditions for the markers used in chapter 2.

					PCR-mix				PCR-cycle					
Marker	**Primer name[1]**	**Primer sequence 5'-3'**	**Product (bp)**	**Source**	**Primer (µM)**	**Mg (mM)**	**dNTP (mM)**	**Taq (U)**	**[PREMIT]**	**[DENAT**	**ANNEAL**	**EXT]**	**CYCLES**	**FINAL**
ND2-1	L5219	**F:** CCCATACCCCGAAAATGATG	547	1										
	H5766	**R:** GGATGAGAAGGCTAGGATTTTKCG		1										
ND2-2	L5758	**F:** GGCTGAATRGGMCTNAAYCARAC	555	1	0.2	2	0.2	1.6	95º-15'	95º-30''	58/55º-30''	72º-90''	11+26	72º-10'
	H6313	**R:** CTCTTATTTAAGGCTTTGAAGGC		1										
FIB5	Fib5	**F:** CGCCATACAGAGTATACTGTGACAT	576	2	0.2	2	0.2	1.6	95º-15'	95º-30''	56º-30''	72º-75''	37	72º-10'
	Fib6	**R:** GCCATCCTGGCGATTCTGAA		2										
CHD1	CDH1R	**F:** TATCGTCAGTTTCCTTTTCAGGT	-	3	0.2	2	0.2	0.6	95º-15'	95º-30''	57º-30''	72º-90''	37	72º-10'
(sexing)	CHH1F	**R:** CCT TTTATTGATCCATCAAGCCT												

[1]:Primer name given as the position of the 3' end base on the chicken mtDNA genome (Desjardins & Morais 1990) - L: light strand, H: heavy strand; 2:Primer name as in source - F: forward, R: reverse.
PREMIT: Initial denaturation; DENAT: Denaturation; ANNEAL: Annealing; EXT: Extension; FINAL: Final extension.
SOURCE: 1 - Sorenson *et al.* (1999); 2 - Bowie & Fjeldså 2005; 3 - Chun-I *et al.* 2010.

Appendix 8. Sampling details for *Columba malherbii* individuals obtained for chapter 2. Country: STP (São Tomé and Príncipe),

Species	Code	Sex	Country	Locality	Sample
Columba malherbii	CMst2	F	STP- São Tomé	Alto Douro	Blood
	CMst3	M	STP- São Tomé	Quija	Feathers
	CMst4	F	STP- São Tomé	S. Miguel 2	Blood
	CMst5	M	STP- São Tomé	Água Izé	Tissue (Heart)
	CMst6	M	STP- São Tomé	Água Izé	Tissue (Heart)
	CMst7	M	STP- São Tomé	Água Izé	Tissue (Heart)
	CMst15	F	STP- São Tomé	Água Izé	Blood
	CMst16	M	STP- São Tomé	Monte Estoril	Tissue (Toes)
	CMst17	M	STP- São Tomé	Monte Estoril	Tissue (Toes)
	CMst32	F	STP- São Tomé	Belém	Tissue (Toes)
	CMst33	F	STP- São Tomé	Margão (Bate Pá)	Tissue (Toes)
	CMst38	M	STP- São Tomé	Margão (Bate Pá)	Tissue (Toes)
	CMst55	F	STP- São Tomé	Pedrómo	Tissue (Toes)
	CMst56	M	STP- São Tomé	Pedrómo	Tissue (Toes)
	CMst86	M	STP- São Tomé	Santa Catarina	Tissue (Toes)
	CMst87	M	STP- São Tomé	Santa Catarina	Tissue (Toes)
	CMst95	M	STP- São Tomé	Maia	Tissue (Toes)
	CMst96	F	STP- São Tomé	Maia	Tissue (Toes)
	CMst98	F	STP- São Tomé	Muro da Trindade	Tissue (Toes)
	CMst99	M	STP- São Tomé	Muro da Trindade	Tissue (Toes)
	CMst113	M	STP- São Tomé	Apolónia	Tissue (Toes)
	CMst114	F	STP- São Tomé	Apolónia	Tissue (Toes)
	CMst121	M	STP- São Tomé	São João Angolares	Tissue (Toes)
	CMst122	M	STP- São Tomé	São João Angolares	Tissue (Toes)
	CMst129	M	STP- São Tomé	Folha Fede	Tissue (Toes)
	CMst130	M	STP- São Tomé	Folha Fede	Tissue (Toes)
	CMst152	F	STP- São Tomé	Morro Trindade	Tissue (Toes)
	CMst153	F	STP- São Tomé	Morro Trindade	Tissue (Toes)
	CMp4	M	STP- Príncipe	Ribeira Izé	Tissue (Muscle)
	CMp5	M	STP- Príncipe	Ribeira Izé	Tissue (Muscle)
	CMp6	F	STP- Príncipe	Azeitona (Bairro)	Tissue (Muscle)
	CMp7	M	STP- Príncipe	Alto Camarão	Tissue (Muscle)
	CMp8	M	STP- Príncipe	Nova Cuba	Tissue (Muscle)
	CMp9	M	STP- Príncipe	Nova Cuba	Tissue (Muscle)
	CMp10	F	STP- Príncipe	Nova Cuba	Tissue (Muscle)
	CMp11	M	STP- Príncipe	Nova Cuba	Tissue (Muscle)
	CMp12	M	STP- Príncipe	Nova Cuba	Tissue (Muscle)

Appendix 9. Sampling details for *Columba thomensis* individuals obtained for chapter 2. Country: STP (São Tomé and Príncipe),

Species	Code	Sex	Country	Locality	Sample
Columba thomensis	CT1	M	STP- São Tomé	Lagoa Amélia	Blood
	CT2	F	STP- São Tomé	Bom Sucesso	Tissue (Heart)
	CT3	M	STP- São Tomé	Bom Sucesso	Tissue (Heart)
	CT4	M	STP- São Tomé	Maia	Tissue (Toes)
	CT5	M	STP- São Tomé	São Manuel (Santa Catarina)	Tissue (Toes)
	CT6	M	STP- São Tomé	São Manuel (Santa Catarina)	Tissue (Toes)
	CT7	M	STP- São Tomé	São Manuel (Santa Catarina)	Tissue (Toes)
	CT8	M	STP- São Tomé	São Manuel (Santa Catarina)	Tissue (Toes)
	CT9	M	STP- São Tomé	São Manuel (Santa Catarina)	Tissue (Toes)
	CT10	M	STP- São Tomé	Ponta Furada	Tissue (Toes)
	CT11	M	STP- São Tomé	Ponta Furada	Tissue (Toes)
	CT12	M	STP- São Tomé	Ponta Furada	Tissue (Toes)
	CT13	F	STP- São Tomé	Ponta Furada	Tissue (Toes)
	CT14	F	STP- São Tomé	Macambrará	Tissue (Toes)
	CT15	F	STP- São Tomé	Macambrará	Tissue (Toes)
	CT16	F	STP- São Tomé	Macambrará	Tissue (Toes)
	CT17	F	STP- São Tomé	Macambrará	Tissue (Toes)
	CT18	M	STP- São Tomé	Macambrará	Tissue (Toes)
	CT19	F	STP- São Tomé	Macambrará	Tissue (Toes)
	CT20	M	STP- São Tomé	Macambrará	Tissue (Toes)
	CT21	M	STP- São Tomé	Macambrará	Tissue (Toes)
	CT22	M	STP- São Tomé	Macambrará	Tissue (Toes)
	CT23	M	STP- São Tomé	Apolónia	Tissue (Toes)
	CT24	F	STP- São Tomé	Apolónia	Tissue (Toes)
	CT25	F	STP- São Tomé	Apolónia	Tissue (Toes)
	CT26	M	STP- São Tomé	Apolónia	Tissue (Toes)
	CT27	F	STP- São Tomé	Apolónia	Tissue (Toes)
	CT28	F	STP- São Tomé	Apolónia	Tissue (Toes)
	CT29	F	STP- São Tomé	Apolónia	Tissue (Toes)
	CT33	M	STP- São Tomé	Ponta Furada	Tissue (Toes)

Appendix 10. Sampling details for *Treron* individuals obtained for this study. Country: STP (São Tomé and Príncipe), EG (Equatorial Guinea).

Species	Code	Sex	Country	Locality	Sample
Treron sanctithomae	TS1	M	STP- São Tomé	Alto Douro	Blood
	TS2	F	STP- São Tomé	Lagoa Amélia	Feathers
	TS3	M	STP- São Tomé	Lagoa Amélia	Tissue (Toes)
	TS4	F	STP- São Tomé	Lagoa Amélia	Tissue (Toes)
	TS5	F	STP- São Tomé	Lagoa Amélia	Tissue (Toes)
	TS31	M	STP- São Tomé	Campo Grande (Bemposta)	Tissue (Toes)
	TS32	M	STP- São Tomé	Campo Grande (Bemposta)	Tissue (Toes)
	TS33	F	STP- São Tomé	Campo Grande (Bemposta)	Tissue (Toes)
	TS61	M	STP- São Tomé	Santa Catarina	Tissue (Toes)
	TS62	M	STP- São Tomé	Santa Catarina	Tissue (Toes)
	TS63	F	STP- São Tomé	Santa Catarina	Tissue (Toes)
	TS68	M	STP- São Tomé	Maia	Tissue (Toes)
	TS70	F	STP- São Tomé	Maia	Tissue (Toes)
	TS97	F	STP- São Tomé	Macambrará	Tissue (Toes)
	TS98	M	STP- São Tomé	Macambrará	Tissue (Toes)
	TS99	M	STP- São Tomé	Macambrará	Tissue (Toes)
	TS118	F	STP- São Tomé	Apolónia	Tissue (Toes)
	TS119	M	STP- São Tomé	Apolónia	Tissue (Toes)
	TS120	M	STP- São Tomé	Apolónia	Tissue (Toes)
	TS124	F	STP- São Tomé	São João Angolares	Tissue (Toes)
	TS125	M	STP- São Tomé	São João Angolares	Tissue (Toes)
	TS126	M	STP- São Tomé	São João Angolares	Tissue (Toes)
	TS127	F	STP- São Tomé	São João Angolares	Tissue (Toes)
	TS128	M	STP- São Tomé	São João Angolares	Tissue (Toes)
	TS129	M	STP- São Tomé	São João Angolares	Tissue (Toes)
	TS141	F	STP- São Tomé	São Nicolau	Tissue (Toes)
	TS142	M	STP- São Tomé	São Nicolau	Tissue (Toes)
	TS143	F	STP- São Tomé	São Nicolau	Tissue (Toes)
	TS152	F	STP- São Tomé	Colónia Açoreana	Tissue (Toes)
	TS153	M	STP- São Tomé	Colónia Açoreana	Tissue (Toes)
	TS154	F	STP- São Tomé	Colónia Açoreana	Tissue (Toes)
	TS155	M	STP- São Tomé	Colónia Açoreana	Tissue (Toes)
	TS156	F	STP- São Tomé	Colónia Açoreana	Tissue (Toes)
Treron calva	TCb1	F	EG-Bioko	Moka	Blood
	TCb2	F	EG-Bioko	Moka	Tissue (Heart)
	TCp9	F	STP-Príncipe	Ribeira Izé	Tissue (Muscle)
	TCp10	M	STP-Príncipe	Ribeira Izé - Top	Tissue (Muscle)
	TCp11	M	STP-Príncipe	Ribeira Izé - Top	Tissue (Muscle)
	TCp12	F	STP-Príncipe	Ribeira Izé - Top	Tissue (Muscle)
	TCp13	M	STP-Príncipe	Azeitona (Bairro)	Tissue (Muscle)
	TCp14	M	STP-Príncipe	Azeitona (Bairro)	Tissue (Muscle)
	TCp15	F	STP-Príncipe	Azeitona (Bairro)	Tissue (Muscle)
	TCp16	M	STP-Príncipe	São Joaquim	Tissue (Heart)
	TCp17	F	STP-Príncipe	O Que Daniel	Tissue (Muscle)
	TCp18	F	STP-Príncipe	O Que Daniel	Tissue (Muscle)
	TCp19	M	STP-Príncipe	Nova Cuba	Tissue (Heart)
	TCp20	F	STP-Príncipe	Nova Cuba	Tissue (Muscle)
	TCp21	M	STP-Príncipe	Nova Cuba	Tissue (Muscle)
	TCp22	F	STP-Príncipe	Nova Cuba	Tissue (Muscle)
	Treron_ calva_mainland	M	Angola	Kumbira, Conda	Tissue (Toes)

Appendix 11. Estimation of the best partition of the data and correspondent model of evolution used in chapter 2.

	Marker	Partition/DNA model	Log-likelihood
C. malherbii	**ND2**	[1st, 2nd, 3rd] HKY+Γ	-1807.24
	FIB5	HKY	-738.98
C. thomensis	**ND2**	[1st, 2nd, 3rd] HKY	-1348.56
	FIB5	HKY	-740.096
T. sanctithomae	**ND2**	[1st, 2nd, 3rd] TrN	-1330.63
	FIB5	HKY	749.72
Treron	**ND2**	[1st, 2nd] HKY [3rd] HKY	-2047.99
	FIB5	HKY	-764.84

Square brackets bound elements of a partition. 1st=first codon position;
2nd =second codon position; 3rd=third codon position